

# Internet Appendix: “ $q^5$ ” (for Online Publication Only)

Kewei Hou\*  
The Ohio State University  
and CAFR

Chen Xue<sup>‡</sup>  
University of Cincinnati

Haitao Mo<sup>†</sup>  
Louisiana State University

Lu Zhang<sup>§</sup>  
The Ohio State University  
and NBER

July 2018

## Abstract

This Online Appendix provides detailed derivation, variable definition, and portfolio construction for the manuscript titled “ $q^5$ .”

---

\*Fisher College of Business, The Ohio State University, 820 Fisher Hall, 2100 Neil Avenue, Columbus OH 43210; and China Academy of Financial Research (CAFR). Tel: (614) 292-0552. E-mail: hou.28@osu.edu.

<sup>†</sup>E. J. Ourso College of Business, Louisiana State University, 2931 Business Education Complex, Baton Rouge, LA 70803. Tel: (225) 578-0648. E-mail: haitaomo@lsu.edu.

<sup>‡</sup>Lindner College of Business, University of Cincinnati, 405 Lindner Hall, Cincinnati, OH 45221. Tel: (513) 556-7078. E-mail: xuecx@ucmail.uc.edu.

<sup>§</sup>Fisher College of Business, The Ohio State University, 760A Fisher Hall, 2100 Neil Avenue, Columbus OH 43210; and NBER. Tel: (614) 292-8644. E-mail: zhanglu@fisher.osu.edu.

## A Proof That Stock and Investment Returns Are Equal

This proof follows Liu, Whited, and Zhang (2009), who in turn construct their proof based on Restroy and Rockinger (1994). Let  $q_{it}$  be the Lagrangian multiplier associated with the capital accumulation equation  $A_{it+1} = (1 - \delta)A_{it} + I_{it}$ . Form the Lagrangian function for the equity value maximization problem of firm  $i$ :

$$\begin{aligned} \mathcal{L} = & \dots + X_{it}A_{it} - \frac{a}{2} \left( \frac{I_{it}}{A_{it}} \right)^2 A_{it} - I_{it} - q_{it}(A_{it+1} - (1 - \delta)A_{it} - I_{it}) \\ + E_t \left[ M_{t+1} \left[ X_{it+1}A_{it+1} - \frac{a}{2} \left( \frac{I_{it+1}}{A_{it+1}} \right)^2 A_{it+1} - I_{it+1} - q_{it+1}(A_{it+2} - (1 - \delta)A_{it+1} - I_{it+1}) \right] \right] + \dots \end{aligned} \quad (\text{A.1})$$

The first-order conditions with respect to  $I_{it}$  and  $A_{it+1}$  are, respectively,

$$q_{it} = 1 + a \frac{I_{it}}{A_{it}}; \quad (\text{A.2})$$

$$q_{it} = E_t \left[ M_{t+1} \left[ X_{it+1} + \frac{a}{2} \left( \frac{I_{it+1}}{A_{it+1}} \right)^2 + (1 - \delta)q_{it+1} \right] \right]. \quad (\text{A.3})$$

To show the marginal  $q$  equals the average  $q$ , we start with  $P_{it} + D_{it} = V_{it}$  and expand  $V_{it}$ :

$$\begin{aligned} P_{it} + X_{it}A_{it} - \frac{a}{2} \left( \frac{I_{it}}{A_{it}} \right)^2 A_{it} - I_{it} &= X_{it}A_{it} - a \frac{I_{it}}{A_{it}} I_{it} + \frac{a}{2} \left( \frac{I_{it}}{A_{it}} \right)^2 A_{it} - I_{it} \\ - q_{it}(A_{it+1} - (1 - \delta)A_{it} - I_{it}) &+ E_t \left[ M_{t+1} \left( X_{it+1}A_{it+1} - a \frac{I_{it+1}}{A_{it+1}} I_{it+1} \right. \right. \\ \left. \left. + \frac{a}{2} \left( \frac{I_{it+1}}{A_{it+1}} \right)^2 A_{it+1} - I_{it+1} - q_{it+1}(A_{it+2} - (1 - \delta)A_{it+1} - I_{it+1}) + \dots \right) \right]. \end{aligned} \quad (\text{A.4})$$

Substituting equations (A.2) and (A.3), and using the linear homogeneity of adjustment costs:

$$P_{it} = \left( 1 + a \frac{I_{it}}{A_{it}} \right) I_{it} + q_{it}(1 - \delta)A_{it} = q_{it}A_{it+1}. \quad (\text{A.5})$$

Finally, we are ready to show the equivalence between the stock and the investment returns:

$$\begin{aligned} r_{it+1}^S &= \frac{P_{it+1} + X_{it+1}A_{it+1} - (a/2) (I_{it+1}/A_{it+1})^2 A_{it+1} - I_{it+1}}{P_{it}} \\ &= \frac{q_{it+1}(I_{it+1} + (1 - \delta)A_{it+1}) + X_{it+1}A_{it+1} - (a/2) (I_{it+1}/A_{it+1})^2 A_{it+1} - I_{it+1}}{q_{it}A_{it+1}} \\ &= \frac{(1 - \delta)q_{it+1} + X_{it+1} + (a/2) (I_{it+1}/A_{it+1})^2}{q_{it}} = r_{it+1}^I, \end{aligned} \quad (\text{A.6})$$

in which the second equality follows from equation (A.2), and the third equality follows from the linear homogeneity of the adjustment costs function. Let  $\Phi_{it} \equiv (a/2) (I_{it}/A_{it})^2 A_{it}$ , its linear homogeneity means that  $\Phi_{it} = I_{it}\partial\Phi_{it}/\partial I_{it} + K_{it}\partial\Phi_{it}/\partial K_{it}$ .

## B Variable Definition and Portfolio Construction

We follow the variable definition and portfolio construction in Hou, Xue, and Zhang (2017). When forming testing deciles, we always use NYSE breakpoints and value-weight decile returns.

### B.1 Momentum

#### B.1.1 Sue1, Standardized Unexpected Earnings

Per Foster, Olsen, and Shevlin (1984), Sue denotes Standardized Unexpected Earnings, and is calculated as the change in split-adjusted quarterly earnings per share (Compustat quarterly item EPSPXQ divided by item AJEXQ) from its value four quarters ago divided by the standard deviation of this change in quarterly earnings over the prior eight quarters (six quarters minimum). At the beginning of each month  $t$ , we split all NYSE, Amex, and NASDAQ stocks into deciles based on their most recent past Sue. Before 1972, we use the most recent Sue computed with quarterly earnings from fiscal quarters ending at least four months prior to the portfolio formation. Starting from 1972, we use Sue computed with quarterly earnings from the most recent quarterly earnings announcement dates (Compustat quarterly item RDQ). For a firm to enter our portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Sue to be within six months prior to the portfolio formation. We do so to exclude stale information on earnings. To avoid potentially erroneous records, we also require the earnings announcement date to be after the corresponding fiscal quarter end. Monthly portfolio returns are calculated for the current month  $t$ , and the portfolios are rebalanced at the beginning of month  $t + 1$ .

#### B.1.2 Abr1, Abr6, and Abr12, Cumulative Abnormal Returns Around Earnings Announcement Dates

We calculate cumulative abnormal stock return (Abr) around the latest quarterly earnings announcement date (Compustat quarterly item RDQ) (Chan, Jegadeesh, and Lakonishok 1996)):

$$\text{Abr}_i = \sum_{d=-2}^{+1} r_{id} - r_{md}, \quad (\text{B.1})$$

in which  $r_{id}$  is stock  $i$ 's return on day  $d$  (with the earnings announced on day 0) and  $r_{md}$  is the market index return. We cumulate returns until one (trading) day after the announcement date to account for the one-day-delayed reaction to earnings news.  $r_{md}$  is the value-weighted market return for the Abr deciles with NYSE breakpoints and value-weighted returns.

At the beginning of each month  $t$ , we split all stocks into deciles based on their most recent past Abr. For a firm to enter our portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Abr to be within six months prior to the portfolio formation. We do so to exclude stale information on earnings. To avoid potentially erroneous records, we also require the earnings announcement date to be after the corresponding fiscal quarter end. Monthly decile returns are calculated for the current month  $t$  (Abr1), and, separately, from month  $t$  to  $t + 5$  (Abr6) and from month  $t$  to  $t + 11$  (Abr12). The deciles are rebalanced monthly. The six-month holding period for Abr6 means that for a given decile in each month there exist six sub-deciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the sub-decile returns as the monthly return of the Abr6 decile. Because quarterly earnings announcement dates are largely unavailable before 1972, the Abr portfolios start in January 1972.

### B.1.3 Re1 and Re6, Revisions in Analyst Earnings Forecasts

Following Chan, Jegadeesh, and Lakonishok (1996), we measure earnings surprise as the revisions in analysts' forecasts of earnings obtained from the Institutional Brokers' Estimate System (IBES). Because analysts' forecasts are not necessarily revised each month, we construct a six-month moving average of past changes in analysts' forecasts:

$$RE_{it} = \sum_{\tau=1}^6 \frac{f_{it-\tau} - f_{it-\tau-1}}{p_{it-\tau-1}}, \quad (\text{B.2})$$

in which  $f_{it-\tau}$  is the consensus mean forecast (IBES unadjusted file, item MEANEST) issued in month  $t - \tau$  for firm  $i$ 's current fiscal year earnings (fiscal period indicator = 1), and  $p_{it-\tau-1}$  is the prior month's share price (unadjusted file, item PRICE). We require both earnings forecasts and share prices to be denominated in US dollars (currency code = USD). We also adjust for any stock splits and require a minimum of four monthly forecast changes when constructing Re. At the beginning of each month  $t$ , we split all stocks into deciles based on their Re. Monthly decile returns are calculated for the current month  $t$  (Re1), and, separately, from month  $t$  to  $t + 5$  (Re6). The deciles are rebalanced monthly. The six-month holding period for Re6 means that for a given decile in each month there exist six sub-deciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the sub-decile returns as the monthly return of the Re6 decile. Because analyst forecast data start in January 1976, the Re portfolios start in July 1976.

### B.1.4 $R^61$ , $R^66$ , and $R^612$ , Prior Six-month Returns

At the beginning of each month  $t$ , we split all stocks into deciles based on their prior six-month returns from month  $t - 7$  to  $t - 2$ . Skipping month  $t - 1$ , we calculate monthly decile returns, separately, for month  $t$  ( $R^61$ ), from month  $t$  to  $t + 5$  ( $R^66$ ), and from month  $t$  to  $t + 11$  ( $R^612$ ). The deciles are rebalanced at the beginning of month  $t + 1$ . The holding period that is longer than one month as in, for instance,  $R^66$ , means that for a given decile in each month there exist six sub-deciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the sub-deciles returns as the monthly return of the  $R^66$  decile. We do not impose a price screen to exclude stocks with prices per share below \$5 as in Jegadeesh and Titman (1993). These stocks are mostly microcaps. Value-weighting returns assigns only tiny weights to these stocks, which in turn do not need to be excluded.

### B.1.5 $R^{111}$ and $R^{116}$ , Prior 11-month Returns

We split all stocks into deciles at the beginning of each month  $t$  based on their prior 11-month returns from month  $t - 12$  to  $t - 2$ . Skipping month  $t - 1$ , we calculate monthly decile returns for month  $t$  ( $R^{111}$ ), and separately, from month  $t$  to  $t + 5$  ( $R^{116}$ ). All the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period that is longer than one month as in  $R^{116}$  means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdecile returns as the monthly return of the  $R^{116}$  decile. Because we exclude financial firms, these decile returns are different from those posted on Kenneth French's Web site.

### B.1.6 Im1, Im6, and Im12, Industry Momentum

We start with the Fama-French (1997) 49-industry classifications. Excluding financial firms from the sample leaves 45 industries. At the beginning of each month  $t$ , we sort industries based on their prior six-month value-weighted returns from  $t - 6$  to  $t - 1$ . Following Moskowitz and Grinblatt (1999), we do not skip month  $t - 1$ . We form nine portfolios ( $9 \times 5 = 45$ ), each of which contains five different industries. We define the return of a given portfolio as the simple average of the five industry returns within the portfolio. We calculate portfolio returns for the nine portfolios for the current month  $t$  (Im1), from month  $t$  to  $t + 5$  (Im6), and from month  $t$  to  $t + 11$  (Im12). The portfolios are rebalanced at the beginning of  $t + 1$ . The holding period that is longer than one month as in, for instance, Im6, means that for a given portfolio in each month there exist six subportfolios, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subportfolio returns as the monthly return of the Im6 portfolio.

### B.1.7 Rs1, Revenue Surprises

Following Jegadeesh and Livnat (2006), we measure revenue surprises (Rs) as changes in revenue per share (Compustat quarterly item SALEQ/(item CSHPRQ times item AJEXQ)) from its value four quarters ago divided by the standard deviation of this change in quarterly revenue per share over the prior eight quarters (six quarters minimum). At the beginning of each month  $t$ , we split stocks into deciles based on their most recent past Rs. Before 1972, we use the most recent Rs computed with quarterly revenue from fiscal quarters ending at least four months prior to the portfolio formation. Starting from 1972, we use Rs computed with quarterly revenue from the most recent quarterly earnings announcement dates (Compustat quarterly item RDQ). Jegadeesh and Livnat find that quarterly revenue data are generally available when earnings are announced. For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Rs to be within six months prior to the portfolio formation. This restriction is imposed to exclude stale revenue information. To avoid potentially erroneous records, we also require the earnings announcement date to be after the corresponding fiscal quarter end. Monthly deciles returns are calculated for the current month  $t$  (Rs1), and the deciles are rebalanced at the beginning of month  $t + 1$ .

### B.1.8 dEf1, dEf6, and dEf12, Changes in Analyst Earnings Forecasts

Following Hawkins, Chamberlin, and Daniel (1984), we define  $dEf \equiv (f_{it-1} - f_{it-2}) / (0.5|f_{it-1}| + 0.5|f_{it-2}|)$ , in which  $f_{it-1}$  is the consensus mean forecast (IBES unadjusted file, item MEANEST) issued in month  $t - 1$  for firm  $i$ 's current fiscal year earnings (fiscal period indicator = 1). We require earnings forecasts to be denominated in US dollars (currency code = USD). We also adjust for any stock splits between months  $t - 2$  and  $t - 1$  when constructing dEf. At the beginning of each month  $t$ , we sort stocks into deciles on the prior month dEf, and calculate returns for the current month  $t$  (dEf1), from month  $t$  to  $t + 5$  (dEf6), and from month  $t$  to  $t + 11$  (dEf12). The deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in, for instance, dEf6, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the dEf6 decile. Because analyst forecast data start in January 1976, the dEf portfolios start in March 1976.

### **B.1.9 Nei1, The Number of Quarters with Consecutive Earnings Increase**

We follow Barth, Elliott, and Finn (1999) and Green, Hand, and Zhang (2013) in measuring Nei as the number of consecutive quarters (up to eight quarters) with an increase in earnings (Compustat quarterly item IBQ) over the same quarter in the prior year. At the beginning of each month  $t$ , we sort stocks into nine portfolios (with  $\text{Nei} = 0, 1, 2, \dots, 7$ , and 8, respectively) based on their most recent past Nei. Before 1972, we use Nei computed with quarterly earnings from fiscal quarters ending at least four months prior to the portfolio formation. Starting from 1972, we use Nei computed with earnings from the most recent quarterly earnings announcement dates (Compustat quarterly item RDQ). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Nei 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. We calculate monthly portfolio returns for the current month  $t$  (Nei1), and the deciles are rebalanced at the beginning of month  $t + 1$ . For sufficient data coverage, the Nei portfolios start in January 1969.

### **B.1.10 52w6, 52-week High**

At the beginning of each month  $t$ , we split stocks into deciles based on 52w, which is the ratio of its split-adjusted price per share at the end of month  $t - 1$  to its highest (daily) split-adjusted price per share during the 12-month period ending on the last day of month  $t - 1$ . Monthly decile returns are calculated from month  $t$  to  $t + 5$  (52w6), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the 52w6 decile. Because a disproportionately large number of stocks can reach the 52-week high at the same time and have 52w equal to one, we use only 52w smaller than one to form the portfolio breakpoints. Doing so helps avoid missing portfolio observations.

### **B.1.11 $\epsilon^6$ 6 and $\epsilon^6$ 12, Six-month Residual Momentum**

We split all stocks into deciles at the beginning of each month  $t$  based on their prior six-month average residual returns from month  $t - 7$  to  $t - 2$  scaled by their standard deviation over the same period. Skipping month  $t - 1$ , we calculate monthly decile returns from month  $t$  to  $t + 5$  ( $\epsilon^6$ 6) and from month  $t$  to  $t + 11$  ( $\epsilon^6$ 12). Residual returns are estimated each month for all stocks over the prior 36 months from month  $t - 36$  to month  $t - 1$  from regressing stock excess returns on the Fama-French three factors. To reduce the noisiness of the estimation, we require returns to be available for all prior 36 months. All the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period that is longer than one month as in  $\epsilon^6$ 6 means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdecile returns as the monthly return of the  $\epsilon^6$ 6 decile.

### **B.1.12 $\epsilon^{11}$ 1, $\epsilon^{11}$ 6, and $\epsilon^{11}$ 12, 11-month Residual Momentum**

We split all stocks into deciles at the beginning of each month  $t$  based on their prior 11-month residual returns from month  $t - 12$  to  $t - 2$  scaled by their standard deviation over the same period. Skipping month  $t - 1$ , we calculate monthly decile returns for month  $t$  ( $\epsilon^{11}$ 1), from month  $t$  to  $t + 5$  ( $\epsilon^{11}$ 6), and from month  $t$  to  $t + 11$  ( $\epsilon^{11}$ 12). Residual returns are estimated each month for all stocks

over the prior 36 months from month  $t - 36$  to month  $t - 1$  from regressing stock excess returns on the Fama-French three factors. To reduce the noisiness of the estimation, we require returns to be available for all prior 36 months. All the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period that is longer than 1 month as in  $\epsilon^{116}$  means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdecile returns as the monthly return of the  $\epsilon^{116}$  decile.

### **B.1.13 Sm1, Segment Momentum**

Following Cohen and Lou (2012), we extract firms' segment accounting and financial information from Compustat segment files. Industries are based on two-digit SIC codes. Standalone firms are those that operate in only one industry with segment sales, reported in Compustat segment files, accounting for more than 80% of total sales reported in Compustat annual files. Conglomerate firms are those that operating in more than one industry with aggregate sales from all reported segments accounting for more than 80% of total sales.

At the end of June of each year, we form a pseudo-conglomerate for each conglomerate firm. The pseudo-conglomerate is a portfolio of the conglomerate's industry segments constructed with solely the standalone firms in each industry. The segment portfolios (value-weighted across standalone firms) are then weighted by the percentage of sales contributed by each industry segment within the conglomerate. At the beginning of each month  $t$  (starting in July), using segment information from the previous fiscal year, we sort all conglomerate firms into deciles based on the returns of their pseudo-conglomerate portfolios in month  $t - 1$ . Monthly deciles are calculated for month  $t$  (Sm1), and the deciles are rebalanced at the beginning of month  $t + 1$ . Because the segment data start in 1976, the Sm portfolios start in July 1977.

### **B.1.14 Ilr1, Ilr6, and Ilr12, Industry Lead-lag Effect in Prior Returns**

We start with the Fama-French (1997) 49-industry classifications. Excluding financial firms from the sample leaves 45 industries. At the beginning of each month  $t$ , we sort industries based on the month  $t - 1$  value-weighted return of the portfolio consisting of the 30% biggest (market equity) firms within a given industry. We form nine portfolios ( $9 \times 5 = 45$ ), each of which contains five different industries. We define the return of a given portfolio as the simple average of the five value-weighted industry returns within the portfolio. The nine portfolio returns are calculated for the current month  $t$  (Ilr1), from month  $t$  to  $t + 5$  (Ilr6), and from month  $t$  to  $t + 11$  (Ilr12), and the portfolios are rebalanced at the beginning of month  $t + 1$ . The holding period that is longer than one month as in, for instance, Ilr6, means that for a given portfolio in each month there exist six subportfolios, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subportfolio returns as the monthly return of the Ilr6 portfolio.

### **B.1.15 Ile1, Industry Lead-lag Effect in Earnings Surprises**

We start with the Fama-French (1997) 49-industry classifications. Excluding financial firms from the sample leaves 45 industries. We calculate Standardized Unexpected Earnings, Sue, as the change in split-adjusted quarterly earnings per share (Compustat quarterly item EPSPXQ divided by item AJEXQ) from its value four quarters ago divided by the standard deviation of this change in quarterly earnings over the prior eight quarters (six quarters minimum). At the beginning of each month  $t$ , we sort industries based on their most recent Sue averaged across the 30% biggest firms

within a given industry.<sup>1</sup> To mitigate the impact of outliers, we winsorize Sue at the 1st and 99th percentiles of its distribution each month. We form nine portfolios ( $9 \times 5 = 45$ ), each of which contains five different industries. We define the return of a given portfolio as the simple average of the five value-weighted industry returns within the portfolio. The nine portfolio returns are calculated for the current month  $t$  (Ile1), and the portfolios are rebalanced at the beginning of month  $t + 1$ .

### B.1.16 Cm1 and Cm12, Customer Momentum

Following Cohen and Frazzini (2008), we extract firms’ principal customers from Compustat segment files. For each firm we determine whether the customer is another company listed on the CRSP/Compustat tape, and we assign it the corresponding CRSP permno number. At the end of June of each year  $t$ , we form a customer portfolio for each firm with identifiable firm-customer relations for the fiscal year ending in calendar year  $t - 1$ . For firms with multiple customer firms, we form equal-weighted customer portfolios. The customer portfolio returns are calculated from July of year  $t$  to June of  $t + 1$ , and the portfolios are rebalanced in June.

At the beginning of each month  $t$ , we sort all stocks into quintiles based on their customer portfolio returns, Cm, in month  $t - 1$ . We do not form deciles because a disproportionate number of firms can have the same Cm, which leads to fewer than ten portfolios in some months. Monthly quintile returns are calculated for month  $t$  (Cm1) and from month  $t$  to  $t + 11$  (Cm12), and the quintiles are rebalanced at the beginning of month  $t + 1$ . The holding period that is longer than one month in Cm12 means that for a given quintile in each month there exist 12 subquintiles, each of which is initiated in a different month in the prior 12-month period. We take the simple average of the subquintile returns as the monthly return of the Cm12 quintile. For sufficient data coverage, we start the Cm portfolios in July 1979.

### B.1.17 Sim1, Cim1, Cim6, and Cim12, Supplier (Customer) industries Momentum

Following Menzly and Ozbas (2010), we use Benchmark Input-Output Accounts at the Bureau of Economic Analysis (BEA) to identify supplier and customer industries for a given industry. BEA Surveys are conducted roughly once every five years in 1958, 1963, 1967, 1972, 1977, 1982, 1987, 1992, 1997, 2002, and 2007. We delay the use of any data from a given survey until the end of the year in which the survey is publicly released during 1964, 1969, 1974, 1979, 1984, 1991, 1994, 1997, 2002, 2007, and 2013, respectively. The BEA industry classifications are based on SIC codes in the surveys from 1958 to 1992 and based on NAICS codes afterwards. In the surveys from 1997 to 2007, we merge three separate industry accounts, 2301, 2302, and 2303 into a single account. We also merge “Housing” (HS) and “Other Real Estate” (ORE) in the 2007 Survey. In the surveys from 1958 to 1992, we merge industry account pairs 1–2, 5–6, 9–10, 11–12, 20–21, and 33–34. We also merge industry account pairs 22–23 and 44–45 in the 1987 and 1992 surveys. We drop miscellaneous industry accounts related to government, import, and inventory adjustments.

At the end of June of each year  $t$ , we assign each stock to an BEA industry based on its reported SIC or NAICS code in Compustat (fiscal year ending in  $t - 1$ ) or CRSP (June of  $t$ ). Monthly value-weighted industry returns are calculated from July of year  $t$  to June of  $t + 1$ , and the industry

---

<sup>1</sup>Before 1972, we use the most recent Sue with earnings from fiscal quarters ending at least four months prior to the portfolio month. Starting from 1972, we use Sue with earnings from the most recent quarterly earnings announcement dates (Compustat quarterly item RDQ). For a firm to enter our portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Sue to be within six months prior to the portfolio month. We also require the earnings announcement date to be after the corresponding fiscal quarter end.

portfolios are rebalanced in June of  $t+1$ . For each industry, we further form two separate portfolios, the suppliers portfolio and the customers portfolios. The share of an industry’s total purchases from other industries is used to calculate the supplier industries portfolio returns, and the share of the industry’s total sales to other industries is used to calculate the customer industries portfolio returns.

At the beginning of each month  $t$ , we split industries into deciles based on the supplier portfolio returns, Sim, and separately, on the customer portfolio returns, Cim, in month  $t-1$ . We then assign the decile rankings of each industry to its member stocks. Monthly decile returns are calculated for month  $t$  (Sim1 and Cim1), from month  $t$  to  $t+5$  (Cim6), and from month  $t$  to  $t+11$  (Cim12), and the deciles are rebalanced at the beginning of month  $t+1$ . The holding period that is longer than one month as in Cim6 means that for a given decile in each month there exist six subdeciles, each initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Cim6 decile.

## **B.2 Value-versus-growth**

### **B.2.1 Bm, Book-to-market Equity**

At the end of June of each year  $t$ , we split stocks into deciles based on Bm, which is the book equity for the fiscal year ending in calendar year  $t-1$  divided by the market equity (from CRSP) at the end of December of  $t-1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Bm. Monthly decile returns are calculated from July of year  $t$  to June of  $t+1$ , and the deciles are rebalanced in June of  $t+1$ . Following Davis, Fama, and French (2000), we measure book equity as stockholders’ book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders’ equity is the value reported by Compustat (item SEQ), if it is available. If not, we measure stockholders’ equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock.

### **B.2.2 Bmj, Book-to-June-end Market Equity**

Following Asness and Frazzini (2013), at the end of June of each year  $t$ , we sort stocks into deciles based on Bmj, which is book equity per share for the fiscal year ending in calendar year  $t-1$  divided by share price (from CRSP) at the end of June of  $t$ . We adjust for any stock splits between the fiscal year end and the end of June. Book equity per share is book equity divided by the number of shares outstanding (Compustat annual item CSHO). Following Davis, Fama, and French (2000), we measure book equity as stockholders’ book equity, plus balance sheet deferred taxes and investment tax credit (item TXDITC) if available, minus the book value of preferred stock. Stockholders’ equity is the value reported by Compustat (item SEQ), if it is available. If not, we measure stockholders’ equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock. Monthly decile returns are calculated from July of year  $t$  to June of  $t+1$ , and the deciles are rebalanced in June of  $t+1$ .

### B.2.3 Bm<sup>q</sup>12, Quarterly Book-to-market Equity

At the beginning of each month  $t$ , we split stocks into deciles based on Bm<sup>q</sup>, which is the book equity for the latest fiscal quarter ending at least four months ago divided by the market equity (from CRSP) at the end of month  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Bm<sup>q</sup>. We calculate decile returns from month  $t$  to  $t + 11$  (Bm<sup>q</sup>12), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month means that for a given decile in each month there exist 12 subdeciles, each of which is initiated in a different month in the prior 12 months. We take the simple average of the subdecile returns as the monthly return of the Bm<sup>q</sup>12 decile. Book equity is shareholders' equity, plus balance sheet deferred taxes and investment tax credit (Compustat quarterly item TXDITCQ) if available, minus the book value of preferred stock (item PSTKQ). Depending on availability, we use stockholders' equity (item SEQQ), or common equity (item CEQQ) plus the book value of preferred stock, or total assets (item ATQ) minus total liabilities (item LTQ) in that order as shareholders' equity.

Before 1972, the sample coverage is limited for quarterly book equity in Compustat quarterly files. We expand the coverage by using book equity from Compustat annual files as well as by imputing quarterly book equity with clean surplus accounting. Specifically, whenever available we first use quarterly book equity from Compustat quarterly files. We then supplement the coverage for fiscal quarter four with annual book equity from Compustat annual files. Following Davis, Fama, and French (2000), we measure annual book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ), if available. If not, stockholders' equity is the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock.

If both approaches are unavailable, we apply the clean surplus relation to impute the book equity. Specifically, we impute the book equity for quarter  $t$  forward based on book equity from prior quarters. Let  $BEQ_{t-j}$ ,  $1 \leq j \leq 4$  denote the latest available quarterly book equity as of quarter  $t$ , and  $IBQ_{t-j+1,t}$  and  $DVQ_{t-j+1,t}$  be the sum of quarterly earnings and quarterly dividends from quarter  $t-j+1$  to  $t$ , respectively.  $BEQ_t$  can then be imputed as  $BEQ_{t-j} + IBQ_{t-j+1,t} - DVQ_{t-j+1,t}$ . We do not use prior book equity from more than four quarters ago (i.e.,  $1 \leq j \leq 4$ ) to reduce imputation errors. Quarterly earnings are income before extraordinary items (Compustat quarterly item IBQ). Quarterly dividends are zero if dividends per share (item DVPSXQ) are zero. Otherwise, total dividends are dividends per share times beginning-of-quarter shares outstanding adjusted for stock splits during the quarter. Shares outstanding are from Compustat (quarterly item CSHOQ supplemented with annual item CSHO for fiscal quarter four) or CRSP (item SHROUT), and the share adjustment factor is from Compustat (quarterly item AJEXQ supplemented with annual item AJEX for fiscal quarter four) or CRSP (item CFACSHR). Because we use quarterly book equity at least four months after the fiscal quarter end, all the Compustat data used in the imputation are at least four-month lagged prior to the portfolio formation. In addition, we do not impute quarterly book equity backward using future earnings and book equity information to avoid look-ahead bias.

### B.2.4 Rev6 and Rev12, Reversal

To capture the De Bondt and Thaler (1985) long-term reversal (Rev) effect, at the beginning of each month  $t$ , we split stocks into deciles based on the prior returns from month  $t - 60$  to  $t - 13$ . Monthly

decile returns are computed from month  $t$  to  $t + 5$  (Rev6) and from month  $t$  to  $t + 11$  (Rev12). The deciles are rebalanced at the beginning of  $t + 1$ . The holding period longer than one month as in, for instance, Rev6, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdeciles returns as the monthly return of the Rev6 decile. To be included in a portfolio for month  $t$ , a stock must have a valid price at the end of  $t - 61$  and a valid return for  $t - 13$ . In addition, any missing returns from month  $t - 60$  to  $t - 14$  must be  $-99.0$ , which is the CRSP code for a missing ending price.

### **B.2.5 Ep, Earnings-to-price**

At the end of June of each year  $t$ , we split stocks into deciles based on earnings-to-price, Ep, which is income before extraordinary items (Compustat annual item IB) for the fiscal year ending in calendar year  $t - 1$  divided by the market equity (from CRSP) at the end of December of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Ep. Firms with non-positive earnings are excluded. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.2.6 Ep<sup>q</sup>1, Ep<sup>q</sup>6, and Ep<sup>q</sup>12, Quarterly Earnings-to-price**

At the beginning of each month  $t$ , we split stocks into deciles based on quarterly earnings-to-price, Ep<sup>q</sup>, which is income before extraordinary items (Compustat quarterly item IBQ) divided by the market equity (from CRSP) at the end of month  $t - 1$ . Before 1972, we use quarterly earnings from fiscal quarters ending at least four months prior to the portfolio formation. Starting from 1972, we use quarterly earnings from the most recent quarterly earnings announcement dates (item RDQ). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent quarterly earnings 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. Firms with non-positive earnings are excluded. For firms with more than one share class, we merge the market equity for all share classes before computing Ep<sup>q</sup>. We calculate decile returns for the current month  $t$  (Ep<sup>q</sup>1), from month  $t$  to  $t + 5$  (Ep<sup>q</sup>6), and from month  $t$  to  $t + 11$  (Ep<sup>q</sup>12), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in, for instance, Ep<sup>q</sup>6, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Ep<sup>q</sup>6 decile.

### **B.2.7 Cp, Cash Flow-to-price**

At the end of June of each year  $t$ , we split stocks into deciles based on cash flow-to-price, Cf, which is cash flows for the fiscal year ending in calendar year  $t - 1$  divided by the market equity (from CRSP) at the end of December of  $t - 1$ . Cash flows are income before extraordinary items (Compustat annual item IB) plus depreciation (item DP)). For firms with more than one share class, we merge the market equity for all share classes before computing Cp. Firms with non-positive cash flows are excluded. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.2.8 Cp<sup>q1</sup>, Cp<sup>q6</sup>, and Cp<sup>q12</sup>, Quarterly Cash Flow-to-price

At the beginning of each month  $t$ , we split stocks into deciles based on quarterly cash flow-to-price, Cp<sup>q</sup>, which is cash flows for the latest fiscal quarter ending at least four months ago divided by the market equity (from CRSP) at the end of month  $t - 1$ . Quarterly cash flows are income before extraordinary items (Compustat quarterly item IBQ) plus depreciation (item DPQ). For firms with more than one share class, we merge the market equity for all share classes before computing Cp<sup>q</sup>. Firms with non-positive cash flows are excluded. We calculate decile returns for the current month  $t$  (Ep<sup>q1</sup>), from month  $t$  to  $t + 5$  (Ep<sup>q6</sup>), and from month  $t$  to  $t + 11$  (Ep<sup>q12</sup>), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in, for instance, Ep<sup>q6</sup>, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Ep<sup>q6</sup> decile.

### B.2.9 Nop, Net Payout Yield

Per Boudoukh, Michaely, Richardson, and Roberts (2007), total payouts are dividends on common stock (Compustat annual item DVC) plus repurchases. Repurchases are the total expenditure on the purchase of common and preferred stocks (item PRSTKC) plus any reduction (negative change over the prior year) in the value of the net number of preferred stocks outstanding (item PSTKRV). Net payouts equal total payouts minus equity issuances, which are the sale of common and preferred stock (item SSTK) minus any increase (positive change over the prior year) in the value of the net number of preferred stocks outstanding (item PSTKRV).

At the end of June of each year  $t$ , we sort stocks into deciles based on net payouts for the fiscal year ending in calendar year  $t - 1$  divided by the market equity (from CRSP) at the end of December of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Nop. Firms with non-positive total payouts (zero net payouts) are excluded. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ . Because the data on total expenditure and the sale of common and preferred stocks start in 1971, the Nop portfolios start in July 1972.

### B.2.10 Em, Enterprise Multiple

Enterprise multiple, Em, is enterprise value divided by operating income before depreciation (Compustat annual item OIBDP). Enterprise value is the market equity plus the total debt (item DLC plus item DLTT) plus the book value of preferred stocks (item PSTKRV) minus cash and short-term investments (item CHE). At the end of June of each year  $t$ , we split stocks into deciles based on Em for the fiscal year ending in calendar year  $t - 1$ . The Market equity (from CRSP) is measured at the end of December of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Em. Firms with negative enterprise value or operating income before depreciation are excluded. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.2.11 Em<sup>q1</sup>, Em<sup>q6</sup>, and Em<sup>q12</sup>, Quarterly Enterprise Multiple

Em<sup>q</sup>, is enterprise value scaled by operating income before depreciation (Compustat quarterly item OIBDPQ). Enterprise value is the market equity plus total debt (item DLCQ plus item DLTTQ) plus the book value of preferred stocks (item PSTKQ) minus cash and short-term investments (item

CHEQ). At the beginning of each month  $t$ , we split stocks into deciles on  $Em^q$  for the latest fiscal quarter ending at least four months ago. The Market equity (from CRSP) is measured at the end of month  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing  $Em^q$ . Firms with negative enterprise value or operating income before depreciation are excluded. Monthly decile returns are calculated for the current month  $t$  ( $Em^{q1}$ ), from month  $t$  to  $t + 5$  ( $Em^{q6}$ ), and from month  $t$  to  $t + 11$  ( $Em^{q12}$ ), and the deciles are rebalanced at the beginning of  $t + 1$ . The holding period longer than one month as in  $Em^{q6}$  means that for a given decile in each month there exist six subdeciles, each initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the  $Em^{q6}$  decile. For sufficient data coverage, the  $EM^q$  portfolios start in January 1975.

### **B.2.12 Sp, Sales-to-price**

At the end of June of each year  $t$ , we sort stocks into deciles based on sales-to-price,  $Sp$ , which is sales (Compustat annual item SALE) for the fiscal year ending in calendar year  $t - 1$  divided by the market equity (from CRSP) at the end of December of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing  $Sp$ . Firms with non-positive sales are excluded. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.2.13 $Sp^{q1}$ , $Sp^{q6}$ , and $Sp^{q12}$ , Quarterly Sales-to-price**

At the beginning of each month  $t$ , we sort stocks into deciles based on quarterly sales-to-price,  $Sp^q$ , which is sales (Compustat quarterly item SALEQ) divided by the market equity at the end of month  $t - 1$ . Before 1972, we use quarterly sales from fiscal quarters ending at least four months prior to the portfolio formation. Starting from 1972, we use quarterly sales from the most recent quarterly earnings announcement dates (item RDQ). Sales are generally announced with earnings during quarterly earnings announcements (Jegadeesh and Livnat 2006). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent quarterly sales 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. Firms with non-positive sales are excluded. For firms with more than one share class, we merge the market equity for all share classes before computing  $Sp^q$ . Monthly decile returns are calculated for the current month  $t$  ( $Sp^{q1}$ ), from month  $t$  to  $t + 5$  ( $Sp^{q6}$ ), and from month  $t$  to  $t + 11$  ( $Sp^{q12}$ ), and the deciles are rebalanced at the beginning of  $t + 1$ . The holding period longer than one month as in  $Sp^{q6}$  means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the  $Sp^{q6}$  decile.

### **B.2.14 Ocp, Operating Cash Flow-to-price**

At the end of June of each year  $t$ , we sort stocks into deciles based on operating cash flows-to-price,  $Ocp$ , which is operating cash flows for the fiscal year ending in calendar year  $t - 1$  divided by the market equity (from CRSP) at the end of December of  $t - 1$ . Operating cash flows are measured as funds from operation (Compustat annual item FOPT) minus change in working capital (item WCAP) prior to 1988, and then as net cash flows from operating activities (item OANCF) starting from 1988. For firms with more than one share class, we merge the market equity for all share classes

before computing Ocp. Firms with non-positive operating cash flows are excluded. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ . Because the data on funds from operation start in 1971, the Ocp portfolios start in July 1972.

### B.2.15 Ocp<sup>q</sup>1, Quarterly Operating Cash Flow-to-price

At the beginning of each month  $t$ , we split stocks on quarterly operating cash flow-to-price, Ocp<sup>q</sup>, which is operating cash flows for the latest fiscal quarter ending at least four months ago divided by the market equity at the end of month  $t - 1$ . Operating cash flows are measured as the quarterly change in year-to-date funds from operation (Compustat quarterly item FOPTY) minus change in quarterly working capital (item WCAPQ) prior to 1988, and then as the quarterly change in year-to-date net cash flows from operating activities (item OANCFY) starting from 1988. For firms with more than one share class, we merge the market equity for all share classes before computing Ocp<sup>q</sup>. Firms with non-positive operating cash flows are excluded. Monthly decile returns are calculated for the current month  $t$ , and the deciles are rebalanced at the beginning of  $t + 1$ . Because the data on year-to-date funds from operation start in 1984, the Ocp<sup>q</sup> portfolios start in January 1985.

### B.2.16 Ir, Intangible Return

Following Daniel and Titman (2006), at the end of June of each year  $t$ , we perform the cross-sectional regression of each firm's past five-year log stock return on its five-year-lagged log book-to-market and five-year log book return:

$$r(t - 5, t) = \gamma_0 + \gamma_1 bm_{t-5} + \gamma_2 r^B(t - 5, t) + u_t \quad (\text{B.3})$$

in which  $r(t - 5, t)$  is the past five-year log stock return from the end of year  $t - 6$  to the end of  $t - 1$ ,  $bm_{t-5}$  is the five-year-lagged log book-to-market, and  $r^B(t - 5, t)$  is the five-year log book return. The five-year-lagged log book-to-market is computed as  $bm_{t-5} = \log(B_{t-5}/M_{t-5})$ , in which  $B_{t-5}$  is the book equity for the fiscal year ending in calendar year  $t - 6$  and  $M_{t-5}$  is the market equity (from CRSP) at the end of December of  $t - 6$ . For firms with more than one share class, we merge the market equity for all share classes before computing  $bm_{t-5}$ . The five-year log book return is computed as  $r^B(t - 5, t) = \log(B_t/B_{t-5}) + \sum_{s=t-5}^{t-1} (r_s - \log(P_s/P_{s-1}))$ , in which  $B_t$  is the book equity for the fiscal year ending in calendar year  $t - 1$ ,  $r_s$  is the stock return from the end of year  $s - 1$  to the end of year  $s$ , and  $P_s$  is the stock price per share at the end of year  $s$ . Following Davis, Fama, and French (2000), we measure book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ), if it is available. If not, we measure stockholders' equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock.

A firm's intangible return, Ir, is defined as its residual from the annual cross-sectional regression. At the end of June of each year  $t$ , we sort stocks based on Ir for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of year  $t + 1$ .

### B.2.17 Vhp and Vfp, (Analyst-based) Intrinsic Value-to-market

Following Frankel and Lee (1998), at the end of June of each year  $t$ , we implement the residual income model to estimate the intrinsic value:

$$Vh_t = B_t + \frac{(E_t[\text{Roe}_{t+1}] - r)}{(1+r)}B_t + \frac{(E_t[\text{Roe}_{t+2}] - r)}{(1+r)r}B_{t+1} \quad (\text{B.4})$$

$$Vf_t = B_t + \frac{(E_t[\text{Roe}_{t+1}] - r)}{(1+r)}B_t + \frac{(E_t[\text{Roe}_{t+2}] - r)}{(1+r)^2}B_{t+1} + \frac{(E_t[\text{Roe}_{t+3}] - r)}{(1+r)^2r}B_{t+2} \quad (\text{B.5})$$

in which  $Vh_t$  is the historical Roe-based intrinsic value and  $Vf_t$  is the analysts earnings forecast-based intrinsic value.  $B_t$  is the book equity (Compustat annual item CEQ) for the fiscal year ending in calendar year  $t - 1$ . Future book equity is computed using the clean surplus accounting:  $B_{t+1} = (1 + (1 - k)E_t[\text{Roe}_{t+1}])B_t$ , and  $B_{t+2} = (1 + (1 - k)E_t[\text{Roe}_{t+2}])B_{t+1}$ .  $E_t[\text{Roe}_{t+1}]$  and  $E_t[\text{Roe}_{t+2}]$  are the return on equity expected for the current and next fiscal years.  $k$  is the dividend payout ratio, measured as common stock dividends (item DVC) divided by earnings (item IBCOM) for the fiscal year ending in calendar year  $t - 1$ . For firms with negative earnings, we divide dividends by 6% of average total assets (item AT).  $r$  is a constant discount rate of 12%. When estimating  $Vh_t$ , we replace all Roe expectations with most recent  $\text{Roe}_t$ :  $\text{Roe}_t = Ni_t / [(B_t + B_{t-1})/2]$ , in which  $Ni_t$  is earnings for the fiscal year ending in  $t - 1$ , and  $B_t$  and  $B_{t-1}$  are the book equity from the fiscal years ending in  $t - 1$  and  $t - 2$ .

When estimating  $Vf_t$ , we use analyst earnings forecasts from IBES to construct Roe expectations. Let  $\text{Fy1}$  and  $\text{Fy2}$  be the one-year-ahead and two-year-ahead consensus mean forecasts (IBES unadjusted file, item MEANEST; fiscal period indicator = 1 and 2) reported in June of year  $t$ . Let  $s$  be the number of shares outstanding 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. Then  $E_t[\text{Roe}_{t+1}] = s\text{Fy1} / [(B_{t+1} + B_t)/2]$ , in which  $B_{t+1} = (1 + s(1 - k)\text{Fy1})B_t$ . Analogously,  $E_t[\text{Roe}_{t+2}] = s\text{Fy2} / [(B_{t+2} + B_{t+1})/2]$ , in which  $B_{t+2} = (1 + s(1 - k)\text{Fy2})B_{t+1}$ . Let  $\text{Ltg}$  denote the long-term earnings growth rate forecast from IBES (item MEANEST; fiscal period indicator = 0). Then  $E_t[\text{Roe}_{t+3}] = s\text{Fy2}(1 + \text{Ltg}) / [(B_{t+3} + B_{t+2})/2]$ , in which  $B_{t+3} = (1 + s(1 - k)\text{Fy2}(1 + \text{Ltg}))B_{t+2}$ . If  $\text{Ltg}$  is missing, we set  $E_t[\text{Roe}_{t+3}]$  to be  $E_t[\text{Roe}_{t+2}]$ . Firms are excluded if their expected Roe or dividend payout ratio is higher than 100%. We also exclude firms with negative book equity.

At the end of June of each year  $t$ , we sort stocks into deciles on the ratios of  $Vh$  and  $Vf$  scaled by the market equity (from CRSP) at the end of December of  $t - 1$ , denoted  $Vhp$  and  $Vfp$ , respectively. For firms with more than one share class, we merge the market equity for all share classes before computing intrinsic value-to-market. Firms with non-positive intrinsic value are excluded. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ . Because analyst forecast data start in 1976, the  $Vfp$  deciles start in July 1977.

### B.2.18 Ebp, Enterprise Book-to-price

Following Penman, Richardson, and Tuna (2007), we measure enterprise book-to-price,  $Ebp$ , as the ratio of the book value of net operating assets (net debt plus book equity) to the market value of net operating assets (net debt plus market equity). Net debt is financial liabilities minus financial assets. We measure financial liabilities as the sum of long-term debt (Compustat annual item DLTT), debt in current liabilities (item DLC), carrying value of preferred stock (item PSTK), and

preferred dividends in arrears (item DVPA, zero if missing), less preferred treasury stock (item TSTKP, zero if missing). We measure financial assets as cash and short-term investments (item CHE). Book equity is common equity (item CEQ) plus any preferred treasury stock (item TSTKP, zero if missing) less any preferred dividends in arrears (item DVPA, zero if missing). Market equity is the number of common shares outstanding times share price (from CRSP). At the end of June of each year  $t$ , we sort stocks into deciles based on Ebp from the fiscal year ending in calendar year  $t - 1$ . Market equity is measured at the end of December of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Ebp. We exclude firms with non-positive book or market value of net operating assets. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.2.19 Dur, Equity Duration

Following Dechow, Sloan, and Soliman (2004), we calculate firm-level equity duration, Dur, as:

$$\text{Dur} = \frac{\sum_{t=1}^T t \times \text{CD}_t / (1+r)^t}{\text{ME}} + \left( T + \frac{1+r}{r} \right) \frac{\text{ME} - \sum_{t=1}^T \text{CD}_t / (1+r)^t}{\text{ME}}, \quad (\text{B.6})$$

in which  $\text{CD}_t$  is the net cash distribution in year  $t$ , ME is market equity,  $T$  is the length of forecasting period, and  $r$  is the cost of equity. Market equity is price per share times shares outstanding (Compustat annual item PRCC\_F times item CSHO). Net cash distribution,  $\text{CD}_t = \text{BE}_{t-1}(\text{ROE}_t - g_t)$ , in which  $\text{BE}_{t-1}$  is the book equity at the end of year  $t - 1$ ,  $\text{ROE}_t$  is return on equity in year  $t$ , and  $g_t$  is the book equity growth in  $t$ . Following Dechow et al., we use autoregressive processes to forecast ROE and book equity growth in future years. We model ROE as a first-order autoregressive process with an autocorrelation coefficient of 0.57 and a long-run mean of 0.12, and the growth in book equity as a first-order autoregressive process with an autocorrelation coefficient of 0.24 and a long-run mean of 0.06. For the starting year ( $t = 0$ ), we measure ROE as income before extraordinary items (item IB) divided by one-year lagged book equity (item CEQ), and the book equity growth rate as the annual change in sales (item SALE). Nissim and Penman (2001) show that past sales growth is a better indicator of future book equity growth than past book equity growth. Finally, we use a forecasting period of  $T = 10$  years and a cost of equity of  $r = 0.12$ . Firms are excluded if book equity ever becomes negative during the forecasting period. At the end of June of each year  $t$ , we sort stocks into deciles based on Dur constructed with data from the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

## B.3 Investment

### B.3.1 Aci, Abnormal Corporate Investment

At the end of June of year  $t$ , we measure abnormal corporate investment, Aci, as  $\text{Ce}_{t-1} / [(\text{Ce}_{t-2} + \text{Ce}_{t-3} + \text{Ce}_{t-4}) / 3] - 1$ , in which  $\text{Ce}_{t-j}$  is capital expenditure (Compustat annual item CAPX) scaled by sales (item SALE) for the fiscal year ending in calendar year  $t - j$ . The last three-year average capital expenditure is designed to project the benchmark investment in the portfolio formation year. We exclude firms with sales less than ten million dollars. At the end of June of each year  $t$ , we sort stocks into deciles based on Aci. Monthly decile returns are computed from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.2 I/A, Investment-to-assets**

At the end of June of each year  $t$ , we sort stocks into deciles based on investment-to-assets, I/A, which is measured as total assets (Compustat annual item AT) for the fiscal year ending in calendar year  $t - 1$  divided by total assets for the fiscal year ending in  $t - 2$  minus one. Monthly decile returns are computed from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.3 Ia<sup>q6</sup> and Ia<sup>q12</sup>, Quarterly Investment-to-assets**

Quarterly investment-to-assets, Ia<sup>q</sup>, is defined as quarterly total assets (Compustat quarterly item ATQ) divided by four-quarter-lagged total assets minus one. At the beginning of each month  $t$ , we sort stocks into deciles based on Ia<sup>q</sup> for the latest fiscal quarter ending at least four months ago. Monthly decile returns are calculated from month  $t$  to  $t + 5$  (Ia<sup>q6</sup>) and from month  $t$  to  $t + 11$  (Ia<sup>q12</sup>), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in, for instance, Ia<sup>q6</sup>, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Ia<sup>q6</sup> decile.

### **B.3.4 dPia, Changes in PPE and Inventory-to-assets**

Changes in PPE and Inventory-to-assets, dPia, is defined as the annual change in gross property, plant, and equipment (Compustat annual item PPEGT) plus the annual change in inventory (item INVT) scaled by one-year-lagged total assets (item AT). At the end of June of each year  $t$ , we sort stocks into deciles based on dPia for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are computed from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.5 Noa and dNoa, (Changes in) Net Operating Assets**

Following Hirshleifer, Hou, Teoh, and Zhang (2004), 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). Noa is net operating assets scaled by one-year-lagged total assets. Changes in net operating assets, dNoa, is the annual change in net operating assets scaled by one-year-lagged total assets. At the end of June of each year  $t$ , we sort stocks into deciles based on Noa, and separately, on dNOA, for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are computed from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.6 dLno, Changes in Long-term Net Operating Assets**

Following Fairfield, Whisenant, and Yohn (2003), we measure changes in long-term net operating assets as the annual change in net property, plant, and equipment (Compustat item PPENT) plus the change in intangibles (item INTAN) plus the change in other long-term assets (item AO) minus the change in other long-term liabilities (item LO) and plus depreciation and amortization expense (item DP). dLno is the change in long-term net operating assets scaled by the average of total assets (item AT) from the current and prior years. At the end of June of each year  $t$ , we sort stocks

into deciles based on  $dLno$  for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.7 Ig, Investment Growth**

At the end of June of each year  $t$ , we sort stocks into deciles based on investment growth,  $Ig$ , which is the growth rate in capital expenditure (Compustat annual item CAPX) from the fiscal year ending in calendar year  $t - 2$  to the fiscal year ending in  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.8 2Ig, Two-year Investment Growth**

At the end of June of each year  $t$ , we sort stocks into deciles based on two-year investment growth,  $2Ig$ , which is the growth rate in capital expenditure (Compustat annual item CAPX) from the fiscal year ending in calendar year  $t - 3$  to the fiscal year ending in  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.9 Nsi, Net Stock Issues**

At the end of June of year  $t$ , we measure net stock issues,  $Nsi$ , as the natural log of the ratio of the split-adjusted shares outstanding at the fiscal year ending in calendar year  $t - 1$  to the split-adjusted shares outstanding at the fiscal year ending in  $t - 2$ . The split-adjusted shares outstanding is shares outstanding (Compustat annual item CSHO) times the adjustment factor (item AJEX). At the end of June of each year  $t$ , we sort stocks with negative  $Nsi$  into two portfolios (1 and 2), stocks with zero  $Nsi$  into one portfolio (3), and stocks with positive  $Nsi$  into seven portfolios (4 to 10). Monthly decile returns are from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.10 dIi, % Change in Investment - % Change in Industry Investment**

Following Abarbanell and Bushee (1998), we define the  $\%d(\cdot)$  operator as the percentage change in the variable in the parentheses from its average over the prior two years, e.g.,  $\%d(\text{Investment}) = [\text{Investment}(t) - E[\text{Investment}(t)]]/E[\text{Investment}(t)]$ , in which  $E[\text{Investment}(t)] = [\text{Investment}(t-1) + \text{Investment}(t-2)]/2$ .  $dIi$  is defined as  $\%d(\text{Investment}) - \%d(\text{Industry investment})$ , in which investment is capital expenditure in property, plant, and equipment (Compustat annual item CAPXV). Industry investment is the aggregate investment across all firms with the same two-digit SIC code. Firms with non-positive  $E[\text{Investment}(t)]$  are excluded and we require at least two firms in each industry. At the end of June of each year  $t$ , we sort stocks into deciles based on  $dIi$  for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.11 Cei, Composite Equity Issuance**

At the end of June of each year  $t$ , we sort stocks into deciles based on composite equity issuance,  $Cei$ , which is the log growth rate in the market equity not attributable to stock return,  $\log(\text{ME}_t/\text{ME}_{t-5}) - r(t-5, t)$ .  $r(t-5, t)$  is the cumulative log stock return from the last trading day of June in year  $t - 5$  to the last trading day of June in year  $t$ , and  $\text{ME}_t$  is the market equity (from CRSP) on the last trading day of June in year  $t$ . Monthly decile returns are from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.3.12 Ivg, Inventory Growth

At the end of June of each year  $t$ , we sort stocks into deciles based on inventory growth, Ivg, which is the annual growth rate in inventory (Compustat annual item INVT) from the fiscal year ending in calendar year  $t - 2$  to the fiscal year ending in  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.3.13 Ivc, Inventory Changes

At the end of June of each year  $t$ , we sort stocks into deciles based on inventory changes, Ivc, which is the annual change in inventory (Compustat annual item INVT) scaled by the average of total assets (item AT) for the fiscal years ending in  $t - 2$  and  $t - 1$ . We exclude firms that carry no inventory for the past two fiscal years. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.3.14 Oa, Operating Accruals

Prior to 1988, we use the balance sheet approach in Sloan (1996) to measure operating accruals, Oa, 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, Oa 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. Starting from 1988, we follow Hribar and Collins (2002) to measure Oa using the statement of cash flows as net income (item NI) minus net cash flow from operations (item OANCF). Doing so helps mitigate measurement errors that can arise from nonoperating activities such as acquisitions and divestitures. Data from the statement of cash flows are only available since 1988. At the end of June of each year  $t$ , we sort stocks into deciles on Oa for the fiscal year ending in calendar year  $t - 1$  scaled by total assets (item AT) for the fiscal year ending in  $t - 2$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.3.15 dWc and dCoa, Changes in Net Non-cash Working Capital and in Current Operating Assets

Richardson, Sloan, Soliman, and Tuna (2005, Table 10) show that several components of total accruals also forecast returns in the cross section. dWc is the change in net non-cash working capital. Net non-cash working capital is current operating asset (Coa) minus current operating liabilities (Col), with  $Coa = \text{current assets (Compustat annual item ACT)} - \text{cash and short term investments (item CHE)}$  and  $Col = \text{current liabilities (item LCT)} - \text{debt in current liabilities (item DLC)}$ . dCoa is the change in current operating asset. Missing changes in debt in current liabilities are set to zero. At the end of June of each year  $t$ , we sort stocks into deciles based, separately, on dWc and dCoa for the fiscal year ending in calendar year  $t - 1$ , all scaled by total assets (item AT) for the fiscal year ending in calendar year  $t - 2$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.3.16 dNco and dNca, Changes in Net Non-current Operating Assets and in Non-current Operating Assets

dNco is the change in net non-current operating assets. Net non-current operating assets are non-current operating assets (Nca) minus non-current operating liabilities (Ncl), with Nca = total assets (Compustat annual item AT) – current assets (item ACT) – long-term investments (item IVAO), and Ncl = total liabilities (item LT) – current liabilities (item LCT) – long-term debt (item DLTT). dNca is the change in non-current operating assets. Missing changes in long-term investments and long-term debt are set to zero. At the end of June of each year  $t$ , we sort stocks into deciles based, separately, on dNco and dNca for the fiscal year ending in calendar year  $t - 1$ , all scaled by total assets for the fiscal year ending in calendar year  $t - 2$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.3.17 dFin, dFnl, and dBe, Changes in Net Financial Assets, in Financial Liabilities, and in Book Equity

dFin is the change in net financial assets. Net financial assets are financial assets (Fna) minus financial liabilities (Fnl), with Fna = short-term investments (Compustat annual item IVST) + long-term investments (item IVAO), and Fnl = long-term debt (item DLTT) + debt in current liabilities (item DLC) + preferred stock (item PSTK). dFnl is the change in financial liabilities. dBe is the change in book equity (item CEQ). Missing changes in debt in current liabilities, long-term investments, long-term debt, short-term investments, and preferred stocks are set to zero (at least one change has to be non-missing when constructing any variable). At the end of June of each year  $t$ , we sort stocks into deciles based, separately, on dFin, dFnl, and dBe for the fiscal year ending in calendar year  $t - 1$ , all scaled by total assets (item AT) for the fiscal year ending in calendar year  $t - 2$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.3.18 Dac, Discretionary Accruals

We measure discretionary accruals, Dac, using the modified Jones model from Dechow, Sloan, and Sweeney (1995):

$$\frac{Oa_{i,t}}{A_{i,t-1}} = \alpha_1 \frac{1}{A_{i,t-1}} + \alpha_2 \frac{dSALE_{i,t} - dREC_{i,t}}{A_{i,t-1}} + \alpha_3 \frac{PPE_{i,t}}{A_{i,t-1}} + e_{i,t}, \quad (\text{B.7})$$

in which  $Oa_{i,t}$  is operating accruals for firm  $i$  (see Appendix B.3.14),  $A_{t-1}$  is total assets (Compustat annual item AT) at the end of year  $t - 1$ ,  $dSALE_{i,t}$  is the annual change in sales (item SALE) from year  $t - 1$  to  $t$ ,  $dREC_{i,t}$  is the annual change in net receivables (item RECT) from year  $t - 1$  to  $t$ , and  $PPE_{i,t}$  is gross property, plant, and equipment (item PPEGT) at the end of year  $t$ . We estimate the cross-sectional regression (B.7) for each two-digit SIC industry and year combination, formed separately for NYSE/AMEX firms and for NASDAQ firms. We require at least six firms for each regression. The discretionary accrual for stock  $i$  is defined as the residual from the regression,  $e_{i,t}$ . At the end of June of each year  $t$ , we sort stocks into deciles based on Dac for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.19 Poa, Percent Operating Accruals**

Accruals are traditionally scaled by total assets. Hafzalla, Lundholm, and Van Winkle (2011) show that scaling accruals by the absolute value of earnings (percent accruals) is more effective in selecting firms for which the differences between sophisticated and naive forecasts of earnings are the most extreme. To construct the percent operating accruals (Poa) deciles, at the end of June of each year  $t$ , we sort stocks into deciles based on operating accruals scaled by the absolute value of net income (Compustat annual item NI) for the fiscal year ending in calendar year  $t - 1$ . See Appendix B.3.14 for the measurement of operating accruals. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.20 Pta, Percent Total Accruals**

At the end of June of each year  $t$ , we sort stocks into deciles on percent total accruals, Pta, calculated as total accruals scaled by the absolute value of net income (Compustat annual item NI) for the fiscal year ending in calendar year  $t - 1$ . See Appendix ?? for the measurement of total accruals. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of year  $t + 1$ .

### **B.3.21 Pda, Percent Discretionary Accruals**

At the end of June of each year  $t$ , we split stocks into deciles based on percent discretionary accruals, Pda, calculated as the discretionary accruals, Dac, for the fiscal year ending in calendar year  $t - 1$  multiplied with total assets (Compustat annual item AT) for the fiscal year ending in  $t - 2$  scaled by the absolute value of net income (item NI) for the fiscal year ending in  $t - 1$ . See Appendix B.3.18 for the measurement of discretionary accruals. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### **B.3.22 Ndf, Net Debt Financing**

Ndf is net debt financing, Ndf (Bradshaw, Richardson, and Sloan 2006). Ndf is the cash proceeds from the issuance of long-term debt (item DLTIS) less cash payments for long-term debt reductions (item DLTR) plus the net changes in current debt (item DLCCH, zero if missing). At the end of June of each year  $t$ , we sort stocks into deciles based on Ndf for the fiscal year ending in calendar year  $t - 1$  scaled by the average of total assets for fiscal years ending in  $t - 2$  and  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ . Because the data on financing activities start in 1971, the portfolios start in July 1972.

## **B.4 Profitability**

### **B.4.1 Roe1 and Roe6, Return on Equity**

Return on equity, Roe, is income before extraordinary items (Compustat quarterly item IBQ) divided by one-quarter-lagged book equity (Hou, Xue, and Zhang 2015). Book equity is shareholders' equity, plus balance sheet deferred taxes and investment tax credit (item TXDITCQ) if available, minus the book value of preferred stock (item PSTKQ). Depending on availability, we use stockholders' equity (item SEQQ), or common equity (item CEQQ) plus the book value of preferred stock, or total assets (item ATQ) minus total liabilities (item LTQ) in that order as shareholders' equity.

Before 1972, the sample coverage is limited for quarterly book equity in Compustat quarterly files. We expand the coverage by using book equity from Compustat annual files as well as by imputing quarterly book equity with clean surplus accounting. Specifically, whenever available we first use quarterly book equity from Compustat quarterly files. We then supplement the coverage for fiscal quarter four with annual book equity from Compustat annual files. Following Davis, Fama, and French (2000), we measure annual book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ), if available. If not, stockholders' equity is the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock.

If both approaches are unavailable, we apply the clean surplus relation to impute the book equity. First, if available, we backward impute the beginning-of-quarter book equity as the end-of-quarter book equity minus quarterly earnings plus quarterly dividends. Quarterly earnings are income before extraordinary items (Compustat quarterly item IBQ). Quarterly dividends are zero if dividends per share (item DVPSXQ) are zero. Otherwise, total dividends are dividends per share times beginning-of-quarter shares outstanding adjusted for stock splits during the quarter. Shares outstanding are from Compustat (quarterly item CSHOQ supplemented with annual item CSHO for fiscal quarter four) or CRSP (item SHROUT), and the share adjustment factor is from Compustat (quarterly item AJEXQ supplemented with annual item AJEX for fiscal quarter four) or CRSP (item CFACSHR). Because we impose a four-month lag between earnings and the holding period month (and the book equity in the denominator of ROE is one-quarter-lagged relative to earnings), all the Compustat data in the backward imputation are at least four-month lagged prior to the portfolio formation. If data are unavailable for the backward imputation, we impute the book equity for quarter  $t$  forward based on book equity from prior quarters. Let  $BEQ_{t-j}$ ,  $1 \leq j \leq 4$  denote the latest available quarterly book equity as of quarter  $t$ , and  $IBQ_{t-j+1,t}$  and  $DVQ_{t-j+1,t}$  be the sum of quarterly earnings and quarterly dividends from quarter  $t-j+1$  to  $t$ , respectively.  $BEQ_t$  can then be imputed as  $BEQ_{t-j} + IBQ_{t-j+1,t} - DVQ_{t-j+1,t}$ . We do not use prior book equity from more than four quarters ago (i.e.,  $1 \leq j \leq 4$ ) to reduce imputation errors.

At the beginning of each month  $t$ , we sort all stocks into deciles based on their most recent past Roe. Before 1972, we use the most recent Roe computed with quarterly earnings from fiscal quarters ending at least four months prior to the portfolio formation. Starting from 1972, we use Roe computed with quarterly earnings from the most recent quarterly earnings announcement dates (Compustat quarterly item RDQ). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Roe 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. Monthly decile returns are calculated for the current month  $t$  (Roe1) and from month  $t$  to  $t+5$  (Roe6). The deciles are rebalanced monthly. The holding period that is longer than one month as in, for instance, Roe6, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdeciles returns as the monthly return of the Roe6 decile.

#### **B.4.2 dRoe1, dRoe6, and dRoe12, Changes in Return on Equity**

Change in return on equity, dRoe, is return on equity minus its value from four quarters ago. See Appendix B.4.1 for the measurement of return on equity. At the beginning of each month  $t$ , we sort all stocks into deciles on their most recent past dRoe. Before 1972, we use the most recent dRoe with quarterly earnings from fiscal quarters ending at least four months ago. Starting from 1972, we use dRoe computed with quarterly earnings from the most recent quarterly earnings announcement dates (Compustat quarterly item RDQ). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent dRoe 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. Monthly decile returns are calculated for the current month  $t$  (dRoe1), from month  $t$  to  $t + 5$  (dRoe6), and from month  $t$  to  $t + 11$  (dRoe12). The deciles are rebalanced monthly. The holding period that is longer than one month as in, for instance, dRoe6, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdeciles returns as the monthly return of the dRoe6 decile.

#### **B.4.3 Roa1, Return on Assets**

Return on assets, Roa, is income before extraordinary items (Compustat quarterly item IBQ) divided by one-quarter-lagged total assets (item ATQ). At the beginning of each month  $t$ , we sort all stocks into deciles based on Roa computed with quarterly earnings from the most recent earnings announcement dates (item RDQ). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Roa 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. Monthly decile returns are calculated for month  $t$ , and the deciles are rebalanced at the beginning of  $t + 1$ . For sufficient data coverage, the Roa portfolios start in January 1972.

#### **B.4.4 dRoa1 and dRoa6, Changes in Return on Assets**

Change in return on assets, dRoa, is return on assets minus its value from four quarters ago. See Appendix B.4.3 for the measurement of return on assets. At the beginning of each month  $t$ , we sort all stocks into deciles based on dRoa computed with quarterly earnings from the most recent earnings announcement dates (Compustat quarterly item RDQ). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent dRoa 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. Monthly decile returns are calculated for month  $t$  (dRoa1) and from month  $t$  to  $t + 5$  (dRoa6), and the deciles are rebalanced at the beginning of  $t + 1$ . The holding period that is longer than one month as in, for instance, dRoa6, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdecile returns as the monthly return of the dRoa6 decile. For sufficient data coverage, the dRoa portfolios start in January 1973.

#### **B.4.5 Rna<sup>q1</sup>, Rna<sup>q6</sup>, Ato<sup>q1</sup>, Ato<sup>q6</sup>, and Ato<sup>q12</sup>, Quarterly Return on Net Operating Assets, Quarterly Asset Turnover**

Quarterly return on net operating assets,  $Rna^q$ , is quarterly operating income after depreciation (Compustat quarterly item OIADPQ) divided by one-quarter-lagged net operating assets (Noa). Noa is operating assets minus operating liabilities. Operating assets are total assets (item ATQ) minus cash and short-term investments (item CHEQ), and minus other investment and advances (item IVAOQ, zero if missing). Operating liabilities are total assets minus debt in current liabilities (item DLCQ, zero if missing), minus long-term debt (item DLTTQ, zero if missing), minus minority interests (item MIBQ, zero if missing), minus preferred stocks (item PSTKQ, zero if missing), and minus common equity (item CEQQ). Quarterly asset turnover,  $Ato^q$ , is quarterly sales divided by one-quarter-lagged Noa. At the beginning of each month  $t$ , we sort stocks into deciles based on  $Rna^q$  for the latest fiscal quarter ending at least four months ago. Separately, we sort stocks into deciles based on  $Ato^q$  computed with quarterly sales from the most recent earnings announcement dates (item RDQ). Sales are generally announced with earnings during quarterly earnings announcements (Jegadeesh and Livnat 2006). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent  $Ato^q$  to be within six months prior to the portfolio formation. This restriction is imposed to exclude stale information. To avoid potentially erroneous records, we also require the earnings announcement date to be after the corresponding fiscal quarter end. Monthly decile returns are calculated for month  $t$  ( $Rna^{q1}$  and  $Ato^{q1}$ ), from month  $t$  to  $t+5$  ( $Rna^{q6}$  and  $Ato^{q6}$ ), and from month  $t$  to  $t+11$  ( $Ato^{q12}$ ). The deciles are rebalanced at the beginning of  $t+1$ . The holding period that is longer than one month as in, for instance,  $Ato^{q6}$ , means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdecile returns as the monthly return of the  $Ato^{q6}$  decile. For sufficient data coverage, the  $Rna^q$  portfolios start in January 1976 and the  $Ato^q$  portfolios start in January 1972.

#### **B.4.6 Cto<sup>q1</sup>, Cto<sup>q6</sup>, and Cto<sup>q12</sup>, Quarterly Capital Turnover**

Quarterly capital turnover,  $Cto^q$ , is quarterly sales (Compustat quarterly item SALEQ) scaled by one-quarter-lagged total assets (item ATQ). At the beginning of each month  $t$ , we sort stocks into deciles based on  $Cto^q$  computed with quarterly sales from the most recent earnings announcement dates (item RDQ). Sales are generally announced with earnings during quarterly earnings announcements (Jegadeesh and Livnat 2006). For a firm to enter the portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent  $Ato^q$  to be within six months prior to the portfolio formation. This restriction is imposed to exclude stale information. To avoid potentially erroneous records, we also require the earnings announcement date to be after the corresponding fiscal quarter end. Monthly decile returns are calculated for month  $t$  ( $Cto^{q1}$ ), from month  $t$  to  $t+5$  ( $Cto^{q6}$ ), and from month  $t$  to  $t+11$  ( $Cto^{q12}$ ). The deciles are rebalanced at the beginning of  $t+1$ . The holding period that is longer than one month as in, for instance,  $Cto^{q6}$ , means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdecile returns as the monthly return of the  $Cto^{q6}$  decile. For sufficient data coverage, the  $Cto^q$  portfolios start in January 1972.

#### **B.4.7 Gpa, Gross Profits-to-assets**

Following Novy-Marx (2013), we measure gross profits-to-assets,  $Gpa$ , as total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS) divided by total assets (item AT, the

denominator is current, not lagged, total assets). At the end of June of each year  $t$ , we sort stocks into deciles based on  $Gpa$  for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

#### **B.4.8 $Gla^q1$ , $Gla^q6$ , and $Gla^q12$ , Quarterly Gross Profits-to-lagged Assets**

$Gla^q$ , is quarterly total revenue (Compustat quarterly item  $REVTQ$ ) minus cost of goods sold (item  $COGSQ$ ) divided by one-quarter-lagged total assets (item  $ATQ$ ). At the beginning of each month  $t$ , we sort stocks into deciles based on  $Gla^q$  for the fiscal quarter ending at least four months ago. Monthly decile returns are calculated for month  $t$  ( $Gla^q1$ ), from month  $t$  to  $t + 5$  ( $Gla^q6$ ), and from month  $t$  to  $t + 11$  ( $Gla^q12$ ). The deciles are rebalanced at the beginning of  $t + 1$ . The holding period that is longer than one month as in, for instance,  $Gla^q6$ , means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdecile returns as the monthly return of the  $Gla^q6$  decile. For sufficient data coverage, the  $Gla^q$  portfolios start in January 1976.

#### **B.4.9 $Ole^q1$ and $Ole^q6$ , Quarterly Operating Profits-to-lagged Equity**

Quarterly operating profits-to-lagged equity,  $Ole^q$ , is quarterly total revenue (Compustat quarterly item  $REVTQ$ ) minus cost of goods sold (item  $COGSQ$ , zero if missing), minus selling, general, and administrative expenses (item  $XSGAQ$ , zero if missing), and minus interest expense (item  $XINTQ$ , zero if missing), scaled by one-quarter-lagged book equity. We require at least one of the three expense items ( $COGSQ$ ,  $XSGAQ$ , and  $XINTQ$ ) to be non-missing. Book equity is shareholders' equity, plus balance sheet deferred taxes and investment tax credit (item  $TXDITCQ$ ) if available, minus the book value of preferred stock (item  $PSTKQ$ ). Depending on availability, we use stockholders' equity (item  $SEQQ$ ), or common equity (item  $CEQQ$ ) plus the book value of preferred stock, or total assets (item  $ATQ$ ) minus total liabilities (item  $LTQ$ ) in that order as shareholders' equity.

At the beginning of each month  $t$ , we split stocks on  $Ole^q$  for the fiscal quarter ending at least four months ago. Monthly decile returns are calculated for month  $t$  ( $Ole^q1$ ) and from month  $t$  to  $t + 5$  ( $Ole^q6$ ), and the deciles are rebalanced at the beginning of  $t + 1$ . The holding period longer than one month as in  $Ole^q6$  means that for a given decile in each month there exist six subdeciles, each initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the  $Ole^q6$  decile. For sufficient data coverage, the  $Ole^q$  portfolios start in January 1972.

#### **B.4.10 $Opa$ , Operating Profits-to-assets**

Following Ball, Gerakos, Linnainmaa, and Nikolaev (2015), we measure operating profits-to-assets,  $Opa$ , as total revenue (Compustat annual item  $REVT$ ) minus cost of goods sold (item  $COGS$ ), minus selling, general, and administrative expenses (item  $XSGA$ ), and plus research and development expenditures (item  $XRD$ , zero if missing), scaled by book assets (item  $AT$ , the denominator is current, not lagged, total assets). At the end of June of each year  $t$ , we sort stocks into deciles based on  $Opa$  for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

#### **B.4.11 Ola<sup>q1</sup>, Ola<sup>q6</sup>, and Ola<sup>q12</sup>, Quarterly Operating Profits-to-lagged Assets**

Quarterly operating profits-to-lagged assets,  $Ola^q$ , is quarterly total revenue (Compustat quarterly item REVTQ) minus cost of goods sold (item COGSQ), minus selling, general, and administrative expenses (item XSGAQ), plus research and development expenditures (item XRDQ, zero if missing), scaled by one-quarter-lagged book assets (item ATQ). At the beginning of each month  $t$ , we sort stocks into deciles based on  $Ola^q$  for the fiscal quarter ending at least four months ago. Monthly decile returns are calculated for month  $t$  ( $Ola^{q1}$ ), from month  $t$  to  $t + 5$  ( $Ola^{q6}$ ), and from month  $t$  to  $t + 11$  ( $Ola^{q12}$ ). The deciles are rebalanced at the beginning of  $t + 1$ . The holding period longer than one month as in  $Ola^{q6}$  means that for a given decile in each month there exist six subdeciles, each initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the  $Ola^{q6}$  decile. For sufficient data coverage, the  $Ola^q$  portfolios start in January 1976.

#### **B.4.12 Cop, Cash-based Operating Profitability**

Following Ball, Gerakos, Linnainmaa, and Nikolaev (2016), we measure cash-based operating profitability,  $Cop$ , as total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS), minus selling, general, and administrative expenses (item XSGA), plus research and development expenditures (item XRD, zero if missing), minus change in accounts receivable (item RECT), minus change in inventory (item INVT), minus change in prepaid expenses (item XPP), plus change in deferred revenue (item DRC plus item DRLT), plus change in trade accounts payable (item AP), and plus change in accrued expenses (item XACC), all scaled by book assets (item AT, the denominator is current, not lagged, total assets). All changes are annual changes in balance sheet items and we set missing changes to zero. At the end of June of each year  $t$ , we sort stocks into deciles based on  $Cop$  for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

#### **B.4.13 Cla, Cash-based Operating Profits-to-lagged Assets**

Cash-based operating profits-to-lagged assets,  $Cla$ , is total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS), minus selling, general, and administrative expenses (item XSGA), plus research and development expenditures (item XRD, zero if missing), minus change in accounts receivable (item RECT), minus change in inventory (item INVT), minus change in prepaid expenses (item XPP), plus change in deferred revenue (item DRC plus item DRLT), plus change in trade accounts payable (item AP), and plus change in accrued expenses (item XACC), all scaled by one-year-lagged book assets (item AT). All changes are annual changes in balance sheet items and we set missing changes to zero. At the end of June of each year  $t$ , we sort stocks into deciles based on  $Cla$  for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

#### **B.4.14 Cla<sup>q1</sup>, Cla<sup>q6</sup>, and Cla<sup>q12</sup>, Quarterly Cash-based Operating Profits-to-lagged Assets**

Quarterly cash-based operating profits-to-lagged assets,  $Cla$ , is quarterly total revenue (Compustat quarterly item REVTQ) minus cost of goods sold (item COGSQ), minus selling, general, and administrative expenses (item XSGAQ), plus research and development expenditures (item XRDQ, zero if missing), minus change in accounts receivable (item RECTQ), minus change in inventory

(item INVTQ), plus change in deferred revenue (item DRCQ plus item DRLTQ), and plus change in trade accounts payable (item APQ), all scaled by one-quarter-lagged book assets (item ATQ). All changes are quarterly changes in balance sheet items and we set missing changes to zero. At the beginning of each month  $t$ , we split stocks on Cla<sup>q</sup> for the fiscal quarter ending at least four months ago. Monthly decile returns are calculated for month  $t$  (Cla<sup>q1</sup>), from month  $t$  to  $t + 5$  (Cla<sup>q6</sup>), and from month  $t$  to  $t + 11$  (Cla<sup>q12</sup>). The deciles are rebalanced at the beginning of  $t + 1$ . The holding period longer than one month as in Cla<sup>q6</sup> means that for a given decile in each month there exist six subdeciles, each initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Cla<sup>q6</sup> decile. For sufficient data coverage, the Cla<sup>q</sup> portfolios start in January 1976.

#### B.4.15 F<sup>q1</sup>, F<sup>q6</sup>, and F<sup>q12</sup>, Quarterly Fundamental Score

To construct quarterly F-score, F<sup>q</sup>, we use quarterly accounting data and the same nine binary signals from Piotroski (2000). Among the four signals related to profitability: (i) Roa is quarterly income before extraordinary items (Compustat quarterly item IBQ) scaled by one-quarter-lagged total assets (item ATQ). If the firm's Roa is positive, the indicator variable F<sub>Roa</sub> equals one and zero otherwise. (ii) Cf/A is quarterly cash flow from operation scaled by one-quarter-lagged total assets. Cash flow from operation is the quarterly change in year-to-date net cash flow from operating activities (item OANCFY) if available, or the quarterly change in year-to-date funds from operation (item FOPTY) minus the quarterly change in working capital (item WCAPQ). If the firm's Cf/A is positive, the indicator variable F<sub>Cf/A</sub> equals one and zero otherwise. (iii) dRoa is the current quarter's Roa less the Roa from four quarters ago. If dRoa is positive, the indicator variable F<sub>dROA</sub> is one and zero otherwise. Finally, (iv) the indicator F<sub>Acc</sub> equals one if Cf/A > Roa and zero otherwise.

Among the three signals related changes in capital structure and a firm's ability to meet future debt obligations: (i) dLever is the change in the ratio of total long-term debt (Compustat quarterly item DLTTQ) to the average of current and one-quarter-lagged total assets. F<sub>dLever</sub> is one if the firm's leverage ratio falls, i.e., dLever < 0, relative to its value four quarters ago, and zero otherwise. (ii) dLiquid measures the change in a firm's current ratio between the current quarter and four quarters ago, in which the current ratio is the ratio of current assets (item ACTQ) to current liabilities (item LCTQ). An improvement in liquidity (dLiquid > 0) is a good signal about the firm's ability to service current debt obligations. The indicator F<sub>dLiquid</sub> equals one if the firm's liquidity improves and zero otherwise. (iii) The indicator, Eq, equals one if the firm does not issue common equity during the past four quarters and zero otherwise. The issuance of common equity is sales of common and preferred stocks minus any increase in preferred stocks (item PSTKQ). To measure sales of common and preferred stocks, we first compute the quarterly change in year-to-date sales of common and preferred stocks (item SSTKY) and then take the total change for the past four quarters. Issuing equity is interpreted as a bad signal (inability to generate sufficient internal funds to service future obligations). For the remaining two signals, (i) dMargin is the firm's current gross margin ratio, measured as gross margin (item SALEQ minus item COGSQ) scaled by sales (item SALEQ), less the gross margin ratio from four quarters ago. The indicator F<sub>dMargin</sub> equals one if dMargin > 0 and zero otherwise. (ii) dTurn is the firm's current asset turnover ratio, measured as (item SALEQ) scaled by one-quarter-lagged total assets (item ATQ), minus the asset turnover ratio from four quarters ago. The indicator, F<sub>dTurn</sub>, equals one if dTurn > 0 and zero otherwise.

The composite score, F<sup>q</sup>, is the sum of the individual binary signals:

$$F^q \equiv F_{Roa} + F_{dRoa} + F_{Cf/A} + F_{Acc} + F_{dMargin} + F_{dTurn} + F_{dLever} + F_{dLiquid} + Eq. \quad (B.8)$$

At the beginning of each month  $t$ , we sort stocks based on Fq for the fiscal quarter ending at least four quarters ago to form seven portfolios: low ( $F^q = 0,1,2$ ), 3, 4, 5, 6, 7, and high ( $F^q = 8, 9$ ). Monthly portfolio returns are calculated for month  $t$  ( $F^q1$ ), from month  $t$  to  $t + 5$  ( $F^q6$ ), and from month  $t$  to  $t + 11$  ( $F^q12$ ), and the portfolios are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in, for instance,  $F^q6$ , means that for a given portfolio in each month there exist six subportfolios, each of which is initiated in a different month in prior six months. We take the simple average of the subportfolio returns as the monthly return of the  $F^q6$  portfolio. For sufficient data coverage, the  $F^q$  portfolios start in January 1985.

#### B.4.16 Fp<sup>q6</sup>, Failure Probability

Failure probability (Fp) is from Campbell, Hilscher, and Szilagyi (2008, Table IV, Column 3):

$$\begin{aligned} Fp_t \equiv & -9.164 - 20.264NIMTAAVG_t + 1.416TLMTA_t - 7.129EXRETAVG_t \\ & + 1.411SIGMA_t - 0.045RSIZE_t - 2.132CASHMTA_t + 0.075MB_t - 0.058PRICE_t \end{aligned} \quad (B.9)$$

in which

$$NIMTAAVG_{t-1,t-12} \equiv \frac{1 - \phi^3}{1 - \phi^{12}} (NIMTA_{t-1,t-3} + \dots + \phi^9 NIMTA_{t-10,t-12}) \quad (B.10)$$

$$EXRETAVG_{t-1,t-12} \equiv \frac{1 - \phi}{1 - \phi^{12}} (EXRET_{t-1} + \dots + \phi^{11} EXRET_{t-12}), \quad (B.11)$$

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  $\equiv \log(1 + R_{it}) - \log(1 + R_{S\&P500,t})$  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. We winsorize the variables on the right-hand side of equation (C.1) at the 1th and 99th percentiles of their distributions each month.

At the beginning of each month  $t$ , we split stocks into deciles based on Fp calculated with ac-

counting data from the fiscal quarter ending at least four months ago. We calculate decile returns from month  $t$  to  $t + 5$  (Fp<sup>q6</sup>), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period that is longer than one month means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six-month period. We take the simple average of the subdeciles returns as the monthly return of the Fp<sup>q6</sup> decile. For sufficient data coverage, the quarterly Fp deciles start in January 1976.

## B.5 Intangibles

### B.5.1 Oca and Ioca, (Industry-adjusted) Organizational Capital-to-assets

Following Eisfeldt and Papanikolaou (2013), we construct the stock of organization capital, Oc, using the perpetual inventory method:

$$Oc_{it} = (1 - \delta)Oc_{it-1} + SG\&A_{it}/CPI_t, \quad (\text{B.12})$$

in which  $Oc_{it}$  is the organization capital of firm  $i$  at the end of year  $t$ ,  $SG\&A_{it}$  is selling, general, and administrative (SG&A) expenses (Compustat annual item XSGA) in  $t$ ,  $CPI_t$  is the average consumer price index during year  $t$ , and  $\delta$  is the annual depreciation rate of Oc. The initial stock of Oc is  $Oc_{i0} = SG\&A_{i0}/(g + \delta)$ , in which  $SG\&A_{i0}$  is the first valid SG&A observation (zero or positive) for firm  $i$  and  $g$  is the long-term growth rate of SG&A. We assume a depreciation rate of 15% for Oc and a long-term growth rate of 10% for SG&A. Missing SG&A values after the starting date are treated as zero. For portfolio formation at the end of June of year  $t$ , we require SG&A to be non-missing for the fiscal year ending in calendar year  $t - 1$  because this SG&A value receives the highest weight in Oc. In addition, we exclude firms with zero Oc. Organizational Capital-to-assets, Oca, is Oc scaled by total assets (item AT). We also industry-standardize Oca using the FF (1997) 17-industry classification. To calculate the industry-adjusted Oca, Ioca, we demean a firm's Oca by its industry mean and then divide the demeaned Oca by the standard deviation of Oca within its industry. To alleviate the impact of outliers, we winsorize Oca at the 1 and 99 percentiles of all firms each year before the industry standardization. At the end of June of each year  $t$ , we sort stocks into deciles based on Oca, and separately, on Ioca, for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.5.2 Adm, Advertising Expense-to-market

At the end of June of each year  $t$ , we sort stocks into deciles based on advertising expenses-to-market, Adm, which is advertising expenses (Compustat annual item XAD) for the fiscal year ending in calendar year  $t - 1$  divided by the market equity (from CRSP) at the end of December of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Adm. We keep only firms with positive advertising expenses. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ . Because sufficient XAD data start in 1972, the Adm portfolios start in July 1973.

### B.5.3 Rdm, R&D Expense-to-market

At the end of June of each year  $t$ , we sort stocks into deciles based on R&D-to-market, Rdm, which is R&D expenses (Compustat annual item XRD) for the fiscal year ending in calendar year  $t - 1$  divided by the market equity (from CRSP) at the end of December of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Rdm. We

keep only firms with positive R&D expenses. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ . Because the accounting treatment of R&D expenses was standardized in 1975, the Rdm portfolios start in July 1976.

#### **B.5.4 Rdm<sup>q1</sup>, Rdm<sup>q6</sup>, and Rdm<sup>q12</sup>, Quarterly R&D Expense-to-market**

At the beginning of each month  $t$ , we split stocks into deciles based on quarterly R&D-to-market, Rdm<sup>q</sup>, which is quarterly R&D expense (Compustat quarterly item XRDQ) for the fiscal quarter ending at least four months ago scaled by the market equity (from CRSP) at the end of  $t - 1$ . For firms with more than one share class, we merge the market equity for all share classes before computing Rdm<sup>q</sup>. We keep only firms with positive R&D expenses. We calculate decile returns for the current month  $t$  (Rdm<sup>q1</sup>), from month  $t$  to  $t + 5$  (Rdm<sup>q6</sup>), and from month  $t$  to  $t + 11$  (Rdm<sup>q12</sup>), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in, for instance, Rdm<sup>q6</sup>, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Rdm<sup>q6</sup> decile. Because the quarterly R&D data start in late 1989, the Rdm<sup>q</sup> portfolios start in January 1990.

#### **B.5.5 Ol, Operating Leverage**

Following Novy-Marx (2011), operating leverage, Ol, is operating costs scaled by total assets (Compustat annual item AT, the denominator is current, not lagged, total assets). Operating costs are cost of goods sold (item COGS) plus selling, general, and administrative expenses (item XSGA). At the end of June of year  $t$ , we sort stocks into deciles based on Ol for the fiscal year ending in calendar year  $t - 1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

#### **B.5.6 Ol<sup>q1</sup>, Ol<sup>q6</sup>, and Ol<sup>q12</sup>, Quarterly Operating Leverage**

At the beginning of each month  $t$ , we split stocks into deciles based on quarterly operating leverage, Ol<sup>q</sup>, which is quarterly operating costs divided by assets (Compustat quarterly item ATQ) for the fiscal quarter ending at least four months ago. Operating costs are the cost of goods sold (item COGSQ) plus selling, general, and administrative expenses (item XSGAQ). We calculate decile returns for the current month  $t$  (Ol<sup>q1</sup>), from month  $t$  to  $t + 5$  (Ol<sup>q6</sup>), and from month  $t$  to  $t + 11$  (Ol<sup>q12</sup>), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in, for instance, Ol<sup>q6</sup>, means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Ol<sup>q6</sup> decile. For sufficient data coverage, the Ol<sup>q</sup> portfolios start in January 1972.

#### **B.5.7 Hs, Industry Concentration in Sales**

Following Hou and Robinson (2006), we measure a firm's industry concentration with the Herfindahl index,  $\sum_{i=1}^{N_j} s_{ij}^2$ , in which  $s_{ij}$  is the market share of firm  $i$  in industry  $j$ , and  $N_j$  is the total number of firms in the industry. We calculate the market share of a firm using sales (Compustat annual item SALE). Industries are defined by three-digit SIC codes. We exclude financial firms (SIC between 6000 and 6999) and firms in regulated industries. Following Barclay and Smith (1995), the regulated

industries include: railroads (SIC=4011) through 1980, trucking (4210 and 4213) through 1980, airlines (4512) through 1978, telecommunication (4812 and 4813) through 1982, and gas and electric utilities (4900 to 4939). To improve the accuracy of the concentration measure, we exclude an industry if the market share data are available for fewer than five firms or 80% of all firms in the industry. We measure industry concentration as the average Herfindahl index during the past three years. Industry concentration calculated with sales is denoted  $Hs$ . At the end of June of each year  $t$ , we sort stocks into deciles based on  $Hs$  for the fiscal year ending in calendar year  $t-1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t+1$ , and the deciles are rebalanced in June of  $t+1$ .

### B.5.8 Etr, Effective Tax Rate

Following Abarbanell and Bushee (1998), we measure effective tax rate,  $Etr$ , as:

$$Etr(t) = \left[ \frac{\text{TaxExpense}(t)}{\text{EBT}(t)} - \frac{1}{3} \sum_{\tau=1}^3 \frac{\text{TaxExpense}(t-\tau)}{\text{EBT}(t-\tau)} \right] \times dEPS(t), \quad (\text{B.13})$$

in which  $\text{TaxExpense}(t)$  is total income taxes (Compustat annual item TXT) paid in year  $t$ ,  $\text{EBT}(t)$  is pretax income (item PI) plus amortization of intangibles (item AM), and  $dEPS$  is the change in split-adjusted earnings per share (item EPSPX divided by item AJEX) between years  $t-1$  and  $t$ , deflated by stock price (item PRCLF) at the end of  $t-1$ . At the end of June of each year  $t$ , we sort stocks into deciles based on  $Etr$  for the fiscal year ending in calendar year  $t-1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t+1$ , and the deciles are rebalanced in June of  $t+1$ .

### B.5.9 Rer, Industry-adjusted Real Estate Ratio

Following Tuzel (2010), we measure the real estate ratio as the sum of buildings (Compustat annual item PPENB) and capital leases (item PPENLS) divided by net property, plant, and equipment (item PPENT) prior to 1983. From 1984 onward, the real estate ratio is the sum of buildings at cost (item FATB) and leases at cost (item FATL) divided by gross property, plant, and equipment (item PPEGT). Industry-adjusted real estate ratio,  $Rer$ , is the real estate ratio minus its industry average. Industries are defined by two-digit SIC codes. To alleviate the impact of outliers, we winsorize the real estate ratio at the 1st and 99th percentiles of its distribution each year before computing  $Rer$ . Following Tuzel (2010), we exclude industries with fewer than five firms. At the end of June of each year  $t$ , we sort stocks into deciles based on  $Rer$  for the fiscal year ending in calendar year  $t-1$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t+1$ , and the deciles are rebalanced in June of  $t+1$ . Because the real estate data start in 1969, the  $Rer$  portfolios start in July 1970.

### B.5.10 Eprd, Earnings Predictability

Following Francis, Lafond, Olsson, and Schipper (2004), we estimate earnings predictability,  $Eprd$ , from a first-order autoregressive model for annual split-adjusted earnings per share (Compustat annual item EPSPX divided by item AJEX). At the end of June of each year  $t$ , we estimate the autoregressive model in the ten-year rolling window up to the fiscal year ending in calendar year  $t-1$ . Only firms with a complete ten-year history are included.  $Eprd$  is measured as the residual volatility. We sort stocks into deciles based on  $Eprd$ . Monthly decile returns are calculated from July of year  $t$  to June of  $t+1$ , and the deciles are rebalanced in June of  $t+1$ .

### B.5.11 Etl, Earnings Timeliness

Following Francis, Lafond, Olsson, and Schipper (2004), we measure earnings timeliness, Etl, from the following rolling-window regression:

$$\text{EARN}_{it} = \alpha_{i0} + \alpha_{i1} \text{NEG}_{it} + \beta_{i1} R_{it} + \beta_{i2} \text{NEG}_{it} R_{it} + e_{it}, \quad (\text{B.14})$$

in which  $\text{EARN}_{it}$  is earnings (Compustat annual item IB) for the fiscal year ending in calendar year  $t$ , scaled by the fiscal year-end market equity.  $R_{it}$  is firm  $i$ 's 15-month stock return ending three months after the end of fiscal year ending in calendar year  $t$ .  $\text{NEG}_{it}$  equals one if  $R_{it} < 0$ , and zero otherwise. For firms with more than one share class, we merge the market equity for all share classes. We measure Etl as the  $R^2$  from the regression in (B.14). At the end of June of each year  $t$ , we sort stocks into deciles based on Etl calculated over the ten-year rolling window up to the fiscal year ending in calendar year  $t - 1$ . Only firms with a complete ten-year history are included. Monthly decile returns are calculated from July of year  $t$  to June of  $t + 1$ , and the deciles are rebalanced in June of  $t + 1$ .

### B.5.12 Alm<sup>q</sup>1, Alm<sup>q</sup>6, and Alm<sup>q</sup>12, Quarterly Asset Liquidity

We measure quarterly asset liquidity as cash + 0.75 × noncash current assets + 0.50 × tangible fixed assets. Cash is cash and short-term investments (Compustat quarterly item CHEQ). Noncash current assets is current assets (item ACTQ) minus cash. Tangible fixed assets is total assets (item ATQ) minus current assets (item ACTQ), minus goodwill (item GDWLQ, zero if missing), and minus intangibles (item INTANQ, zero if missing). Alm<sup>q</sup> is quarterly asset liquidity scaled by one-quarter-lagged market value of assets. Market value of assets is total assets plus market equity (item PRCCQ times item CSHOQ) minus book equity (item CEQQ). At the beginning of each month  $t$ , we sort stocks into deciles based on Alm<sup>q</sup> for the fiscal quarter ending at least four months ago. Monthly decile returns are calculated for the current month  $t$  (Alm<sup>q</sup>1), from month  $t$  to  $t + 5$  (Alm<sup>q</sup>6), and from month  $t$  to  $t + 11$  (Alm<sup>q</sup>12). The deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month as in Alm<sup>q</sup>6 means that for a given decile in each month there exist six subdeciles, each of which is initiated in a different month in the prior six months. We take the simple average of the subdecile returns as the monthly return of the Alm<sup>q</sup>6 decile. For sufficient data coverage, the quarterly asset liquidity portfolios start in January 1976.

### B.5.13 $R_a^1$ , $R_a^{[2,5]}$ , $R_n^{[2,5]}$ , $R_a^{[6,10]}$ , $R_n^{[6,10]}$ , $R_a^{[11,15]}$ , and $R_a^{[16,20]}$ , Seasonality

Following Heston and Sadka (2008), at the beginning of each month  $t$ , we sort stocks into deciles based on various measures of past performance, including returns in month  $t - 12$  ( $R_a^1$ ), average returns across months  $t - 24, t - 36, t - 48$ , and  $t - 60$  ( $R_a^{[2,5]}$ ), average returns from month  $t - 60$  to  $t - 13$  except for lags 24, 36, 48, and 60 ( $R_n^{[2,5]}$ ), average returns across months  $t - 72, t - 84, t - 96, t - 108$ , and  $t - 120$  ( $R_a^{[6,10]}$ ), average returns from month  $t - 120$  to  $t - 61$  except for lags 72, 84, 96, 108, and 120 ( $R_n^{[6,10]}$ ), average returns across months  $t - 132, t - 144, t - 156, t - 168$ , and  $t - 180$  ( $R_a^{[11,15]}$ ), and average returns across months  $t - 192, t - 204, t - 216, t - 228$ , and  $t - 240$  ( $R_a^{[16,20]}$ ). Monthly decile returns are calculated for the current month  $t$ , and the deciles are rebalanced at the beginning of month  $t + 1$ .

## B.6 Trading frictions

### B.6.1 Sv1, Systematic Volatility Risk

Following Ang, Hodrick, Xing, and Zhang (2006), we measure systematic volatility risk, Sv, as  $\beta_{dVXO}^i$  from the bivariate regression:

$$r_d^i = \beta_0^i + \beta_{MKT}^i MKT_d + \beta_{dVXO}^i dVXO_d + \epsilon_d^i, \quad (\text{B.15})$$

in which  $r_d^i$  is stock  $i$ 's excess return on day  $d$ ,  $MKT_d$  is the market factor return, and  $dVXO_d$  is the aggregate volatility shock measured as the daily change in the Chicago Board Options Exchange S&P 100 volatility index (VXO). At the beginning of each month  $t$ , we sort stocks into deciles based on  $\beta_{dVXO}^i$  estimated with the daily returns from month  $t - 1$ . We require a minimum of 15 daily returns. Monthly decile returns are calculated for the current month  $t$ , and the deciles are rebalanced at the beginning of month  $t + 1$ . Because the VXO data start in January 1986, the Sv portfolios start in February 1986.

### B.6.2 Dtv12, Dollar Trading Volume

At the beginning of each month  $t$ , we sort stocks into deciles based on their average daily dollar trading volume, Dtv, over the prior six months from  $t - 6$  to  $t - 1$ . We require a minimum of 50 daily observations. Dollar trading volume is share price times the number of shares traded. We adjust the trading volume of NASDAQ stocks per Gao and Ritter (2010).<sup>2</sup> Monthly decile returns are calculated from month  $t$  to  $t + 11$  (Dtv12), and the deciles are rebalanced at the beginning of month  $t + 1$ . The holding period longer than one month for Dtv12, means that for a given decile in each month there exist 12 subdeciles, each of which is initiated in a different month in the prior 12 months. We take the simple average of the subdecile returns as the monthly return of the Dtv12 decile.

### B.6.3 Isff1, Idiosyncratic Skewness per the Fama-French 3-factor Model

At the beginning of each month  $t$ , we sort stocks into deciles based on idiosyncratic skewness, Isff, calculated as the skewness of the residuals from regressing a stock's excess return on the Fama-French three factors using daily observations from month  $t - 1$ . We require a minimum of 15 daily returns. Monthly decile returns are calculated for the current month  $t$ , and the deciles are rebalanced at the beginning of month  $t + 1$ .

### B.6.4 Isq1, Idiosyncratic Skewness per the $q$ -factor Model

At the beginning of each month  $t$ , we sort stocks into deciles based on idiosyncratic skewness, Isq, calculated as the skewness of the residuals from regressing a stock's excess return on the  $q$ -factors

---

<sup>2</sup> We adjust the NASDAQ trading volume to account for the institutional differences between NASDAQ and NYSE-Amex volumes (Gao and Ritter 2010). Prior to February 1, 2001, we divide NASDAQ volume by two. This procedure adjusts for the practice of counting as trades both trades with market makers and trades among market makers. On February 1, 2001, according to the director of research of NASDAQ and Frank Hathaway (the chief economist of NASDAQ), a "riskless principal" rule goes into effect and results in a reduction of approximately 10% in reported volume. From February 1, 2001 to December 31, 2001, we thus divide NASDAQ volume by 1.8. During 2002, securities firms began to charge institutional investors commissions on NASDAQ trades, rather than the prior practice of marking up or down the net price. This practice results in a further reduction in reported volume of approximately 10%. For 2002 and 2003, we divide NASDAQ volume by 1.6. For 2004 and later years, in which the volume of NASDAQ (and NYSE) stocks has mostly been occurring on crossing networks and other venues, we use a divisor of 1.0.

using daily observations from month  $t - 1$ . We require a minimum of 15 daily returns. Monthly decile returns are calculated for the current month, and the deciles are rebalanced at the beginning of month  $t+1$ . Because the q-factors start in January 1967, the Ivq portfolios start in February 1967.

## C Replicating the Stambaugh-Yuan (2017) Factors

To make the document self-contained, we furnish the details of replicating the Stambaugh-Yuan factors in Hou, Mo, Xue, and Zhang (2018).

### C.1 Factor Construction

We describe below the 11 anomaly variables used to construct the Stambaugh-Yuan factors (Appendix C.2). 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, investment-to-assets, and changes in property, plant, and equipment plus change in inventory scaled by 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 component anomalies.

We replicate the Stambaugh-Yuan factors from two separate, independent  $2 \times 3$  sorts, with one on size and MGMT, and another on size and PERF. At the beginning of each month  $t$ , we sort stocks by the NYSE median size into two groups, small and big. Independently, we split stocks based on MGMT, and separately, on PERF, into three groups, low, median, and high, with the 30th and 70th percentiles of the NYSE breakpoints. Taking intersections yields six size-MGMT and six size-PERF portfolios. Monthly value-weighted portfolio returns are calculated for the current month  $t$ , and the portfolios are rebalanced at the beginning of month  $t + 1$ . The MGMT factor is the average of the returns on the two low MGMT portfolios minus the average of the returns on the two high MGMT portfolios. The PERF factor is the average of the returns on the two low PERF portfolios minus the average of the returns on the two high PERF portfolios. Finally, each of the two independent sorts yields a size factor, which is the average of the returns on the three small portfolios minus the average of the returns on the three big portfolios. We take the average of the two size factors as the size factor in the replicated Stambaugh-Yuan model.

### C.2 Variable Definitions

Net stock issues is the annual change in the log of the 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. 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  $t - 1$ ).

Following Sloan (1996), 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.

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.

We measure investment-to-assets as the annual change in total assets (Compustat annual item AT) scaled by one-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 measured as the annual change in gross property, plant, and equipment (Compustat annual item PPEGT) plus the annual change in inventory (item INVT) scaled by one-year-lagged total 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.

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

$$\begin{aligned} \text{Fp}_t \equiv & -9.164 - 20.264\text{NIMTAAVG}_t + 1.416\text{TLMTA}_t - 7.129\text{EXRETAVG}_t \\ & + 1.411\text{SIGMA}_t - 0.045\text{RSIZE}_t - 2.132\text{CASHMTA}_t + 0.075\text{MB}_t - 0.058\text{PRICE}_t \end{aligned} \quad (\text{C.1})$$

in which

$$\text{NIMTAAVG}_{t-1,t-12} \equiv \frac{1 - \phi^3}{1 - \phi^{12}} (\text{NIMTA}_{t-1,t-3} + \dots + \phi^9 \text{NIMTA}_{t-10,t-12}) \quad (\text{C.2})$$

$$\text{EXRETAVG}_{t-1,t-12} \equiv \frac{1 - \phi}{1 - \phi^{12}} (\text{EXRET}_{t-1} + \dots + \phi^{11} \text{EXRET}_{t-12}), \quad (\text{C.3})$$

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 Szilagyι 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 (C.1) at the 1th and 99th percentiles of their distributions each month.

We follow Ohlson (1980, Model One in Table 4) to construct O-score:

$$O \equiv -1.32 - 0.407 \log(\text{TA}) + 6.03\text{TLTA} - 1.43\text{WCTA} + 0.076\text{CLCA} \\ - 1.72\text{OENEG} - 2.37\text{NITA} - 1.83\text{FUTL} + 0.285\text{INTWO} - 0.521\text{CHIN}, \quad (\text{C.4})$$

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  $\text{NI}_s$  and  $\text{NI}_{s-1}$  are the net income for the current and prior years. We winsorize all non-dummy variables on the right-hand side of equation (C.4) 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.

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, the denominator is current, not lagged, total assets). 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 one-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.

## D Replicating the Daniel-Hirshleifer-Sun (2018) Factors

We replicate the Daniel-Hirshleifer-Sun factors as in Hou, Mo, Xue, and Zhang (2018). We replicate the post-earnings-announcement-draft factor (PEAD) by combining standardized unexpected earnings (Sue), the 4-day cumulative abnormal return around the most recent quarterly earnings announcement dates (Abr), and revisions in analysts' earnings forecasts (Re).

Sue is the change in split-adjusted quarterly earnings per share (Compustat quarterly item EP-SPXQ divided by item AJEXQ) from its value four quarters ago divided by the standard deviation of this change in quarterly earnings over the prior eight quarters (six quarters minimum). Before 1972, we use the most recent Sue with earnings from fiscal quarters ending at least four months prior to the portfolio formation. Starting from 1972, we use Sue with quarterly earnings from the most recent quarterly earnings announcement dates (Compustat quarterly item RDQ). For a firm to enter our portfolio formation, we require the end of the fiscal quarter that corresponds to its most recent Sue to be within six months prior to the portfolio formation. Abr is measured as a stock's daily return minus the value-weighted market's daily return cumulated from two days prior to and one day after the most recent quarterly earnings announcement dates. To measure Re, because analysts' earnings forecasts from the Institutional Brokers' Estimate System (IBES) are not necessarily revised each month, we construct a 6-month moving average of past revisions,  $\sum_{\tau=1}^6 (f_{it-\tau} - f_{it-\tau-1}) / p_{it-\tau-1}$ , in which  $f_{it-\tau}$  is the consensus mean forecast (IBES unadjusted file, item MEANEST) issued in month  $t - \tau$  for firm  $i$ 's current fiscal year earnings (fiscal period indicator = 1), and  $p_{it-\tau-1}$  is the prior month's share price (unadjusted file, item PRICE). We require both earnings forecasts and share prices to be denominated in US dollars (currency code = USD). We also adjust for any stock splits and require a minimum of four monthly forecast changes when constructing Re.

At the beginning of each month  $t$ , we calculate a stock's NYSE percentiles on each of the three PEAD variables, and then take their simple average as the stock's ranked PEAD value. When taking the simple average, we use the available NYSE percentiles, allowing us to extend the sample backward to January 1967. This approach follows Stambaugh and Yuan (2017).

We use the same approach to replicate the financing factor (FIN) by combining the net share issuance and the composite share issuance in annual sorts. At the end of June of each year  $t$ , net share issuance is the natural log of the ratio of split-adjusted shares outstanding for fiscal year ending in calendar year  $t - 1$  (the common share outstanding, Compustat annual item CSHO, times the adjustment factor, item AJEX) to the split-adjusted shares outstanding for fiscal year ending in  $t - 2$ . The composite share issuance is the log growth rate of the market equity not attributable to stock return,  $\log (Me_t / Me_{t-5}) - r(t - 5, t)$ , in which  $r(t - 5, t)$  is the cumulative log stock return from the last trading day of June in year  $t - 5$  to the last trading day of June in year  $t$ , and  $Me_t$  is the market equity from CRSP on the last trading day of June in year  $t$ .

Finally, armed with the composite FIN and PEAD scores, we split stocks based on their NYSE breakpoints of the 30th and 70th percentiles in double  $2 \times 3$  sorts with size.