**Replication data and code for**

# What is the Expected Return on a Stock?

Ian Martin Christian Wagner

September, 2018

We provide data and R-code to replicate the main results of our paper "What is the Expected Return on a Stock?", forthcoming in the *Journal of Finance*.

- The code has been tested using R version 3.3.3 and requires the R-packages 'car', 'latex2exp', 'MASS', and 'plm'.
- The Rdata-file 'SVIX_replication_data.Rdata' contains
  - the data.frame 'data', which contains monthly data for the empirical analysis of one-year horizon returns of S&P 500 firms.
  - the data.frame 'data.daily', which contains daily data used to produce some descriptive statistics.
  - the variable 'rx.maturity', which takes the value 12, indicating that the maturity of options and the horizon for returns is 12 months.
- The R-code replicating our empirical results is in the file 'SVIX_replication_master.r'; it uses some functions defined in the file 'SVIX_replication_functions.r'.
- The code is designed to offer maximum transparency (hopefully also useful for readers who do not use R) rather than maximum efficiency. The code and the functions are tailormade for the replication exercise using the aforementioned datafiles, is not intended for other purposes, and comes with no warranty. Furthermore, the rules of the JF codesharing policy apply.

The R-code replicates the main results of our paper. More specifically, it replicates all results of our empirical analysis of one-year horizon returns of S&P 500 firms, with the exception of analyses that require data which we cannot provide due to copyright protection. In these cases, we explain in the code comments how users who have access to these data can extend the code accordingly. Table I below summarizes how the notation from the paper translates into variable names in the R-code.

Table I: Notation in the paper and variable names in the R-code

| Variable | Paper | R-code |
|---|---|---|
| Panel A. Variables in data.frame 'data' as stored in 'SVIX_replication_data.Rdata' | | |
| Time | $t$ | date |
| Firm identifier | $i$ | gvkey |
| Risk-neutral variance of firm $i$ | $\text{SVIX}^2_{i,t}$ | svix2 |
| Average risk-neutral variance | $\overline{\text{SVIX}}^2_t$ | basket.svix2 |
| Risk-neutral variance of the index | $\text{SVIX}^2_t$ | index.svix2 |
| Market cap weight | $\text{w}_{i,t}$ | mkt.cap.weight |
| CAPM beta | $\text{Beta}_{i,t}$ or $\widehat{\beta}_{i,t}$ | beta |
| Book-to-market | $\text{B/M}_{i,t}$ | btm |
| Realized excess-of-market return of firm $i$: | | |
| – regression analysis | $\frac{R_{i,t+1}-R_{m,t+1}}{R_{f,t+1}}$ | rxm |
| – out-of-sample comparison with other forecasts | $R_{i,t+1}-R_{m,t+1}$ | rxm.os |
| Realized excess return of firm $i$: | | |
| – regression analysis | $\frac{R_{i,t+1}-R_{f,t+1}}{R_{f,t+1}}$ | rx |
| – out-of-sample comparison with other forecasts | $R_{i,t+1}-R_{f,t+1}$ | rx.os |
| Riskless rate of return | $R_{f,t+1}-1$ | rf |
| Panel B. Some other variables frequently used in the code | | |
| Risk-neutral excess variance of firm $i$ | $\text{SVIX}^2_{i,t}-\overline{\text{SVIX}}^2_t$ | svix2_basket.svix2 |
| Expected excess-of-market return of firm $i$: | | |
| – regression-related analysis | $\frac{\mathbb{E}_t\, R_{i,t+1}-R_{m,t+1}}{R_{f,t+1}}$ | E.rxm |
| – out-of-sample comparison with other forecasts | $\mathbb{E}_t\, R_{i,t+1}-R_{m,t+1}$ | E.rxm.os |
| Expected excess return of firm $i$: | | |
| – regression-related analysis | $\frac{\mathbb{E}_t\, R_{i,t+1}-R_{f,t+1}}{R_{f,t+1}}$ | E.rx |
| – out-of-sample comparison with other forecasts | $\mathbb{E}_t\, R_{i,t+1}-R_{f,t+1}$ | E.rx.os |