1 Empirical Application

1.1 Data

In the empirical section, for the return of assets, we use the monthly excess returns from Standard Poor (S&P) 500 index component companies.[1] We prepared three data sets for different time spams: 10 years (January 2008 to December 2017), 20 years (January 1998 to December 2017), and 30 years (January 1989 to December 2017). Because of the components companies of the index is constantly changing, for each of the datasets, the companies amount (n) is different, the dimensions of the data set is showing in the table (1).

explain why some stocks are missing: they are not going public.
some missing spots full fill those gaps

Table 1: Data Set Dimensions

	Time Spam	Companies Amount (n)	Observations Amount (T)
10 Years	January 2008 - December 2017	419	120
20 Years	January 1998 - December 2017	342	240
30 Years	January 1988 - December 2017	242	360

The one-month U.S. treasury bill return rate was set as the risk free return r_{ft}. For company i, we calculates the companies return at month t (r_{it}) use the following formula:

$$r_{it} = \frac{p_{it} - p_{it-1}}{p_{it-1}}$$

and calculate the access return $x_{it} = r_{it} - r_{ft}$. Here the p_{it} and p_{it-1} are the company's close stock price at the first day of month t and t-1. The price is adjusted for the dividends and splits.[2]

With regard of the factors, we use 146 different risk factors, including the market factors as market return minus risk free rate form Feng, Giglio, and Xiu (2020)

[1] The data was obtained from the Global Finance Data, Osiris, and Yahoo Finance

[2] The data is adjusted base on the Central for Research in Security Price (CRSP) method.

References

Feng, G., Giglio, S., & Xiu, D. (2020, 6). Taming the factor zoo: A test of new factors. *The Journal of Finance*, *75*, 1327-1370. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12883 doi: 10.1111/jofi.12883