Supplementary Appendix

VARIABLE DEFINITIONS

We describe the 11 anomaly variables used to replicate the Stambaugh-Yuan (2017) factors. At the beginning of each month, we rank stocks into percentiles (1 to 100) based on each anomaly. The rankings are created such that high rankings are associated with lower future average returns. The first composite measure, MGMT (management), is the average of the six percentile rankings in net stock issues, composite equity issuance, accruals, net operating assets, asset growth, and investment-to-assets. The second composite measure, PERF (performance), is the average of the five percentile rankings in failure probability, O-score, momentum, gross profitability, and return on assets. In any given month, an anomaly variable needs at least 30 stocks with non-missing values in order to be included in the composite measure. In addition, we compute a composite measure for a stock only if it has non-missing values for at least three of the (six or five) component anomalies.

Net stock issues. Net stock issues is the annual change in the log of split-adjusted shares outstanding. The split-adjusted shares outstanding is shares outstanding (Compustat annual item CSHO) times the adjustment factor (item AJEX). At the beginning of each month, we use the latest net stock issues from fiscal year ending at least four months ago.

Composite equity issuance. Following Stambaugh and Yuan (2017), at the beginning of month t, we measure composite equity issuance as the growth rate in market equity minus the cumulative stock return from month $t-16$ to $t-5$ (skipping month $t-4$ to $t-1$).

Accruals. We measure accruals as changes in noncash working capital minus depreciation, in which the noncash working capital is changes in noncash current assets minus changes in current liabilities less short-term debt and taxes payable. In particular, accruals equals (dCA – dCASH) – (dCL – dSTD – dTP) – DP, in which dCA is the change in current assets (Compustat annual item ACT), dCASH is the change in cash or cash equivalents (item CHE), dCL is the change in current liabilities (item LCT), dSTD is the change in debt included in current liabilities (item DLC), dTP is the change in income taxes payable (item TXP), and DP is depreciation and amortization (item DP). Missing changes in income taxes payable are set to zero. We scale accruals by average total assets from the previous and current years. At the beginning of each month, we use the latest accruals from fiscal year ending at least four months ago.

Net operating assets. We measure net operating assets as operating assets minus operating liabilities. Operating assets are total assets (Compustat annual item AT) minus cash and short-term investment (item CHE). Operating liabilities are total assets minus debt included in current liabilities (item DLC, zero if missing), minus long-term debt (item DLTT, zero if missing), minus minority interests (item MIB, zero if missing), minus preferred stocks (item PSTK, zero if missing), and minus common equity (item CEQ). We scale net operating assets by one-year-lagged total assets. At the beginning of each month, we use the latest net operating assets from fiscal year ending at least four months ago.

Asset growth. Asset growth is the annual change in total assets (Compustat annual item AT) scaled by 1-year-lagged total assets. At the beginning of each month, we use the latest asset growth from fiscal year ending at least four months ago.

Changes in PPE and Inventory-to-assets are the annual change in gross property, plant, and equipment (Compustat annual item PPEGT) plus the annual change in inventory (item INVT) scaled by 1-year-lagged assets (item AT). At the beginning of each month, we use the latest investment-to-assets from fiscal year ending at least four months ago.

Failure Probability. At the beginning of month t, we follow Campbell, Hilscher, and Szilagyi (2008, Table IV, Column 3) to construct failure probability:

$$\text{Fp}_t \equiv -9.164 - 20.264\text{NIMTAAVG}_t + 1.416\text{TLMTA}_t - 7.129\text{EXRETAVG}_t + 1.411\text{SIGMA}_t - 0.045\text{RSIZE}_t - 2.132\text{CASHMTA}_t + 0.075\text{MB}_t - 0.058\text{PRICE}_t \quad \text{(A1)}$$

in which

$$\text{NIMTAAVG}_{t-1,t-12} \equiv \frac{1-\phi^3}{1-\phi^{12}}\left(\text{NIMTA}_{t-1,t-3} + \cdots + \phi^9\text{NIMTA}_{t-10,t-12}\right) \quad \text{(A2)}$$

$$\text{EXRETAVG}_{t-1,t-12} \equiv \frac{1-\phi}{1-\phi^{12}}\left(\text{EXRET}_{t-1} + \cdots + \phi^{11}\text{EXRET}_{t-12}\right), \quad \text{(A3)}$$

and $\phi = 2^{-1/3}$. NIMTA is net income (Compustat quarterly item NIQ) divided by the sum of market equity (share price times the number of shares outstanding from CRSP) and total liabilities (item LTQ). The moving average NIMTAAVG captures the idea that a long history of losses is a better predictor of bankruptcy than one large quarterly loss in a single month. EXRET $\equiv \log(1 + R_{it}) - \log(1 + R_{\text{S\&P500},t})$ is the monthly log excess return on each firm's equity relative to the S&P 500 index. The moving average EXRETAVG captures the idea that a sustained decline in stock market value is a better predictor of bankruptcy than a sudden stock price decline in a single month.

TLMTA is total liabilities divided by the sum of market equity and total liabilities. SIGMA is the annualized three-month rolling sample standard deviation: $\sqrt{\frac{252}{N-1}\sum_{k\in\{t-1,t-2,t-3\}} r_k^2}$, in which k is the index of trading days in months $t-1, t-2$, and $t-3$, r_k is the firm-level daily return, and N is the total number of trading days in the three-month period. SIGMA is treated as missing if there are less than five nonzero observations over the three months in the rolling window. RSIZE is the relative size of each firm measured as the log ratio of its market equity to that of the S&P 500 index. CASHMTA, aimed to capture the liquidity position of the firm, is cash and short-term investments (Compustat quarterly item CHEQ) divided by the sum of market equity and total liabilities (item LTQ). MB is the market-to-book equity, in which we add 10% of the difference between the market equity and the book equity to the book equity to alleviate measurement issues for extremely small book equity values (Campbell, Hilscher, and Szilagyi 2008). For firm-month observations that still have negative book equity after this adjustment, we replace these negative values with \$1 to ensure that the market-to-book ratios for these firms are in the right tail of the distribution. PRICE is each firm's log price per share, truncated above at \$15. We further eliminate stocks with prices less than \$1 at the portfolio formation date. Variables requiring quarterly accounting data are from fiscal quarter ending at least four months ago to ensure the availability of balance sheet items. We winsorize the variables on the right-hand side of equation (A1) at the 1th and 99th percentiles of their distributions each month.

Ohlson's O-score. The O-score is defined as:

$$O \equiv -1.32 - 0.407\log(\text{TA}) + 6.03\text{TLTA} - 1.43\text{WCTA} + 0.076\text{CLCA} - 1.72\text{OENEG} - 2.37\text{NITA} - 1.83\text{FUTL} + 0.285\text{INTWO} - 0.521\text{CHIN}, \quad \text{(A4)}$$

in which TA is total assets (Compustat annual item AT). TLTA is the leverage ratio defined as total debt (item DLC plus item DLTT) divided by total assets. WCTA is working capital (item ACT minus item LCT) divided by total assets. CLCA is current liability (item LCT) divided by current assets (item ACT). OENEG is one if total liabilities (item LT) exceeds total assets and zero otherwise. NITA is net income (item NI) divided by total assets. FUTL is the fund provided by operations (item PI plus item DP) divided by total liabilities. INTWO is equal to one if net income is negative for the last two years and zero otherwise. CHIN is $(\text{NI}_s - \text{NI}_{s-1})/(|\text{NI}_s| + |\text{NI}_{s-1}|)$, in which NI_s and NI_{s-1} are the net income for the current and prior years. We winsorize all non-dummy variables on the right-hand side of equation (A4) at the 1th and 99th percentiles of their distributions each year. At the beginning of each month, we use the latest O-score from fiscal year ending at least four months ago.

Momentum. At the beginning of each month t, we measure momentum as the 11-month cumulative return from month $t-12$ to $t-2$ (skipping month $t-1$).

Gross Profitability is total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS) divided by total assets (item AT). At the beginning of each month, we use the latest gross profitability from fiscal year ending at least four months ago.

Return on Assets is income before extraordinary items (Compustat quarterly item IBQ) divided by 1-quarter-lagged total assets (item ATQ). At the beginning of each month, we use return on assets computed with quarterly earnings from the most recent earnings announcement dates (item RDQ). For a firm to enter our sample, we require the end of the fiscal quarter that corresponds to its most recent return on assets to be within six months prior to the portfolio formation. This restriction is imposed to exclude stale earnings information. To avoid potentially erroneous records, we also require the earnings announcement date to be after the corresponding fiscal quarter end.

SUPPLEMENTARY RESULTS ON SPANNING TESTS

Table A1 shows spanning tests on the q-factor and q^5 models versus the replicated Stambaugh-Yuan model (Panel A) as well as the replicated Daniel-Hirshleifer-Sun model (Panel B), following their exact sample selection criteria, respectively.

ESTIMATING THE INTERNAL RATE OF RETURN

The Gebhardt, Lee, and Swaminathan (2001, GLS) Procedure. At the end of June in each year t, we estimate the IRR from the following nonlinear equation:

$$P_t = Be_t + \sum_{\tau=1}^{11} \frac{(E_t[\text{Roe}_{t+\tau}] - \text{IRR}) \times Be_{t+\tau-1}}{(1+\text{IRR})^\tau} + \frac{(E_t[\text{Roe}_{t+12}] - \text{IRR}) \times Be_{t+11}}{\text{IRR} \times (1+\text{IRR})^{11}}, \quad \text{(A5)}$$

in which P_t is the market equity in year t, $Be_{t+\tau}$ is the book equity, and $E_t[\text{Roe}_{t+\tau}]$ is the expected return on equity for year $t+\tau$ based on information available in year t.

We measure current book equity, Be_t, using the latest accounting data from the fiscal year ending between March of year $t-1$ to February of t. This practice implies that for the IRR estimates at the end of June in t, we impose at least a four-month lag to ensure that the accounting information is released to the public. The definition of book equity follows Davis, Fama, and French (2000). We apply clean surplus accounting to construct future book equity as $Be_{t+\tau} = Be_{t+\tau-1} + Be_{t+\tau-1}E_t[\text{Roe}_{t+\tau}](1-k), 1 \leq \tau \leq 11$, in which k is the dividend payout ratio in year t. Dividend payout ratio is dividends (Compustat annual item DVC) divided by earnings (item IB) for profitable firms, or dividends divided by 6% of total assets (item AT) for firms with zero or negative earnings. We drop a firm if its book equity is zero or negative in any year. We construct the expected Roe for the

Table A1 Spanning Tests, Supplementary Results (January 1967–December 2016)

Panel A: The replicated Stambaugh-Yuan model with their sample criterion

	$\overline{R}$	α	MKT	SMB	MGMT	PERF	R^2		$\overline{R}$	α	R_{Mkt}	R_{Me}	$R_{\text{I/A}}$	R_{Roe}	R_{Eg}	R^2
R_{Me}	0.31	−0.00	0.01	1.03	−0.01	0.01	91	SMB	0.31	0.06	0.01	0.87	0.00	−0.05		92
	2.43	−0.11	0.40	26.50	−0.47	0.29			2.51	1.42	0.38	22.52	0.07	−2.16		
$R_{\text{I/A}}$	0.41	0.09	0.02	0.06	0.69	−0.02	67			0.11	−0.00	0.86	0.02	−0.03	−0.07	92
	4.92	1.59	1.19	2.82	18.38	−0.68				2.56	−0.15	22.12	0.66	−0.94	−2.56	
R_{Roe}	0.55	0.31	0.01	−0.22	0.15	0.54	44	MGMT	0.44	0.19	−0.12	−0.10	0.86	−0.02		73
	5.25	3.51	0.49	−3.90	1.91	10.01			4.26	3.26	−7.00	−4.33	23.19	−0.61		
R_{Eg}	0.82	0.56	−0.03	−0.12	0.43	0.29	55			−0.01	−0.09	−0.07	0.79	−0.11	0.31	76
	9.81	9.66	−2.26	−5.35	13.88	11.58				−0.17	−6.20	−3.57	22.49	−3.49	8.18	
								PERF	0.44	0.24	−0.09	0.09	−0.37	0.66		41
									3.70	1.92	−2.78	1.72	−4.00	10.78		
										0.02	−0.05	0.12	−0.46	0.55	0.35	44
										0.20	−1.75	2.30	−4.92	8.08	4.36	

	$\alpha_{\text{I/A}}, \alpha_{\text{Roe}} = 0$	$\alpha_{\text{I/A}}, \alpha_{\text{Roe}}, \alpha_{\text{Eg}} = 0$	$\alpha_{\text{MGMT}}, \alpha_{\text{PERF}} = 0$ in q	$\alpha_{\text{MGMT}}, \alpha_{\text{PERF}} = 0$ in q^5
GRS	7.91	34.93	11.09	0.03
p	0.00	0.00	0.00	0.97

Panel B: The replicated Daniel-Hirshleifer-Sun model with their sample criterion

	$\overline{R}$	α	MKT	FIN	PEAD	R^2		$\overline{R}$	α	MKT	R_{Me}	$R_{\text{I/A}}$	R_{Roe}	R_{Eg}	R^2
R_{Me}	0.31	0.63	0.00	−0.46	−0.23	23	FIN	0.33	0.00	−0.16	−0.22	0.86	0.23		69
	2.43	4.25	0.07	−3.74	−3.16			2.55	0.04	−6.77	−3.79	13.46	4.32		
$R_{\text{I/A}}$	0.41	0.32	−0.01	0.43	−0.07	47			−0.06	−0.15	−0.22	0.83	0.20	0.09	69
	4.92	4.31	−0.26	8.65	−1.98				−0.65	−6.83	−3.46	11.92	3.32	1.54	
R_{Roe}	0.55	−0.16	0.04	0.32	0.79	59	PEAD	0.74	0.44	0.00	0.02	−0.10	0.61		49
	5.25	−2.21	1.91	6.43	19.47			8.00	5.34	−0.17	0.54	−1.64	11.96		
R_{Eg}	0.82	0.54	−0.08	0.28	0.31	47			0.33	0.02	0.03	−0.15	0.56	0.18	50
	9.81	7.35	−4.54	8.24	8.68				4.25	0.78	1.01	−2.31	9.10	2.92	

	$\alpha_{\text{I/A}}, \alpha_{\text{Roe}} = 0$	$\alpha_{\text{I/A}}, \alpha_{\text{Roe}}, \alpha_{\text{Eg}} = 0$	$\alpha_{\text{FIN}}, \alpha_{\text{PEAD}} = 0$ in q	$\alpha_{\text{FIN}}, \alpha_{\text{PEAD}} = 0$ in q^5
GRS	14.39	34.82	22.31	9.57
p	0.00	0.00	0.00	0.00

$\overline{R}$ is the average return, α the intercept, and R^2 its goodness-of-fit in percent. $R_{\text{Mkt}}, R_{\text{Me}}, R_{\text{I/A}}$, and R_{Roe} are the market, size, investment, and Roe factors in the q-factor and q^5 models, respectively, and R_{Eg} the expected growth factor in the q^5 model. In Panel A, MKT, SMB, MGMT, and PERF are the market, size, management, and performance factors in the Stambaugh-Yuan model. In Panel B, MKT, FIN, and PEAD are the market, financing, and post-earnings-announcement-drift factors in the Daniel-Hirshleifer-Sun model. The t-values (reported in the rows beneath the corresponding estimates) are adjusted for heteroscedasticity and autocorrelations.

first three years ahead, using analyst earnings forecasts from the Institutional Brokers' Estimated System (IBES) or forecasts from cross-sectional regressions. After year $t+3$, we assume that the expected firm-level Roe mean-reverts linearly to the historical industry median Roe by year $t+12$, and becomes a perpetuity afterwards. We use the Fama-French (1997) 48 industry classification. We use at least five and up to ten years of past Roe data from non-loss firms to compute the industry median Roe.

GLS (2001) use a per share basis with analysts' earnings forecasts. P_t is the June-end share price from CRSP. Be_t is book equity per share calculated as book equity divided by the number of shares outstanding reported in June from IBES (unadjusted file, item SHOUT). When IBES shares are not available, we use shares from CRSP (daily item SHROUT) on the IBES pricing date (item PRDAYS) that corresponds to the IBES report. At the end of June in each year t, we construct the expected Roe for year $t+1$ to $t+3$ as $E_t[\text{Roe}_{t+\tau}] = \text{FEPS}_{t+\tau}/Be_{t+\tau-1}$, in which $\text{FEPS}_{t+\tau}$ is the consensus mean forecast of earnings per share from IBES (unadjusted file, item MEANEST) for year $t+\tau$ (fiscal period indicator $=\tau$) reported in June of t. We require the availability of earnings forecast for years $t+1$ and $t+2$. When the forecast for year $t+3$ is not available, we use the long-term growth rate (item LTG) to compute a three-year-ahead forecast: $\text{FEPS}_{t+3} = \text{FEPS}_{t+2} \times (1+\text{LTG})$. If the long-term growth rate is missing, we replace it with the growth rate implied by the first two forecasts: $\text{FEPS}_{t+3} = \text{FEPS}_{t+2} \times (\text{FEPS}_{t+2}/\text{FEPS}_{t+1})$, when $\text{FEPS}_{t+1}, \text{FEPS}_{t+2} > 0$.

As noted, we measure current book equity Be_t based on the latest accounting data from the fiscal year ending between March of year $t-1$ and February of t. However, firms with fiscal years ending between March of t and May of t can announce their latest earnings before the IBES report in June of t. In response to earnings announcement for the current fiscal year, the analyst forecasts would "roll forward" to the next year. As such, we also need to roll forward book equity by one year for these firms to match with the updated analyst forecasts. In particular, we roll forward their book equity using clean surplus accounting as: $Be_{t-1} + Y_t - D_t$, in which Be_{t-1} is the lagged book equity (relative to the announced earnings), Y_t is the earnings announced after February of t but before the IBES report in June of t, and D_t is dividends.

In the first modified procedure, we follow Hou, van Dijk, and Zhang (2012) to estimate the IRRs at the firm level (not the per share basis), whenever regression-based earnings forecasts (not analysts' earnings forecasts) are used. We use pooled cross-sectional regressions to forecast future earnings for up to three years ahead:

$$Y_{is+\tau} = a + b_1 A_{is} + b_2 D_{is} + b_3 DD_{is} + b_4 Y_{is} + b_5 Y_{is}^{-} + b_6 AC_{is} + \epsilon_{is+\tau}, \qquad \text{(A6)}$$

for $1 \leq \tau \leq 3$, in which Y_{is} is earnings (Compustat annual item IB) of firm i for fiscal year s, A_{is} is total assets (item AT), D_{is} is dividends (item DVC), and DD_{is} is a dummy variable that equals one for dividend payers, and zero otherwise. Y_{is}^{-} is a dummy variable that equals one for negative earnings, and zero otherwise, and AC_{is} is operating accruals.

Prior to 1988, we use the balance-sheet approach of Sloan (1996) to measure operating accruals as changes in noncash working capital minus depreciation, in which the noncash working capital is changes in noncash current assets minus changes in current liabilities less short-term debt and taxes payable. In particular, $AC = (\triangle CA - \triangle CASH) - (\triangle CL - \triangle STD - \triangle TP) - DP$, in which $\triangle CA$ is the change in current assets (Compustat annual item ACT), $\triangle CASH$ is the change in cash or cash equivalents (item CHE), $\triangle CL$ is the change in current liabilities (item LCT), $\triangle STD$ is the change in debt included in current liabilities (item DLC, zero if missing), $\triangle TP$ is the change in income taxes payable (item TXP, zero if missing), and DP is depreciation and amortization (item DP, zero if missing). Starting from 1988, we follow Hribar and Collins (2002) to measure AC using the statement of cash flows as net income (item NI) minus net cash flow from operations (item OANCF).

In equation (A6), regressors with time subscript s are from the fiscal year ending between March of year s and February of $s+1$. Following Hou, van Dijk, and Zhang (2012), we winsorize all the level variables in equation (A6) at the 1st and 99th percentiles of their cross-sectional distributions each year. In June of each year t, we estimate the regressions using the pooled panel data from the previous ten years. With a minimum four-month lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. Differing from the baseline GLS procedure, we forecast the expected earnings as the estimated regression coefficients times the latest values of the (unwinsorized) predictors from the fiscal year ending between March of $t-1$ and February of t.

In the second modified procedure, we use annual cross-sectional regressions per Tang, Wu, and Zhang (2014) to forecast the future ROE for up to three years, $\text{Roe}_{is+\tau} \equiv Y_{is+\tau}/Be_{is+\tau-1}$:

$$\text{Roe}_{is+\tau} = a + b_1 \log\left(\frac{Be_{is}}{P_{is}}\right) + b_2 \log(P_{is}) + b_3 Y_{is}^{-} + b_4 \text{Roe}_{is} + b_5 \frac{A_{is} - A_{is-1}}{A_{is-1}} + \epsilon_{is+\tau}, \tag{A7}$$

for $1 \leq \tau \leq 3$, in which Roe_{is} is return on equity of firm i for fiscal year s, Y_{is} is earnings (Compustat annual item IB), Be_{is} is the book equity, P_{is} is the market equity at the fiscal year end from Compustat or CRSP, Y_{is}^{-} is a dummy variable that equals one for negative earnings and zero otherwise, and A_{is} is total assets (item AT). Regression variables with time subscript s are from the fiscal year ending between March of year s and February of $s+1$. Extremely small firms tend to have extreme regression variables which can affect the Roe regression estimates significantly. To alleviate this problem, we exclude firm-years with total assets less than \$5 million or book equity less than \$2.5 million. Fama and French (2006, p. 496) require firms to have at least \$25 million total assets and \$12.5 million book equity, but state that their results are robust to using the \$5 million total assets and \$2.5 million book equity cutoff. We choose the less restrictive cutoff to enlarge the sample coverage. We also winsorize each variable (except for Y_{it}^{-}) at the 1 and 99 percentiles of its cross-sectional distribution each year to limit the impact of outliers.

In June of each year t, we run the regression (A7) using the previous ten years of data. With a minimum four-month information lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. Differing from the baseline GLS procedure, we directly forecast the expected Roe, $E_t[\text{Roe}_{t+\tau}]$, as the average cross-sectional regression coefficients times the latest values of the predictors from fiscal years ending between March of $t-1$ and February of t. We implement this modified GLS procedure at the firm level.

The Easton (2004) Procedure. At the end of June in each year t, we estimate the IRR from:

$$P_t = \frac{E_t[Y_{t+2}] + \text{IRR} \times E_t[D_{t+1}] - E_t[Y_{t+1}]}{\text{IRR}^2}, \tag{A8}$$

in which P_t is the market equity in year t, $E_t[Y_{t+\tau}]$ is the expected earnings for year $t+\tau$ based on information available in year t, and $E_t[D_{t+1}]$ is the expected dividends for year $t+1$.

Expected earnings are based on analyst forecasts from IBES or forecasts from regression models. Expected dividends are expected earnings times the current dividend payout ratio, which is computed as dividends (Compustat annual item DVC) divided by earnings (item IB) for profitable firms, or dividends divided by 6% of total assets (item AT) for firms with zero or negative earnings. When equation (A8) has two positive roots (in very few cases), we use the average as the IRR estimate.

Following Easton (2004), we implement the model on the per share basis with analysts' earnings forecasts. We measure P_t as the June-end share price from CRSP. At the end of June in year t, the expected earnings per share for year $t+\tau$ is the consensus mean forecast from IBES (unadjusted file, item MEANEST) for year $t+\tau$ (fiscal period indicator $=\tau$) reported in June of t.

Instead of analysts' earnings forecasts, we also use pooled cross-sectional regressions in equation (A6) to forecast future earnings for up to two years ahead. In June of each year t, we estimate the regression using the pooled panel data from the previous ten years. With a four-month information lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. We construct the expected earnings as the estimated regression coefficients times the latest values of the (unwinsorized) predictors from the fiscal year ending between March of $t-1$ and February of t. We implement the modified procedure at the firm level.

Finally, we also use annual cross-sectional regressions in equation (A7) to forecast future ROE for up to two years ahead. In June of each year t, we estimate the regression using the previous ten years of data. With a four-month information lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. We forecast the expected Roe, $E_t[\mathrm{Roe}_{t+\tau}]$, as the average regression coefficients times the latest values of the predictors from fiscal years ending between March of $t-1$ and February of t. Expected earnings are then constructed as: $E_t[Y_{t+\tau}] = E_t[Roe_{t+\tau}] \times Be_{t+\tau-1}$, in which $Be_{t+\tau-1}$ is the book equity in year $t+\tau-1$. We measure current book equity Be_t based on the latest accounting data from the fiscal year ending in March of $t-1$ to February of t, and impute future book equity by applying clean surplus accounting recursively. We implement the modified procedure at the firm level.

The Claus and Thomas (2001, CT) Procedure. At the end of June in each year t, we estimate the IRR from:

$$P_t = Be_t + \sum_{\tau=1}^{5} \frac{(E_t[\mathrm{Roe}_{t+\tau}] - \mathrm{IRR}) \times Be_{t+\tau-1}}{(1+\mathrm{IRR})^{\tau}} + \frac{(E_t[\mathrm{Roe}_{t+5}] - \mathrm{IRR}) \times Be_{t+4} \times (1+g)}{(\mathrm{IRR}-g) \times (1+\mathrm{IRR})^{5}}, \tag{A9}$$

in which P_t is the market equity in year t, $Be_{t+\tau}$ is the book equity, $E_t[\mathrm{Roe}_{t+\tau}]$ is the expected Roe for year $t+\tau$ based on information available in year t, and g is the long-term growth rate of abnormal earnings. Abnormal earnings are defined as $(E_t[\mathrm{Roe}_{t+\tau}] - \mathrm{IRR}) \times Be_{t+\tau-1}$.

We measure book equity using the latest accounting data from the fiscal year ending between March of year $t-1$ and February of t. The definition follows Davis, Fama, and French (2000). We apply clean surplus accounting to construct future book equity as $Be_{t+\tau} = Be_{t+\tau-1} + Be_{t+\tau-1}E_t[\mathrm{Roe}_{t+\tau}](1-k), 1 \leq \tau \leq 4$, in which k is the dividend payout ratio in year t. Dividend payout ratio is dividends (Compustat annual item DVC) divided by earnings (item IB) for profitable firms, or dividends divided by 6% of total assets (item AT) for firms with zero or negative earnings. We drop a firm if its book equity is zero or negative in any year. We construct the expected Roe, $E_t[\mathrm{Roe}_{t+\tau}]$, for up to five years ahead, using analysts' earnings forecasts from IBES or regression-based forecasts. Following CT (2001), we set g to the ten-year Treasury bond rate minus 3%.

Following CT (2001), we implement the CT model on the per share basis when using analysts' earnings forecasts. We measure P_t as the June-end share price from CRSP. Book equity per share, B_t, is book equity divided by the number of shares outstanding reported in June from IBES (unadjusted file, item SHOUT). When IBES shares are not available, we use shares from CRSP (daily item SHROUT) on the IBES pricing date (item PRDAYS) that corresponds to the IBES report. As noted, current book equity B_t is based on the latest accounting data from the fiscal year ending between March of $t-1$ and February of t. However, firms with fiscal year ending in March of t to May of t can announce their latest earnings before the IBES report in June of t. To match the updated analyst forecasts, we roll forward their book equity using clean surplus accounting as: $B_{t-1} + Y_t - D_t$, in which B_{t-1} is the lagged book equity (relative to the announced earnings), Y_t is the earnings announced after February of t but before the IBES report in June of t, and D_t is dividends.

At the end of June in each year t, we construct the expected Roe for year $t+1$ to $t+5$ as $E_t[\text{Roe}_{t+\tau}] = \text{FEPS}_{t+\tau}/Be_{t+\tau-1}$, in which $\text{FEPS}_{t+\tau}$ is the consensus mean forecast of earnings per share from IBES (unadjusted file, item MEANEST) for year $t+\tau$ (fiscal period indicator $=\tau$) reported in June of t. We require the availability of earnings forecast for years $t+1$ and $t+2$. When the forecast after year $t+2$ is not available, we use the long-term growth rate (item LTG) to construct it as $\text{FEPS}_{t+\tau} = \text{FEPS}_{t+\tau-1} \times (1+\text{LTG})$. If the long-term growth rate is missing, we replace it with the growth rate implied by the forecasts for the previous two years: $\text{FEPS}_{t+\tau} = \text{FEPS}_{t+\tau-1} \times (\text{FEPS}_{t+\tau-1}/\text{FEPS}_{t+\tau-2})$, when $\text{FEPS}_{t+\tau-2}$ and $\text{FEPS}_{t+\tau-1}$ are both positive.

Instead of analysts' earnings forecasts, we also use pooled cross-sectional regressions in equation (A6) to forecast future earnings for up to five years ahead. In June of each year t, we estimate the regression using the pooled panel data from the previous ten years. With a four-month information lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. We forecast the expected earnings as the estimated regression coefficients times the latest values of the (unwinsorized) predictors from the fiscal year ending between March of $t-1$ and February of t. We implement this modified CT procedure at the firm level. Finally, we also use annual cross-sectional regressions in equation (A7) to forecast future Roe for up to five years ahead. In June of each year t, we estimate the regression using the previous ten years of data. With a four-month information lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. We directly forecast the expected Roe, $E_t[\text{Roe}_{t+\tau}]$, as the average cross-sectional regression coefficients times the latest values of the predictors from fiscal years ending between March of $t-1$ and February of t. We implement the modified CT procedure at the firm level.

The Ohlson and Juettner-Nauroth (2005, OJ) Procedure. At the end of June in each year t, we construct the IRR as:

$$\text{IRR} = A + \sqrt{A^2 + \frac{E_t[Y_{t+1}]}{P_t} \times (g - (\gamma - 1))}, \tag{A10}$$

in which

$$A \equiv \frac{1}{2}\left((\gamma - 1) + \frac{E_t[D_{t+1}]}{P_t}\right), \tag{A11}$$

$$g \equiv \frac{1}{2}\left(\frac{E_t[Y_{t+3}] - E_t[Y_{t+2}]}{E_t[Y_{t+2}]} + \frac{E_t[Y_{t+5}] - E_t[Y_{t+4}]}{E_t[Y_{t+4}]}\right). \tag{A12}$$

P_t is the market equity in year t, $E_t[Y_{t+\tau}]$ is the expected earnings for year $t+\tau$ based on information available in t, and $E_t[D_{t+1}]$ is the expected dividends for year $t+1$.

Expected earnings are based on analyst forecasts from IBES or forecasts from regression models. Expected dividends are expected earnings times the current dividend payout ratio, which is computed as dividends (Compustat annual item DVC) divided by earnings (item IB) for profitable firms, or dividends divided by 6% of total assets (item AT) for firms with zero or negative earnings. We follow Gode and Mohanram (2003) and use the average of forecasted near-term growth rate and five-year growth rate as an estimate of g. We require $E_t[Y_{t+2}]$ and $E_t[Y_{t+4}]$ to be positive so that g is well defined. Following Gode and Mohanram, we implement the OJ model on the per share basis with analysts' earnings forecasts. We measure P_t as the June-end share price from CRSP. At the end of June in year t, the expected earnings per share for year $t+\tau$ is the consensus mean forecast from IBES (unadjusted file, item MEANEST) for year $t+\tau$ (fiscal period indicator $= \tau$) reported in June of t. We require the availability of earnings forecast for years $t+1$ and $t+2$. When the forecast after year $t+2$ is not available, we use the long-term growth rate (item LTG) to construct it as: $\text{FEPS}_{t+\tau} = \text{FEPS}_{t+\tau-1} \times (1+\text{LTG})$. If the long-term growth rate is missing, we replace it with the growth rate implied by the forecasts for the previous two years: $\text{FEPS}_{t+\tau} = \text{FEPS}_{t+\tau-1} \times (\text{FEPS}_{t+\tau-1}/\text{FEPS}_{t+\tau-2})$, when $\text{FEPS}_{t+\tau-2}$ and $\text{FEPS}_{t+\tau-1}$ are both positive.

Instead of analysts' earnings forecasts, we also use pooled cross-sectional regressions in equation (A6) to forecast future earnings for up to five years ahead. In June of each year t, we estimate the regression using the pooled panel data from the previous ten years. With a four-month information lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. We construct the expected earnings as the estimated regression coefficients times the latest values of the (unwinsorized) predictors from the fiscal year ending between March of $t-1$ and February of t. We implement the modified OJ procedure at the firm level.

We also use annual cross-sectional regressions in equation (A7) to forecast future Roe for up to five years ahead. In June of each year t, we estimate the regression using the previous ten years of data. With a four-month information lag, the accounting data are from fiscal years ending between March of $t-10$ and February of t. We forecast the expected Roe, $E_t[\text{Roe}_{t+\tau}]$, as the average cross-sectional regression coefficients times the latest values of the predictors from fiscal years ending between March of $t-1$ and February of t. Expected earnings are then constructed as: $E_t[Y_{t+\tau}] = E_t[\text{Roe}_{t+\tau}] \times Be_{t+\tau-1}$, in which $Be_{t+\tau-1}$ is the book equity in year $t+\tau-1$. We measure current book equity Be_t based on the latest accounting data from the fiscal year ending in March of $t-1$ to February of t, and impute future book equity by applying clean surplus accounting recursively. We implement the modified OJ procedure at the firm level.