An Implementation of Quality Minus Junk

Ryan Kwon, Anthoney Tsou

February 11, 2015

1 Abstract

The qmj package produces quality scores for companies based on the work of Asness et. al (2013). It measures the quality of each of the 3000 largest US companies from the Russell 3000 Index based on profitability, growth, safety, and payouts, using the latest available data from Google Finance. The package includes tools to automatically gather relevant financial documents and stock data, allowing users to update their data whenever desired. The package also provides utilities for analyzing the scores of individual companies, various plotting and filtering tools, and generally helps separate the list of companies into "junk" stocks, which are expected to underperform relative to the market, and "quality" stocks, which are expected to outperform.

Introduction

qmj implements the methodology of the work of *Asness et. al* (2013). Within the paper, Asness uses several financial measures to calculate the relative profitability, growth, safety, and payouts of a company within a given universe, which they use to provide an overall quality score for a company.

This quality score is used as the basis of a portfolio which longs quality companies, which are likely to outperform the market, while "junk" companies shorted as they are expected to underperform relative to the market.

qmj provides tools to practically apply their results, coming equipped with pre-compiled recent data in addition to providing tools to automatically update or analyze that information.

Data

We demonstrate the use of **qmj** based on a subset of companies from the Russell 3000 Index.

```
> library(qmj)
> data(companies) #Stores company names and tickers from the
> #Russell 3000 index
> data(financials) #Stores financial documents for the given
> #list of companies.
> data(prices) # Stores price returns and closing stock prices
> #for the past two years.
> data(quality) #Stores the quality scores and the scores of
> #its components.

> #And more detailed data sets into what makes up quality
> data(profitability)
> data(growth)
> data(payouts)
> data(safety)
```

Though **qmj** keeps datasets updated, it has a few functions that can extract information directly from Google Finance to grab the most recent data.

```
> #raw_prices <- get_prices(companies)
> #raw_data <- get_info(companies)
```

get_prices() takes a data frame of companies, organized by name and ticker, and returns the daily prices and returns for the past two years including the most recent trading day. **get_info()** also takes a data frame of companies, organized by name and ticker, and grabs the most recent company 10-K financial statements. Thus, **get_info()** does not need to be called often since it will only grab new data once per year. Both functions will return a data frame that can be organized easily. An easy way to make the data more readable is through tidy functions in the **qmj** package.

```
> #clean_prices <- tidy_prices(raw_prices)
> #clean_data <- tidyinfo(raw_data)
```

tidy_prices() takes as input the result of **get_prices()**, which is assigned here as **raw_prices**, and organizes the data into columns for ticker, date, price, and price return. **tidyinfo()** takes as input the result of **get_info()**, which is assigned here as **raw_data**, and organizes the data into columns for ticker, year, and various items found in company financial statements such as total assets and net income. The column names themselves are abbreviations that are used in the Appendix.

Conclusion

In the **qmj** package, we automate AQR's method of assigning quality scores for publicly traded companies in today's market. The package itself provides convenient datasets and utility functions, and it also takes advantage of R's robust nature to allow seamless interaction with functions in the base R package and other packages.

Ryan Kwon
Williams College
Williamstown, MA
USA
rhk1@williams.edu

Anthoney Tsou
Williams College
Williamstown, MA
USA
at6@williams.edu

Bibliography

Asness, Clifford S., Andrea Frazzini, and Lasse H. Pedersen. "Quality Minus Junk." AQR (2013)

Appendix

We calculate quality scores for publicly traded companies in the Russell 3000 Index by summing the z-scores for each company's profitability, growth, safety, and payouts. We attempt to perform the same calculations as AQR does, but we have a few adjustments given the availability of data from public sources.

Profitability

Profitability is composed of six variables: gross profits over assets ($GPOA$), return on equity (ROE), return on assets (ROA), cash flow over assets ($CFOA$), gross margin ($GMAR$), and accruals (ACC). $GPOA$ is calculated as gross profits ($GPROF$) over total assets (TA).

$$GPOA = \frac{GPROF}{TA}$$

ROE is calculated as net income (NI) over book equity (BE), which is shareholders' equity (the difference of Total Liabilities and Shareholders' Equity ($TLSE$) with Total Liabilities (TL)) - preferred stock (the sum of redeemable preferred stock (RPS) and non redeemable preferred stock ($NRPS$)).

$$ROE = \frac{NI}{BE}$$

ROA is calculated as NI over TA.

$$ROA = \frac{NI}{TA}$$

$CFOA$ is calculated as NI + depreciation ($DP.DPL$) - changes in working capital (CWC) - capital expenditures (CX) all over TA.

$$CFOA = \frac{NI + DP.DPL - CWC - CX}{TA}$$

$GMAR$ is calculated as $GPROF$ over total revenue ($TREV$).

$$GMAR = \frac{GPROF}{TREV}$$

Finally, ACC is calculated as $DP.DPL$ - CWC all over TA.

$$ACC = \frac{DP.DPL - CWC}{TA}$$

We then standardize all components of profitability to z-scores and then standardize all profitability scores into z-scores.

$$Profitability = z(z_{gpoa} + z_{roe} + z_{roa} + z_{cfoa} + z_{gmar} + z_{acc})$$

Growth

Growth is measured by differences in profitability across a time span of four years. Though AQR recommends measuring growth across a time span of five years, public information that is both consistent and well-organized in 10-K forms is only available for a time span of four years, and it is still too early in the most recent year (2015) for most companies to have submitted a 10-K form. Thus, we measure growth using a time span of four years, which we will update once this year's 10-K form is submitted for each company in the Russell 3000 Index. As of now,

$$Growth = z(z_{\Delta gpoa_{t,t-4}} + z_{\Delta roe_{t,t-4}} + z_{\Delta roa_{t,t-4}} + z_{\Delta cfoa_{t,t-4}} + z_{\Delta gmar_{t,t-4}} + z_{\Delta acc_{t,t-4}})$$

Safety

Safety is composed of six variables: beta (BAB), idiosyncratic volatility ($IVOL$), leverage (LEV), Ohlson's O (O), Altman's Z (Z), and earnings volatility ($EVOL$). BAB is calculated as the negative covariance of each company's daily price returns ($pret_{c_i}$) relative to the benchmark daily market price returns ($pret_{mkt}$), in this case the S&P 500, over the variance of $pret_{mkt}$.

$$BAB = \frac{-cov(pret_{c_i}, pret_{mkt})}{var(pret_{mkt})}$$

$IVOL$ is the standard deviation of daily beta-adjusted excess returns. In other words, $IVOL$ is found by running a regression on each company's price returns and the benchmark, then taking the standard deviation of the residuals. Leverage is -(total debt (TD) over TA).

$$Leverage = -\frac{TD}{TA}$$

$$O = -(-1.32 - 0.407 * log\left(\frac{ADJASSET}{CPI}\right) + 6.03 * TLTA - 1.43 * WCTA$$
$$+ 0.076 * CLCA - 1.72 * OENEG - 2.37 * NITA - 1.83 * FUTL$$
$$+ 0.285 * INTWO - 0.521 * CHIN)$$

$ADJASSET$ is adjusted total assets, which is $TA + 0.1$ * (market equity (ME, calculated as average price per share for the most recent year * total number of shares outstanding ($TCSO$) - BE)).

$$ADJASSET = TA + 0.1 * (ME - BE)$$

CPI, the consumer price index, is assumed to be 100, since we only care about the most recent year. $TLTA$ is book value of debt (BD, calculated as TD - minority interest (MI) - (RPS + $NRPS$)) over $ADJASSET$.

$$TLTA = \frac{BD}{ADJASSET}$$

$WCTA$ is current assets (TCA) - current liabilities (TCL) over TA.

$$WCTA = \frac{TCA - TCL}{TA}$$

$CLCA$ is TCL over TCA.

$$CLCA = \frac{TCL}{TCA}$$

$OENEG$ is a dummy variable that is 1 if total liabilities (TL) is greater than TA.

$$OENEG = TL > TA$$

$NITA$ is NI over TA.

$$NITA = \frac{NI}{TA}$$

$FUTL$ is income before taxes (IBT) over TL.

$$FUTL = \frac{IBT}{TL}$$

$INTWO$ is another dummy variable that is 1 if NI for the current year and NI for the previous year are both negative.

$$INTWO = MAX(NI_t, NI_{t-1}) < 0$$

$CHIN$ is NI for the current year - NI for the previous year all over the sum of the absolute value of NI for the current year and the absolute value of NI for the previous year

$$CHIN = \frac{NI_t - NI_{t-1}}{|NI_t| + |NI_{t-1}|}$$

Altman's Z is calculated using weighted averages of working capital (WC, calculated as TCA - TCL),

$$WC = TCA - TCL$$

retained earnings (RE, calculated as NI - dividends per share ($DIVC$) * $TCSO$),

$$RE = NI - DIVC * TCSO$$

earnings before interest and taxes ($EBIT$, calculated as NI - Discontinued Operations(DO) + (IBT - income after tax (IAT)) + interest expense ($NINT$)),

$$EBIT = NI - DO + (IBT - IAT) + NINT$$

ME, and $TREV$, all over TA.

$$Z = \frac{1.2 * WC + 1.4 * RE + 3.3 * EBIT + 0.6 * ME + TREV}{TA}$$

$EBIT$ is likely an overestimate for a given company due to potentially missing information. $EVOL$ is calculated as the standard deviation of ROE for a four year span. AQR recommends the past five years, but for the same reason stated in the Growth section, we use a four year span.

$$EVOL = \sigma \left(\sum_{i=t-4}^{t} ROE_i \right)$$

Likewise, we standardize each variable and then standardize each safety measure, so

$$Safety \ = \ z(z_{bab} \ + \ z_{ivol} \ + \ z_{lev} \ + \ z_o \ + \ z_z \ + \ z_{evol})$$

Payouts

Payouts is composed of three variables: net equity issuance ($EISS$), net debt issuance ($DISS$), and total net payout over profits ($NPOP$). $EISS$ is calculated as the negative log of the ratio of $TCSO$ of the most recent year and $TCSO$ of the previous year.

$$EISS \ = \ -log \left(\frac{TCSO_t}{TCSO_{t-1}} \right)$$

Though AQR uses split-adjusted number of shares, we are currently using $TCSO$ given available information and will adjust for splits in future iterations of qmj. $DISS$ is calculated as the negative log of the ratio of TD of the most recent year and TD of the previous year.

$$DISS \ = \ -log \left(\frac{TD_t}{TD_{t-1}} \right)$$

$NPOP$ is calculated as NI - ΔBE over a four year span all over sum of $GPROF$ for the past four years (for the same reason as explained in the Growth section).

$$NPOP \ = \ \frac{NI - \Delta BE}{\sum_{i=t-4}^{t} GPROF_i}$$