Internet Appendix for "Diagnostic Expectations and Stock Returns"

PEDRO BORDALO, NICOLA GENNAIOLI, RAFAEL LA PORTA, ANDREI SHLEIFER*

This Appendix consists of five Sections. Section I presents the proofs. Section II examines the robustness of the LLTG-HLTG return differential. Section III examines the kernel of truth of long-term growth forecasts by comparing the post-formation growth in earnings of HLTG and non-HLTG firms. Section IV expands on the Coibion and Gorodnichenko analysis presented in the main text, providing evidence on overreaction versus adaptive expectations, on long versus short horizon forecasts, and on the link between overreaction to news and the return spread between HLTG and LLTG portfolios. Finally, Section V extends our model estimation with several exercises.

I. Proofs

Proposition 1: Upon observing $g_{i,t} \equiv x_{i,t} - bx_{i,t-1}$, the analyst's believed distribution of firm fundamentals is given by

$$h^{\theta}(f, g_{i,t}) = h(f, g_{i,t}) \cdot [R(f, g_{i,t})]^{\theta} \cdot Z,$$

where $Z^{-1} = \int h(f, g_{i,t}) \cdot [R(f, g_{i,t})]^{\theta} \cdot df$ and

$$R(f, g_{i,t}) = \exp\left\{\frac{(\hat{f}_{i,t} - a\hat{f}_{i,t-1})(2f - a\hat{f}_{i,t-1} - \hat{f}_{i,t})}{2\sigma_f^2}\right\}.$$

* Citation format: Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta and Andrei Shleifer, Internet Appendix for "Diagnostic Expectations and Stock Returns," *Journal of Finance*, DOI: 10.1111/jofi.12833.

We expand the above expression using the assumption that $h(f, g_{i,t})$ is normally distributed with variance σ_f^2 and mean

$$\hat{f}_{i,t} = a\hat{f}_{i,t-1} + K(g_{i,t} - a\hat{f}_{i,t-1}).$$

We find that

$$h^{\theta}(f, g_{i,t}) = Z \cdot exp \frac{1}{2\sigma_f^2}\left\{-\left(f - \hat{f}_{i,t}\right)^2 + \theta(\hat{f}_{i,t} - a\hat{f}_{i,t-1})(2f - a\hat{f}_{i,t-1} - \hat{f}_{i,t})\right\}.$$

The exponent then reads

$$-\left(f - \hat{f}_{i,t}\right)^2 + \theta(\hat{f}_{i,t} - a\hat{f}_{i,t-1})(2f - a\hat{f}_{i,t-1} - \hat{f}_{i,t}) = -\left(f - \left(\hat{f}_{i,t} + \theta(\hat{f}_{i,t} - a\hat{f}_{i,t-1})\right)\right)^2 + c(\hat{f}_{i,t}, \hat{f}_{i,t-1}),$$

where $c(\hat{f}_{i,t}, \hat{f}_{i,t-1})$ is a constant (does not depend on f). Taking normalization into account, we then find that

$$h^{\theta}(f, g_{i,t}) = \frac{1}{\sqrt{2\pi\sigma_f^2}} e^{-\frac{\left(f - \left(\hat{f}_{i,t} + \theta(\hat{f}_{i,t} - a\hat{f}_{i,t-1})\right)\right)^2}{2\sigma_f^2}},$$

Using equation (3) for the Bayesian expectation $\hat{f}_{i,t}$, the mean of this distribution can be written:

$$\hat{f}_{i,t}^{\theta} = \hat{f}_{i,t} + \theta(\hat{f}_{i,t} - a\hat{f}_{i,t-1}) = a\hat{f}_{i,t-1} + K(1 + \theta)(g_{i,t} - a\hat{f}_{i,t-1}).$$

■

Proposition 2: Denote by $\lambda_H > 0$ the threshold in expected growth rate above which a firm is classified as HLTG (i.e., it is in the top decile). From the definition of LTG in Section IV, firm i is classified as HLTG at time t if

$$LTG_{i,t} = -\varphi_h x_{i,t} + \vartheta_h \hat{f}_{i,t}^{\theta} \geq \lambda_H,$$

where we have defined $\varphi_h \equiv (1 - b^h)$ and $\vartheta_h \equiv a^h \frac{1-(b/a)^h}{1-(b/a)}$. This expression can be written as

$$-\varphi_h x_{i,t} + \vartheta_h a(1 - K')\hat{f}_{i,t-1} + \vartheta_h K'(1 - b)x_{i,t-1} + \vartheta_h K'\left(x_{i,t} - x_{i,t-1}\right) \geq \lambda_H,$$

where $K' \equiv K(1 + \theta)$. The left-hand side of the above condition is a linear combination of mean-zero normally distributed random variables. Denote it by $LHS_{i,t}$. By linear regression, the average growth rate $x_{i,t} - x_{i,t-1}$ experienced by firms whose $LHS_{i,t}$ is equal to λ is given by

$$\mathbb{E}\left[x_{i,t} - x_{i,t-1} | LHS_{i,t} = \lambda\right] = \frac{cov\left(x_{i,t} - x_{i,t-1}, LHS_{i,t}\right)}{var\left(LHS_{i,t}\right)} \lambda.$$

Because for HLTG firms $\lambda \geq \lambda_H > 0$, their pre-formation growth is positive, $\mathbb{E}\left[x_{i,t} - x_{i,t-1} | LHS_{i,t} = \lambda\right] > 0$, if $cov\left(x_{i,t} - x_{i,t-1}, LHS_{i,t}\right) > 0$. This occurs when the expression

$$[K'\vartheta_h(1 + b) - \varphi_h]\left(var\left(x_{i,t}\right) - cov\left(x_{i,t}, x_{i,t-1}\right)\right) + \vartheta_h a(1 - K')cov\left(x_{i,t} - x_{i,t-1}, \hat{f}_{i,t-1}\right)$$

is positive. For convenience, rewrite this as

$$A\left[var\left(x_{i,t}\right) - cov\left(x_{i,t}, x_{i,t-1}\right)\right] + Bcov\left(x_{i,t} - x_{i,t-1}, \hat{f}_{i,t-1}\right).$$

It is useful to rewrite the first term as

$$A\left[(1 - b)var\left(x_{i,t}\right) - a \cdot cov\left(f_{i,t-1}, x_{i,t-1}\right)\right] = A\left[\frac{var(\epsilon)}{1 + b} + \frac{var\left(f_{i,t}\right)}{1 - ab}\frac{1 - a}{1 + b}\right],$$

where we use $var(x_{i,t}) = \frac{1}{1-b^2}\left[var(\epsilon) + \frac{1+ab}{1-ab}var(f_{i,t})\right]$. The second term reads

$$cov\left(x_{i,t} - x_{i,t-1}, a(1-K)\hat{f}_{i,t-2} + K(f_{i,t-1} + \epsilon_{i,t-1})\right) = a(1-K)cov(x_{i,t} - x_{i,t-1}, \hat{f}_{i,t-2}) -$$

$$(1-b)Kcov(x_{i,t-1}, f_{i,t-1}) - (1-b)Kvar(\epsilon) + aKvar(f_{i,t}).$$

We can show that

$$cov(x_{i,t} - x_{i,t-1}, \hat{f}_{i,t-2})$$

$$> a^2 cov(f_{i,t}, \hat{f}_{i,t-2}) - b(1-b)cov(x_{i,t-2}, f_{i,t-2}) - a(1-b)var(f_{i,t}),$$

where we use $cov(x_{i,t-1}, f_{i,t-2}) > cov(x_{i,t-1}, \hat{f}_{i,t-2})$. Thus,

$$cov\left(x_{i,t} - x_{i,t-1}, a(1-K)\hat{f}_{i,t-2} + K(f_{i,t-1} + \epsilon_{i,t-1})\right) > var(f_{i,t})\left[\frac{aK}{1-a^2(1-K)} -\right.$$

$$\left.(1-b)\left(K + a^2(1-K) + \frac{ab}{1-ab}\right)\right] - (1-b)Kvar(\epsilon).$$

Putting the two terms together, we find that

$$var(\epsilon)\left[\frac{A}{1+b} - (1-b)KB\right]$$

$$+ var(f_{i,t})\left[\frac{A}{1+b}\frac{1-a}{1-ab}\right.$$

$$\left.+ B\left[\frac{aK}{1-a^2(1-K)} - (1-b)\left(K + a^2(1-K) + \frac{ab}{1-ab}\right)\right]\right].$$

A sufficient condition that makes both terms positive is

$$b^h + \vartheta_h\left[\frac{K'-a}{1-a}\right] > 1.$$

This condition is easier to satisfy for low mean reversion (large b), large signal-to-noise ratio (large K), and strong overreaction (large θ). It is trivially satisfied when $b = 1$ if $K(1+\theta) \geq a$ (which holds in our estimation). ■

Proposition 3: From the law of motion of earnings we have that

$$\mathbb{E}\left(x_{i,t+h} - x_{i,t} | HLTG_t\right) = \mathbb{E}\left(-\varphi_h x_{i,t} + \vartheta_h f_{i,t} | HLTG_t\right).$$

Because rational estimation errors $u_{i,t} \equiv f_{i,t} - \hat{f}_{i,t}$ are zero on average, we also have that

$$\mathbb{E}\left(x_{i,t+h} - x_{i,t} | HLTG_t\right) = \mathbb{E}\left(-\varphi_h x_{i,t} + \vartheta_h \hat{f}_{i,t} | HLTG_t\right).$$

This implies that the average forecast error entailed in LTG is equal to

$$\mathbb{E}\left(x_{i,t+h} - x_{i,t} - LTG_{i,t} | HLTG_t\right) = \vartheta_h \mathbb{E}\left(\hat{f}_{i,t} - \hat{f}_{i,t}^{\theta} | HLTG_t\right) =$$

$$-\vartheta_h K\theta \mathbb{E}\left(x_{i,t} - b x_{i,t-1} - \; a\hat{f}_{i,t-1} | HLTG_t\right) = -\vartheta_h K\theta \mathbb{E}\left(\eta_{i,t} + \epsilon_{i,t} \; | HLTG_t\right).$$

The expectation $\mathbb{E}\left(\eta_{i,t} + \epsilon_{i,t} \; | HLTG_t\right)$ is positive because HLTG firms have positive recent performance (see Lemma 1). Under rationality, $\theta = 0$, forecast errors are unpredictable. Under diagnostic expectations, $\theta > 0$, forecast errors are predictably negative for the HLTG group. Conversely, the same argument shows that they are predictably positive for the LLTG group. ■

Proposition 4: The average LTG at future date $t+s, s \geq 1$, in the HLTG group is equal to

$$\mathbb{E}\left(LTG_{i,t+s} | HLTG_t\right) = \mathbb{E}\left(-\varphi_h x_{i,t+s} + \vartheta_h \hat{f}_{i,t+s}^{\theta} | HLTG_t\right) =$$

$$\mathbb{E}\left(-\varphi_h x_{i,t+s} + \vartheta_h \hat{f}_{i,t+s} + \vartheta_h \left(\hat{f}_{i,t+s}^{\theta} - \hat{f}_{i,t+s}\right) | HLTG_t\right) =$$

$$\mathbb{E}\big(-\varphi_h x_{i,t+s} + \vartheta_h \hat{f}_{i,t+s} + \vartheta_h K\theta\big(g_{i,t+s} - a\hat{f}_{i,t+s-1}\big)|HLTG_t\big) =$$

$$-\varphi_h b^s x_{i,t} + \vartheta_h a^s \hat{f}_{i,t},$$

where the last equality follows from the fact that within the HLTG group of stocks, $g_{i,t+s} - a\hat{f}_{i,t+s-1}$ is zero on average. This implies that within HLTG stocks, future $LTG_{i,t+s}$ mean reverts on average, as implied by the power terms b^s and a^s. This occurs regardless of whether expectations are rational or diagnostic because $-\varphi_h b^s x_{i,t} + \vartheta_h a^s \hat{f}_{i,t}$ does not depend on θ. Between the formation date t and date $t+1$, however, mean reversion is stronger under diagnostic expectations. In fact, the condition

$$\mathbb{E}\big(g_{i,t+s} - a\hat{f}_{i,t+s-1}|HLTG_t\big) > 0$$

holds if and only if

$$b^h + \vartheta_h K' > 1.$$

This condition is implied by the assumption of Proposition 2 (equation (7)), provided $K' > 1$, which holds in the estimation. We then find

$$\mathbb{E}\big(LTG_{i,t+1}|HLTG_t\big) - \mathbb{E}\big(LTG_{i,t}|HLTG_t\big) = \mathbb{E}\big(LTG_{i,t+1}^{\theta=0}|HLTG_t\big) - \mathbb{E}\big(LTG_{i,t}^{\theta=0}|HLTG_t\big) - \Psi\theta,$$

where $\Psi = \vartheta_h K\mathbb{E}\big(g_{i,t+s} - a\hat{f}_{i,t+s-1}\big) > 0$ (and $\theta > 0$). This is because under diagnostic expectations, the average $LTG_{i,t}$ in the HLTG group is inflated relative to the rational benchmark. The converse holds for stocks in the LLTG group at time t. ■

Proposition 5: The realized return at time t on HLTG stocks is equal to the average

$$\mathbb{E}\left(\frac{P_{i,t}+D_{it}}{P_{i,t-1}}|HLTG_t\right)=R+\mathbb{E}\left(\frac{P_{i,t}-\mathbb{E}^{\theta}_{i,t-1}(P_{it})+D_{i,t}-\mathbb{E}^{\theta}_{t-1}(D_{i,t})}{P_{i,t-1}}|HLTG_t\right).$$

An individual stock i in the HLTG portfolio therefore experiences positive abnormal returns pre formation if

$$P_{i,t}-\mathbb{E}^{\theta}_{t-1}(P_{i,t})+D_{i,t}-\mathbb{E}^{\theta}_{t-1}(D_{i,t})>0.$$

Consider the first term $P_{i,t}-\mathbb{E}^{\theta}_{t-1}(P_{i,t})$. Because prices are equal to discounted future dividends,

$$P_{i,t}-\mathbb{E}^{\theta}_{t-1}(P_{i,t})=\sum_{s\geq 1}\frac{\mathbb{E}^{\theta}_{t}(D_{i,t+s})-\mathbb{E}^{\theta}_{t-1}(D_{i,t+s})}{R^s},$$

which implies that abnormal returns are induced by an upward revision $\mathbb{E}^{\theta}_{t}(D_{i,t+s})-\mathbb{E}^{\theta}_{t-1}(D_{i,t+s})>0$ of investors' beliefs of future dividends. Using previous notation, we have that

$$\mathbb{E}^{\theta}_{t}(D_{i,t+s})=\mathbb{E}^{\theta}_{t}(e^{x_{i,t+s}})=e^{b^s x_{i,t}+\vartheta_s \hat{f}^{\theta}_{i,t}+\frac{1}{2}Var_t(x_{t+s})}.$$

As a result, we have that $\mathbb{E}^{\theta}_{t}(D_{i,t+s})-\mathbb{E}^{\theta}_{t-1}(D_{i,t+s})>0$ if on average in HLTG

$$b^s x_{i,t}+\vartheta_s \hat{f}^{\theta}_{i,t}>b^s\mathbb{E}^{\theta}_{t-1}(x_{i,t})+\vartheta_s\mathbb{E}^{\theta}_{t-1}(f_{i,t}).$$

We thus have

$$\begin{aligned}
&b^s\left(x_{i,t}-\mathbb{E}^{\theta}_{t-1}(x_{i,t})\right)+\vartheta_s\left(\hat{f}^{\theta}_{i,t}-\mathbb{E}^{\theta}_{t-1}(\hat{f}^{\theta}_{i,t})\right)\\
&\qquad=b^s(S_{i,t}-aK\theta S_{i,t-1})+\vartheta_s\big(K(1+\theta)S_{i,t}-aK\theta S_{i,t-1}\big)\\
&\qquad=S_{i,t}(b^s+\vartheta_s K)+\theta K(\vartheta_s S_{i,t}-\vartheta_{s+1}S_{i,t-1}),
\end{aligned}$$

where $S_t=x_{i,t}-bx_{i,t-1}-a\hat{f}_{i,t-1}$ is the news, or surprise, at time t. Intuitively, this suggests that high returns at t are associated with surprises $S_{i,t}$ that are not only positive, but also (for $\theta>$

0) sufficiently large compared to surprises in the previous period. Rewriting the above as $AS_{i,t} - BAS_{i,t-1}$ with $A = b^s + \vartheta_s K(1+\theta)$ and $B = \theta K \vartheta_{s+1}$, we then have

$$\mathbb{E}\left[AS_{i,t} - BAS_{i,t-1} | LTG_{i,t} = \lambda\right] = \frac{cov\left(AS_{i,t} - BAS_{i,t-1}, LTG_{i,t}\right)}{var\left(LTG_{i,t}\right)} \lambda.$$

We can write $LTG_{i,t}$ as

$$LTG_{i,t} = [\vartheta_s K(1+\theta) - \varphi_s]S_{i,t} - [\varphi_s(b + aK) - a\vartheta_s K]S_{i,t-1} + terms\ at\ t-2.$$

Because surprises in period t are uncorrelated with information at different periods, the numerator of the expectation above then reads

$$\begin{aligned} &cov\left(AS_{i,t} - BAS_{i,t-1}, LTG_{i,t}\right) \\ &\quad = A[\vartheta_s K(1+\theta) - \varphi_s]var\left(S_{i,t}\right) - B[\varphi_s(b + aK) - a\vartheta_s K]var\left(S_{i,t-1}\right) \\ &\quad = var\left(S_{i,t}\right)\left[A[\vartheta_s K(1+\theta) - \varphi_s] - B[\varphi_s(b + aK) - a\vartheta_s K]\right]. \end{aligned}$$

his is positive provided that

$$\vartheta_s K(1+\theta) + b^s > 1 + 2\theta K \vartheta_{s+1}[\varphi_s(b + aK) - a\vartheta_s K]$$

(which holds in our estimation). The assumption of Proposition 2 guarantees $\vartheta_s K(1+\theta) + b^s > 1$. Thus, under rational expectations, $\theta = 0$, the condition holds trivially. For $\theta > 0$, a sufficient condition for the above to hold is that $\vartheta_s K > \varphi_s\left(\frac{b}{a} + K\right)$. Under this condition, $\mathbb{E}\left[x_{i,t} + LTG_{i,t} - x_{i,t-1} - LTG_{i,t-1} | LTG_{i,t} \geq \lambda_h\right]$ is positive and the result follows. Note that this is implied by the assumption of Proposition 2 provided that $\frac{b}{a} + K < \frac{1}{1+\theta}$. ■

Proposition 6: As shown in equation (9), a stock's average return going forward into the next period is equal to

$$\frac{\mathbb{E}_t\big(P_{i,t+1}+D_{i,t+1}\big)}{\mathbb{E}_t^{\theta}\big(P_{i,t+1}+D_{i,t+1}\big)}R.$$

Note that $D_{i,t+s}=e^{x_{i,t+s}}$ and, given that price is the discounted sum of future dividends,

$$P_{i,t+1}+D_{i,t+1}=\sum\nolimits_{s\geq 0}\frac{D_{i,t+1+s}}{R^s}=\sum\nolimits_{s\geq 0}e^{x_{i,t+1+s}-slnR},$$

where, as usual, we assume that lnR is large enough that the sum converges. Given lognormality, we have that:

$$\mathbb{E}_t^{\theta}\big(P_{i,t+1}+D_{i,t+1}\big)=\sum\nolimits_{s\geq 0}e^{\mathbb{E}_t^{\theta}(x_{i,t+1+s})-slnR+\frac{1}{2}Var_t(x_{i,t+1+s})},$$

where rational expectations correspond to the special case of $\theta=0$. For $\theta=0$, then, the numerator and the denominator of equation (9) are equal, so that the average realized return is equal to the realized return R for all firms. As a result, the average realized post-formation return of the HLTG and LLTG portfolios should be equal to the required return R.

To see the role of diagnostic expectations, note that θ influences only the expected log dividend $\mathbb{E}_t^{\theta}\big(x_{i,t+1+s}\big)$, but not the perceived variance $Var_t\big(x_{i,t+1+s}\big)$. In particular

$$\mathbb{E}_t^{\theta}\big(x_{i,t+s+1}\big)=b^{s+1}x_t+a^{s+1}\frac{1-(b/a)^{s+1}}{1-(b/a)}\big[\hat{f}_{i,t}+K\theta\big(g_{i,t}-a\hat{f}_{i,t-1}\big)\big].$$

This implies that

$$\frac{\partial \mathbb{E}_t^\theta\left(P_{i,t+1}+D_{i,t+1}\right)}{\partial \theta}$$

$$= K\theta\left(g_{i,t}-a\hat{f}_{i,t-1}\right)\sum_{s\geq 0} a^{s+1}\frac{1-(b/a)^{s+1}}{1-(b/a)} e^{\mathbb{E}_t^\theta(x_{i,t+1+s})-s\ln R+\frac{1}{2}Var_t(x_{i,t+1+s})}.$$

Under diagnostic expectations, pre-formation news drives mispricing. HLTG stocks experience positive surprises before formation, that is, $\left(g_{i,t}-a\hat{f}_{i,t-1}\right)>0$. As a result, the diagnostic expectation $\mathbb{E}_t^\theta\left(P_{i,t+1}+D_{i,t+1}\right)$ is above the rational counterpart, so that realized post-formation returns are on average below the required return R. For LLTG the opposite is true.■

Proposition 7: Regressing $x_{i,t+h}-x_{i,t}-LTG_{i,t}$ on $LTG_{i,t}-LTG_{i,t-k}$ yields the coefficient

$$\beta=\frac{cov\left(x_{i,t+h}-x_{i,t}-LTG_{i,t}, LTG_{i,t}-LTG_{i,t-k}\right)}{var\left(LTG_{i,t}-LTG_{i,t-k}\right)}.$$

The forecast error in the denominator reads

$$x_{i,t+h}-x_{i,t}-\mathbb{E}_t\left(x_{i,t+h}-x_{i,t}\right)-\theta\vartheta_h K\left(x_{i,t}-bx_{i,t-1}-a\hat{f}_{i,t-1}\right).$$

The first two terms include only shocks after t and do not co-vary with any quantity at t. The last term arises only for $\theta>0$, and captures the overreaction to news at t embedded in $LTG_{i,t}$. We thus have

$$\beta \propto -\theta\cdot cov\left(x_{i,t}-bx_{i,t-1}-a\hat{f}_{i,t-1}, LTG_{i,t}-LTG_{i,t-k}\right).$$

Intuitively, a positive covariance means that positive surprises at t tend to be associated with upward revisions in LTG. The second argument reads

$$-\varphi_h\left(x_{i,t}-x_{i,t-1}\right)+\vartheta_h a(1-K')\left(\hat{f}_{i,t-1}-\hat{f}_{i,t-2}\right)+\vartheta_h K'^{\left(x_{i,t}-bx_{i,t-1}-a\hat{f}_{i,t-1}\right)}$$
$$-\vartheta_h K'^{\left(x_{i,t-1}-bx_{i,t-2}-a\hat{f}_{i,t-2}\right)}.$$

The surprise at t does not covary with either the update in beliefs at $t-1$ (second term) or the surprise at $t-1$ (last term), so these drop out. Write the first term as

$$-\varphi_h\left(x_{i,t}-x_{i,t-1}\right)=-\varphi_h\left(x_{i,t}-bx_{i,t-1}-a\hat{f}_{i,t-1}\right)-\varphi_h\left((1-b)x_{i,t-1}-a\hat{f}_{i,t-1}\right).$$

Again, because surprises at t are not predictable from information at $t-1$, the second term drops out. We therefore get $(\vartheta_h K'-\varphi_h)var\left(x_{i,t}-bx_{i,t-1}-a\hat{f}_{i,t-1}\right)$, which is positive if and only if

$$b^h+\vartheta_h K(1+\theta)>1.$$

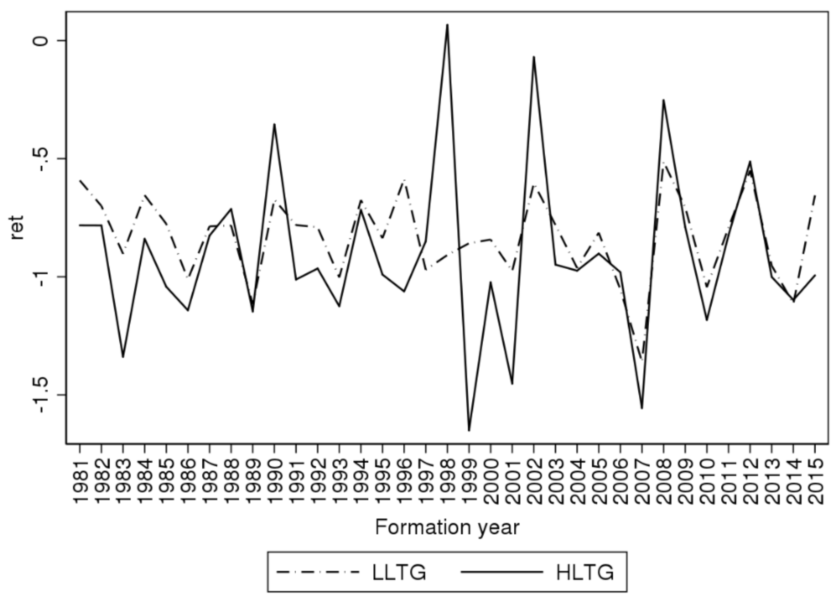
This condition is weaker than that of Proposition 2 provided $K(1+\theta)>1$, which holds in our estimation. ■

II. Robustness of LLTG-HLTG Return Differential

In this section we examine the robustness of the LLTG-HLTG return differential from a number of perspectives. We first examine the performance of LTG portfolios across different subsamples. Next, we relate the LLTG-HLTG spread in returns to the return factors commonly used in the finance literature. We then present two-way sorts of LTG and, alternatively, momentum and size. We conclude by presenting value-weighted returns for LTG portfolios.

We begin by analysing the consistency of the returns of HLTG and LLTG portfolios during subsamples. Panel A in Figure IA.1 illustrates the performance of the LLTG and HLTG portfolios over time. It shows that the HLTG portfolio exhibits extreme volatility, particularly during the 1998 to 2001 period. Panel B in Figure IA.1 splits the sample period roughly in half and shows the performance of LTG portfolios during the periods 1981 to 1997 (left panel) and 1998 to 2015. The LLTG-HLTG spread is roughly 14 percentage points during the first half of the sample and 12 percentage points during the second half. These calculations are sensitive to whether the 1998 formation period is included during the first or the second half of the sample. Specifically, if we do the former, the LLTG-HLTG spread is roughly 9 percentage points during the 1981 to 1998 period and 16 percentage points during the 1999 to 2015 period. Either way, the LLTG-HLTG spread is statistically indistinguishable between subsamples.

Panel A



Panel B

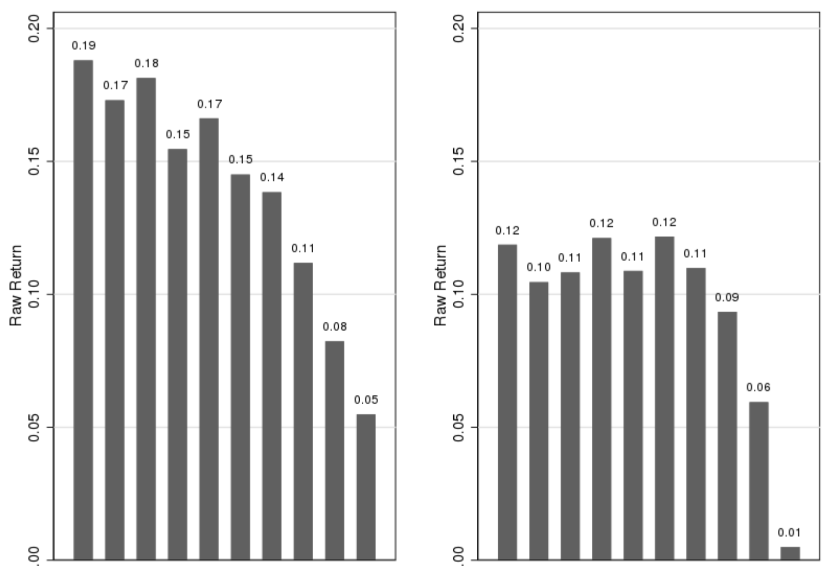


Figure IA.1. Annual returns for portfolios formed on LTG. In December of each year between 1981 and 2015, we form decile portfolios based on ranked analysts' expected growth in EPS. Panel A shows the time series of HLTG and LLTG returns during the full sample. Panel B shows returns for LTG portfolios formed during the periods 1981 to 1997 (on the left) and 1998 to 2015 (on the right). The returns in Panel B are geometric averages of one-year equally weighted returns over the relevant sample periods.

Next, we examine the relationship between the performance of a portfolio that is long LLTG stocks and short HLTG stocks and some of the return factors commonly used in the finance literature (see Fama and French (2015) and the references therein).

The factors that we consider are the five Fama-French factors, momentum (UMD), and betting against beta (BAB). The Fama-French factors are: (1) the excess return on the market relative to the one-month T-bill (Mkt-RF), (2) the difference in the average return of a high book-to-market portfolio and the average return on a low book-to market portfolio (HML), (3) the difference in the average return on a portfolio of small firms and the average return on a portfolio of big stocks (SML), (4) the difference in the average return of a high operating profitability portfolio and a low operating profitability portfolio (RMW), and (5) the difference in the average return of a portfolio of low investment and the average return of a portfolio of high investment (CMA).

Table IA.I shows betas for LTG portfolios. As described in the text, beta increases monotonically with LTG and nearly doubles from the LLTG to HLTG portfolio (0.79 versus 1.51).

Table IA.I

Average Beta is Increasing across LTG Portfolios

We estimate slope coefficients from the following OLS regression

$$ret_{i,t} - rf_t = \alpha_i + \beta_i * (rm_t - rf_{i,t}) + \varepsilon_{i,t},$$

where $ret_{i,t}$ is the monthly return for firm i, rf_t is the risk-free rate (from Ken French's website) in period t, and rm_t is the return on the equally weighted index in period t (also from Ken French's website). We estimate the regression using a rolling window of 60 months. The table reports average βs for LTG portfolios.

	LTG decile										
	1	2	3	4	5	6	7	8	9	10	1-10
β	0.79	0.93	1.00	1.06	1.12	1.18	1.28	1.37	1.46	1.51	-0.72

Turning to the factor analysis, we begin by examining simple correlations of monthly equally weighted portfolio returns in Table IA.II, Panel A. The long-short LLTG-HLTG portfolio is negatively correlated with the market factor, consistent with Table IA.I. It is negatively correlated with the size factor and strongly positively correlated with the book-to-market factor as well as with the investment factor (CMA). As noted in the text, this is unsurprising since – relative to HLTG stocks – LLTG stocks have high book-to-market ratios, low beta, and high profitability (Table I). In contrast, the LTG portfolio only displays a small and insignificant correlation with momentum (UMD). This is confirmed in Panel B, which shows that the LLTG-HLTG spread is approximately constant across momentum buckets (i.e., bottom 30%, middle 40%, and top 30%). Finally, Table IA.II, Panel C reports results of a conventional factor analysis. The monthly excess return of the long-short portfolio is 1.16% when we control only for market risk and falls roughly in half when we control for the standard three Fama-French factors (i.e., MktRF, HML, and SMB). The monthly excess return of the long-short LTG portfolio loses significance if we alternatively control for (1) momentum, (2) betting-against-beta, and (3) profitability (RMW) plus investment (CMA).

Table IA.II

Returns for the spread between LLTG and HLTG

In December of each year between 1981 and 2015, we form decile portfolios based on ranked analysts' expected long-term growth in EPS and compute equally weighted monthly returns for decile portfolios. Panel A reports pair-wise correlations between the return of the portfolio that is long LLTG and short HLTG (LLTG-HLTG) : (1) the excess return on the market relative to the one-month T-Bill (Mkt-RF), (2) the difference in the average return of a high book-to-market portfolio and the average return on a low book-to market portfolio (HML), (3) the difference in the average return on a portfolio of small firms and the average return on a portfolio of big stocks (SML), (4) the difference in the average return on a portfolio with high prior (*t*-12, *t*-2) returns and the average return on a portfolio with low previous returns (UMD), (5) the difference in the average return of a portfolio of low-beta stocks and the average return of a portfolio of high-beta stocks (BAB), (6) the difference in the average return of a high operating profitability portfolio and a low operating profitability portfolio (RMW), and (7) the difference in the average return of a portfolio of low investment and the average return of a portfolio of high investment (CMA). Data on BAB are from the AQR website. All other data on factor returns come from Ken French's website. In Panel B, we independently form ten portfolios based on ranked analysts' expected growth in EPS and three portfolios (i.e., bottom 30%, middle 40%, and top 30%) based on six-month momentum (i.e., July-December of year *t*). The table reports the average one-year return over the subsequent calendar year for equally weighted portfolios. Panel C reports results of OLS regressions of the return of the portfolio that is long LLTG and short HLTG on Mkt-RF, HML, SML, UMD, BAB, RMW, and CMA. [a] denotes significance at the 1% level.

Panel A. Correlates of LLTG-HLTG

	LLTG-HLTG
MktRF	-54%[a]
HML	74%[a]
SMB	-43%[a]
RMW	60%[a]
CMA	63%[a]
BAB	48%[a]
UMD	11%

Panel B. Annual Returns for Portfolios Formed on LTG and Six-Month Momentum

	Returns in the previous 6 months			
LTG	Bottom 30%	Middle	Top 30%	Top-Bottom
LLTG	9.6%	15.9%	17.3%	7.7%
2	9.7%	15.2%	13.9%	4.2%
3	10.6%	15.2%	14.8%	4.2%
4	10.7%	14.4%	14.2%	3.5%

5	10.5%	15.0%	14.5%	3.9%
6	10.3%	14.4%	14.0%	3.6%
7	7.5%	12.3%	15.7%	8.2%
8	6.1%	11.3%	12.4%	6.3%
9	5.7%	7.1%	7.9%	2.2%
HLTG	-1.9%	4.5%	8.0%	9.8%
HLTG-LLTG	-11.5%	-11.4%	-9.3%	2.1%

Panel C. Factor Regressions

Mkt-RF	-0.8761[a]	-0.5327[a]	-0.4581[a]	-0.4953[a]	-0.3512[a]
	(0.0815)	(0.0465)	(0.0436)	(0.0383)	(0.0423)
HML		1.5141[a]	1.6182[a]	1.3270[a]	1.0829[a]
		(0.0813)	(0.0854)	(0.0732)	(0.0901)
SMB		-0.6685[a]	-0.6861[a]	-0.6823[a]	-0.4604[a]
		(0.0744)	(0.0668)	(0.0692)	(0.0717)
UMD			0.2936[a]		
			(0.0759)		
BAB				0.4377[a]	
				(0.0798)	
RMW					0.7918[a]
					(0.1162)
CMA					0.7052[a]
					(0.1768)
Constant	1.1608[a]	0.5526[a]	0.2916	0.1831	0.0298
	(0.3090)	(0.1791)	(0.1956)	(0.1905)	(0.1804)
Observations	408	408	408	408	408
Adjusted R^2	0.293	0.761	0.792	0.802	0.825

We conclude by taking a closer look at the role of size in the performance of LTG portfolios and, related, value-weighted returns. Table IA.III reports average annual compounded returns of portfolios formed independently based on LTG and size. The LLTG-HLTG spread holds within size buckets and ranges from 13.5 percentage points for small stocks to 9.5 percentage points. In contrast, size plays a muted role within LTG buckets, except for the LLTG portfolio where small stocks earn 3.9% higher returns than big stocks. The final column presents value-weighted returns. The LLTG-HLTG spread in (log) returns drops from 13.6%, (t-statistic of 2.22) when returns are equally weighted to 6.4% when returns are value-weighted (t-statistic of 1.2). While the last result lacks significance, this should be interpreted with caution: first, the table shows that equally weighted returns of HLTG portfolios are roughly 3% across all size buckets. Second, value-weighted LTG returns are dominated by a handful of large cap stocks that had high returns, particularly during the early 2000s.

Table IA.III

LTG and Value Weighted Returns

In December of each year between 1981 and 2015, we form decile portfolios based on ranked analysts' expected long-term growth in EPS and form decile portfolios. The first three columns of the table below report average compounded annual returns for independent two-way sorts of LTG and various size portfolios based on market capitalization in December of year *t* (using NYSE breakpoints). *Small* is the portfolio of stocks in the bottom three deciles of the size distribution. *Middle* is the portfolio of stocks in deciles 4 through 7 of the size distribution. *Big* is the portfolio of stocks in the top three deciles of the size distribution. SMB is the difference in the average return of the portfolio that is long *Small* and short *Big*. The last column shows value-weighted returns (WV Ret).

	Size				
LTG	Small	Middle	Big	SMB	VW Ret
1	16.5%	15.0%	12.6%	3.9%	13.5%
2	13.6%	13.0%	13.9%	-0.3%	13.1%
3	12.9%	14.9%	13.8%	-0.9%	12.1%
4	12.9%	14.2%	13.0%	-0.1%	12.1%
5	13.2%	14.6%	12.8%	0.4%	10.4%
6	13.8%	13.2%	13.0%	0.8%	12.8%
7	11.1%	13.2%	11.8%	-0.7%	11.5%
8	10.3%	10.9%	9.9%	0.4%	10.6%
9	8.4%	6.2%	6.4%	2.1%	7.3%
10	3.0%	2.1%	3.1%	-0.1%	7.1%
LLTG-HLTG	13.5%	12.9%	9.5%	4.0%	6.4%

III. Kernel of Truth

We begin by comparing the post-formation growth in earnings of HLTG and LLTG firms. For convenience, Figure IA.2 reproduces Figure 6 (distribution of EPS growth, top left panel) and Figure 7 (distribution of LTG expectations, bottom left panel) for the HLTG portfolio. The middle panels of Figure IA.2 plot the representativeness (i.e., the ratio of the densities in the top left panel) of HLTG verus all other firms. As noted in the text, the most representative future growth realizations for HLTG firms are in the 50% to 90% range of annual growth. Stocks in the LLTG portfolio (see right panels) stand in sharp contrast to those in the HLTG portfolio. The top right panel plots the distribution of earnings growth for all LLTG stocks relative to all other stocks. The right tail of the distribution of growth in earnings for stocks in the LLTG stocks is noticeably thinner than for stocks in all other portfolios. Critically, it is the left tail that is most representative for LLTG stocks (see middle panels). Despite the prominence of the left tail, analysts are not overly pessimistic about the performance of LLTG firms, which narrowly exceed growth expectations (i.e., the post-formation average growth in earnings averages 5% per year rather than the expected 3%).

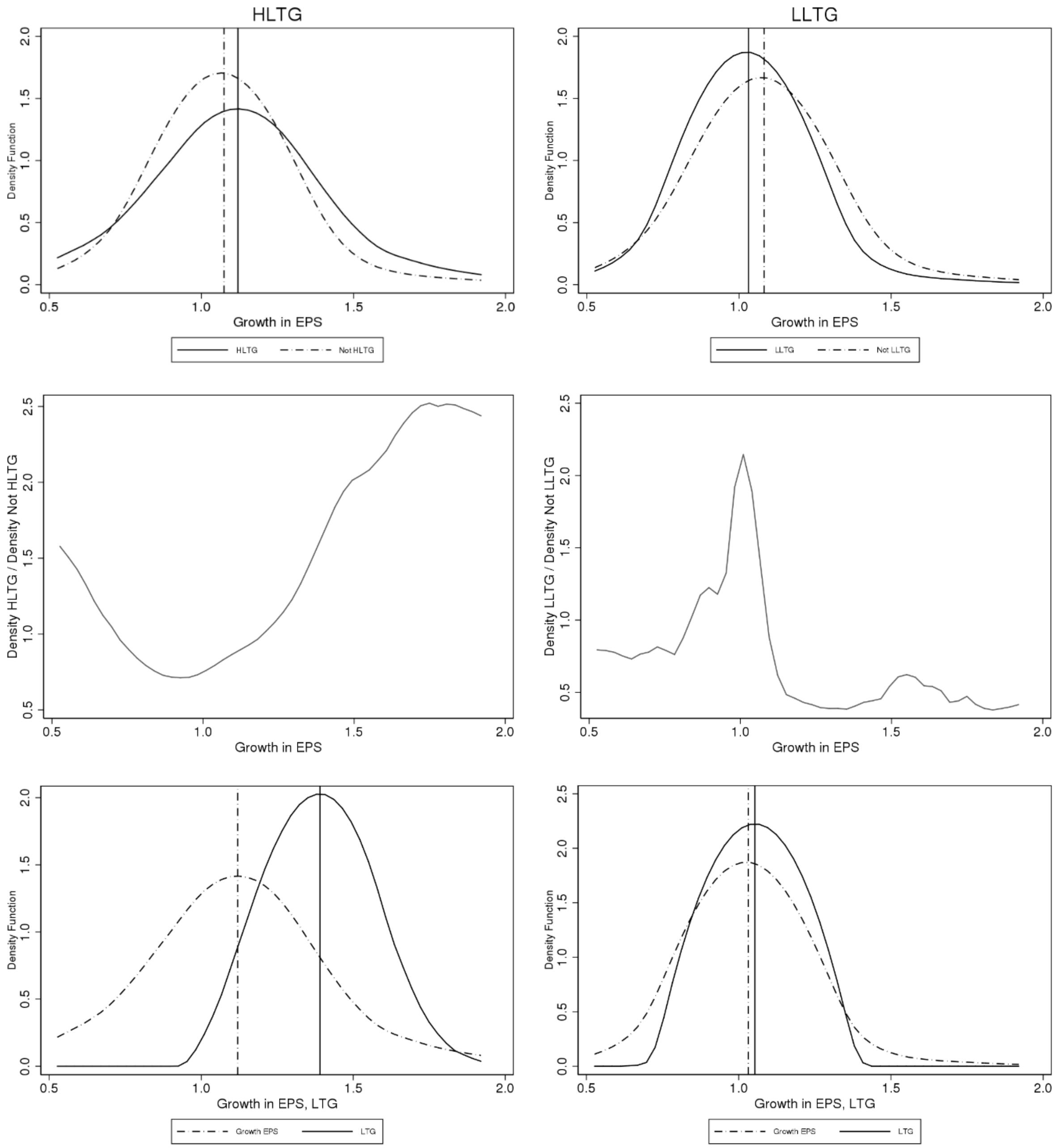


Figure IA.2. **EPS and expectations for the HLGT (left panels) and LLTG (right panels) portfolios**. In December of years (*t*) 1981, 1986, …, and 2011*,* we form decile portfolios based on ranked analysts' expected growth in long-term EPS (*LTG*). For each stock, we compute the gross annual growth rate of EPS between *t* and *t*+5, excluding stocks with negative earnings in year *t* and *t+5*. We estimate the kernel densities for stocks in: (a) the highest (*HLTG*) decile versus all other stocks, and (b) the lowest (*LLTG*) decile versus all other stocks. The top panels show the estimated density kernels of growth in EPS for stocks in the *HLTG* and all other firms on the left and the estimated density kernels of growth in EPS for stocks in the *LLTG* and all other firms on the right. The vertical lines indicate the means of each distribution (1.12 versus1.07 on the left panel and 1.03 versus 1.08 on the right panel). The left middle panel shows the ratio of the densities of growth in EPS for stocks in the *HLTG* portfolio versus stocks in the non-LLTG portfolio. The right middle panel repeats the exercise for stocks in the LLTG portfolio versus stocks in the non-LLTG portfolio. The left bottom panel plots two series: the kernel distribution of the gross annual growth rate in EPS between *t* and *t*+5 for stocks in the HLTG portfolio, and the kernel distribution of the expected growth in long-term EPS at time t for stocks in the HLTG portfolio. The right bottom panel repeats the analysis for stocks in the LLTG portfolio versus stocks in the non-LLTG portfolio. The vertical lines indicate the means of each distribution (1.12 versus 1.39 on the left panel and 1.03 versus 1.05 on the right panel).

The analysis so far focuses on firms with positive earnings. Next, we consider two metrics that make it possible to examine the performance of firms with negative earnings at the cost of shifting the focus away from the variable that theory predicts should matter for pricing stocks. The first such metric is growth in revenue minus operating, general, and sales costs (RMC), which takes advantage of the fact that most (97%) firms with negative earnings have positive values of RMC while staying close to the idea that what matters for expectations is growth. The second metric is defined as the ratio of changes in earnings between t and t+5 scaled to the initial level of sales per share. The latter measure allows us to include all firms in the analysis but is much less directly related to the model's predictions.

Figure IA.3 plots results for growth in RMC in the left panels and changes in earnings in the right panels. Both sets of panels confirm that HLTG firms have a slightly higher mean and a much fatter right tail of exceptional performers. Finally, the evidence in the bottom panels – particularly with regards to growth in RMC – is consistent with the prediction of our model that the representativeness of Googles in the HLTG portfolio leads analysts to overestimate their future performance.

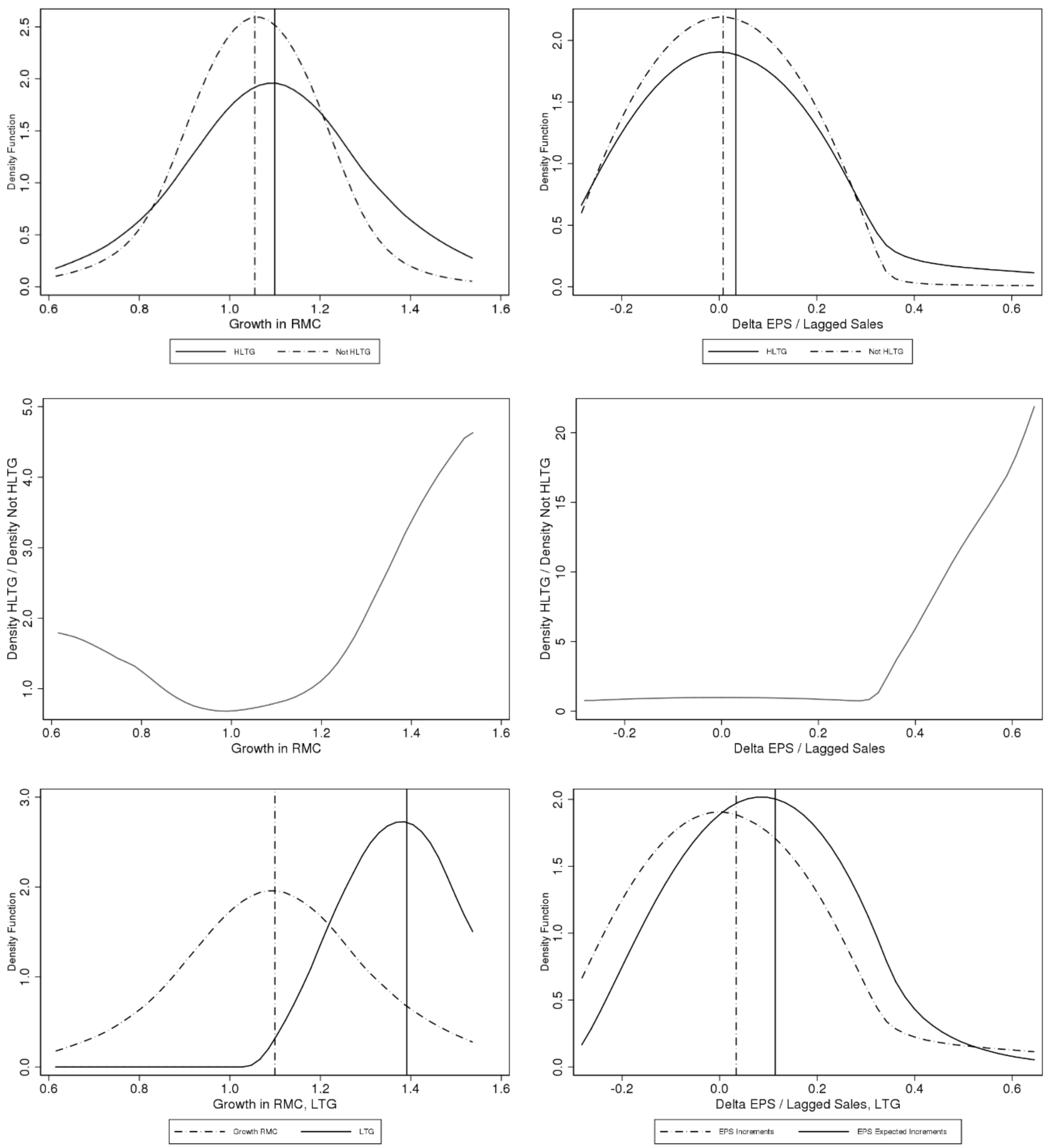


Figure IA.3. Revenues minus cost of goods sold (RMC) (left panels) or change in EPS normalized by lagged sales (right panels) for the HLGT portfolios. In December of years (t) 1981, 1986, …, and 2011*,* we form decile portfolios based on ranked analysts' expected growth in long-term EPS (*LTG*). For each stock, we compute the gross annual growth rate of operating margin (i.e., revenue minus operating, general, and sales costs) per share between t and t+5 as well as the change in EPS between t and t+5 normalized by lagged sales per share. We exclude firms with negative margins in years t and t+5 for the computation of the growth rate of operating margin. The top panel shows the estimated density kernels of growth in operating margins per share for stocks in the *HLTG* portfolio and all other firms on the left and the estimated density kernels of the change in earnings for the *HLTG* portfolio and all other firms on the right. The vertical lines indicate the means of each distribution (1.10 versus 1.06 on the left panel and 0.045 versus 0.01 on the right panel). The left middle panel shows the ratio of the densities of growth in margins for stocks in the *HLTG* portfolio versus stocks in the non-LLTG portfolio on the left and the analogous graph for changes in EPS on the right. The left bottom panel plots two series: the kernel distribution of the gross annual growth rate in operating margins per share between t and t+5 for stocks in the HLTG portfolio, and the kernel distribution of the expected growth in long-term earnings at time t for stocks in the HLTG portfolio. The right bottom panel repeats the analysis for changes in EPS of HLTG stocks. Expected earnings in year t+5 are computed as $EPS_t\,(1+LTG)^5$. The vertical lines indicate the means of each distribution (1.10 versus 1.39 on the left panel and 0.045 versus 0.11 on the right panel).

IV. Coibion and Gorodnichenko Analysis

A. Overreaction to News versus Adaptive Expectations

Adaptive expectations (Equation (11)) predict no overreaction to news after the persistence of the earnings process is accounted for. From (11), the forecast error on an AR(1) process with persistence ρ is $x_{t+1} - x_{t+1}^a = (\rho - 1)x_t + \left(\frac{1-\mu}{\mu}\right)(x_{t+1}^a - x_t^a)$. Controlling for x_t fully accounts for mechanical overreaction in processes with low persistence. The adaptive forecast revision $(x_{t+1}^a - x_t^a)$ should positively predict forecast errors as in the underreaction models. This prediction is not shared by our model because diagnostic expectations overreact to news regardless of the persistence of the data generating process. Table IA.IV reports the results. The coefficients on forecast revision become larger than those estimated in Table II, but they remain mostly negative and statistically significant.

Table IA.IV

Forecast Errors

Each entry in the table corresponds to the estimated coefficient of the forecast errors $(EPS_{t+n}/EPS_t)^{1/n}-LTG_t$ for n=3, 4, and 5 on the variables listed in the first column of the table as well as (log) EPS_t and year fixed effects (not shown).

	Dependent Variable		
	$(EPS_{t+3}/EPS_t)^{1/3}-LTG_t$	$(EPS_{t+4}/EPS_t)^{1/4}-LTG_t$	$(EPS_{t+5}/EPS_t)^{1/5}-LTG_t$
LTG_t-LTG_{t-1}	0.0332	-0.0733	-0.1372[b]
	(0.0725)	(0.0660)	(0.0589)
LTG_t-LTG_{t-2}	-0.0875	-0.1435[b]	-0.1842[a]
	(0.0641)	(0.0691)	(0.0545)

LTG_t-LTG_{t-3}	-0.0956	-0.1184[c]	-0.1701[a]
	(0.0578)	(0.0627)	(0.0517)

B. Underreaction for Forecasts at Short Time Horizons

Table IA.V tests the predictability of forecast errors in the forecast of earnings levels, as opposed to the predictability of errors in LTG forecasts analysed in Table II. The results suggest underreaction for forecasts at short time horizons (i.e., one year ahead), consistent with Bouchaud et al. (2018). As in Table II, as the forecasting horizon increases to three years ahead, the coefficient becomes less positive and even negative in some specifications.

Table IA.V

EPS Forecast Errors at short time horizons

Each entry in the table corresponds to the estimated coefficient of the forecast errors for *t*+1, *t*+3, and *t*+5 on the variables listed in the first column of the table as well as year fixed effects (not shown). All forecast errors are scaled by lagged sales per share (SPS_{t-1}).

	$(EPS_{t+1}-E_tEPS_{t+1}) / SPS_{t-1}$	$(EPS_{t+3}-E_tEPS_{t+3}) / SPS_{t-1}$	$(EPS_{t+5}-E_tEPS_{t+2}*(1+LTG_t)^3) / SPS_{t-1}$
LTG_t-LTG_{t-1}	0.0839	-0.3226[b]	-0.5918[a]
	(0.0554)	(0.1312)	(0.1435)
LTG_t-LTG_{t-2}	0.1629[a]	0.1262[c]	-0.0227
	(0.0275)	(0.0677)	(0.0772)
LTG_t-LTG_{t-3}	0.0825[a]	-0.0664	-0.2145[b]
	(0.0195)	(0.0532)	(0.0919)

C. Overreaction and Predictable Returns

We next try to tie overreaction to news to the return spread between HLTG and LLTG portfolios. We estimate equation (10) by pooling firms at the industry level, using the Fama and French classification. To capture industry- and firm-specific factors, we allow for industry×year fixed effects. This yields an industry level estimate γ_s, where s indexes the industry, which we can correlate with the industry-level LLTG-HLTG spread. These results should be taken with caution, due to the small number of industries.

In Figure IA.4, we compare the post-formation return spread across different terciles of the distribution of industry γ_s. Consistent with our prediction, the extra return obtained by betting against HLTG firms is highest in sectors that feature most overreaction, namely, those in the bottom tercile of γ_s. The return differential is sizable, though given the small sample size it is not statistically significant. Thus, the pattern of LLTG-HLTG return spreads across industries is consistent with a link from overreaction to news to overvaluation of HLTG stocks and thus to abnormally low returns of the HLTG portfolio.

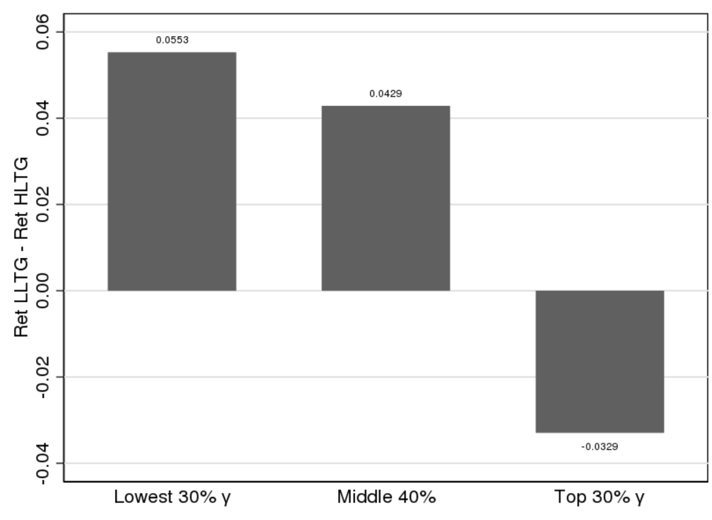


Figure IA.4. Overreaction and return spread across industries. For each of the 48 Fama-French industries, we estimate the regression $(EPS_{t+4}/EPS_t)^{1/4} - (1 + LTG_t) = \alpha_i + \mu_{i,t} + \gamma_i(LTG_t - LTG_{t-3}) + \varepsilon_{i,t}$, where μ_t are year fixed effects, EPS is EPS, and LTG is the forecast of long-term growth in earnings. We rank industries according to γ_{iS} and form the following three groups: (1) 14 industries with the lowest γ_{iS}, (2) 24 industries with intermediate values of γ_{iS}, and (3) 14 industries with the highest γ_{iS}. Finally, for each year and each group, we compute the difference in return for the LLTG (i.e., bottom 30% of LTG) and HLTG (i.e., highest 30% of LTG) portfolios. The graph shows the arithmetic mean of the LLTG-HLTG spread for industries grouped based on γ_{iS} from the regression.

V. Model Estimation

In this section, we extend our estimates with several exercises. We first test the robustness of our estimate $\theta = 0.9$. We then provide the simulation counterparts to the comparative static of overreaction relative to persistence and volatility of the process for growth fundamentals f_t (Table IV) and to the revision of expectations as a function of news (Figure 9).

To test the robustness of our estimate for θ, we evaluate the matching between the target vector of empirical moments $\bar{v} = (0.82, 0.75, 0.70, 0.65, -0.276, -0.126)$ and the predicted counterpart for $\theta \in [0,2]$ (keeping the other parameters constant). Figure IA.5 shows that the match is indeed optimized for $\theta = 0.9$ and that it drops fast as θ departs from this value.

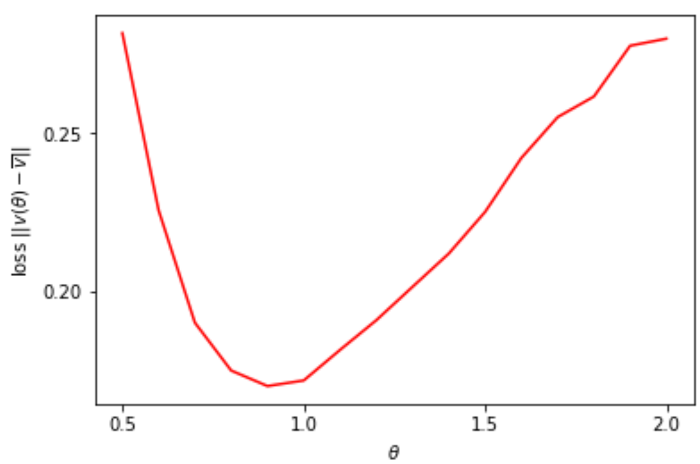


Figure IA.5. Loss function. For $(a, b, \sigma_\eta, \sigma_\epsilon, s) = (0.97, 0.56, 0.138, 0.082, 11)$, as given by our estimation, and each $\theta \in [0,2]$ (in steps of 0.1) we compute the target vector of moments v. The figure plots the loss $\ell(v) = \|v - \bar{v}\|$.

Next, we return to the kernel-of-truth prediction that overreaction increases with the drivers of the signal-to-noise ratio. To illustrate this prediction, we compute the returns of LLTG and HLTG portfolios under our estimates for $(b, \sigma_\epsilon, \theta, s)$ while allowing the persistence a and volatility σ_η of the growth fundamentals f_t to vary over parameter specifications comparable to those in Table IV. Figure IA.6 presents the results. Consistent with the empirical evidence of Table 4, the LLTG-HLTG return spread increases with both a and σ_η.

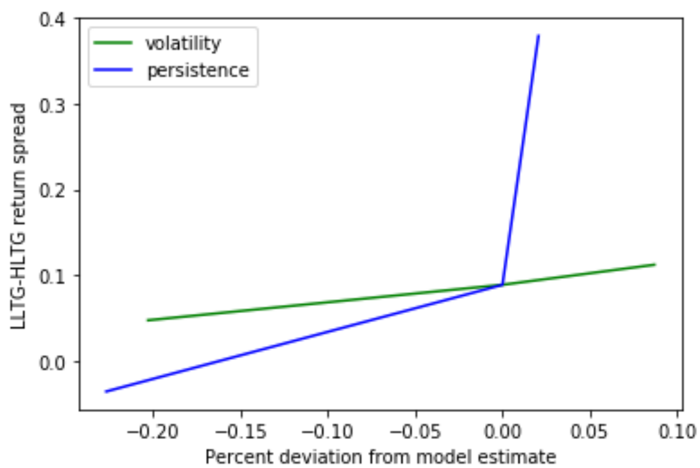


Figure IA.6. Volatility and persistence of the LLTG-HLTG spread in returns. For $(b, \sigma_\epsilon, \theta, s) = (0.56, 0.08, 0.90, 11)$, as given by our estimation, we vary the persistence of fundamentals $a \in \{0.75, 0.97, 0.99\}$ (keeping $\sigma_\eta = 0.138$) or the volatility of fundamentals $\sigma_\eta \in \{0.11, 0.138, 0.17\}$ (keeping $a = 0.97$). For each parameter combination we compute the LLTG-HLTG return spread.

Finally, we confirm numerically the pattern of Figure 9, namely, that high LTG forecasts are on average revised downwards, even conditional on positive news, that is, earnings that exceeded (distorted) expectations. Figure IA.7 below plots the equivalent of Figure 9 for our estimation.

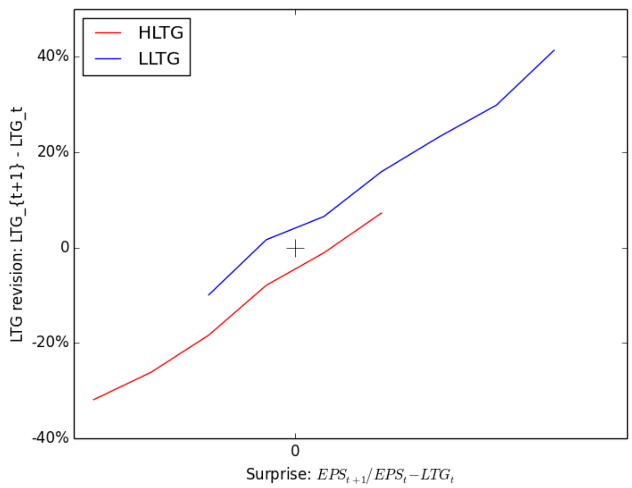


Figure IA.7. Earnings surprises and LTG revisions. For each period t and each firm in the $HLTG_t$ and $LLTG_t$ portfolios, we compute the surprise $\ln EPS_{t+1} - \ln EPS_t - LTG_t$. Pooling the data for all t and both portfolios, we rank the surprises in deciles and plot the average revision of LTG for each portfolio in each decile.

In the model, news and forecast revisions are positively correlated. Crucially, however, they may go in opposite directions. For a range of positive surprises, namely, realized growth above the diagnostic forecasts, forecasts about HLTG firms are still revised downwards. Naturally, for sufficiently large positive surprises (larger than $1.5\sigma_\epsilon$ in our estimation), forecast revisions are positive. The converse holds for LLTG.

References for the Internet Appendix

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar, 2018, Sticky expectations and the profitability anomaly, *Journal of Finance* 74, 639-674 .

Fama, Eugene, and Kenneth French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1 – 22.