Robust Regression and Fama-French 1992 Revisited

R. Douglas Martin
Professor Emeritus
Applied Mathematics and Statistics
University of Washington

Colloquia Presentation
Texas Tech University
February 19, 2021

Outline

Main Reference and R Package

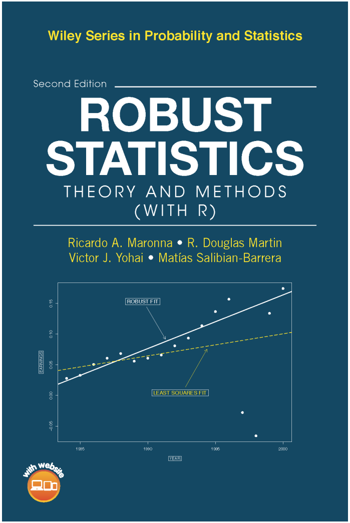
Robust Statistics: Theory and Methods (2019), 2nd ed.
Maronna, Martin, Yohai and Salibian-Barrera (MMYS)

Companion R Package: RobStatTM

To install, load, and view functions and data sets from R:

```
> install.packages("RobStatTM")
> library(RobStatTM)
> ls("package:RobStatTM")
> data(package = "RobStatTM")
```

http://www.wiley.com/go/maronna/robust



ROBUST ESTIMATION OVERVIEW

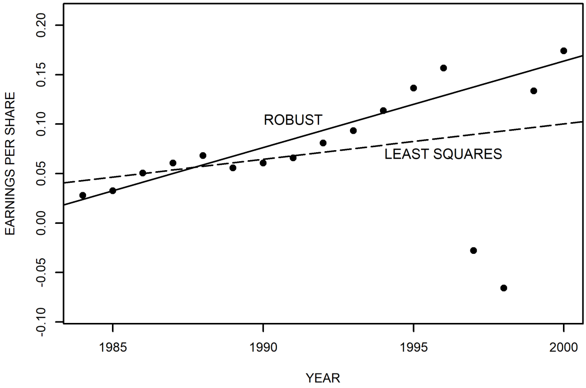
Parametric Models Focus

Location, scale, and linear regression models

1. **Data oriented defintion of a robust estimator**
   - Parameter estimates are not much influenced by outliers
   - Results in a good fit to the bulk of the data
   - Reliable detection of multivariate outlier
   - Confidence intervals and tests are not much influenced by outliers

2. **Theoretical foundation**
   - Good estimator performance at normal and “nearly normal” distributions
   - Universal use of two component nearly normal mixture distributions
   - Performance measures: asymptotic variance and bias
   - Supported by extensive finite sample mean-square error Monte Carlo

Simple Example



- Robust regression provides a good fit to the bulk of the data
- Most robust regression residuals are smaller than for LS, with two robust residuals are larger than for LS

Huber Location M-estimators (1964)

Location and Scale Model

$$r_t = \mu + s \cdot \epsilon_t, \quad t = 1, 2, \ldots, T, \qquad \text{i.i.d. } \epsilon_t \overset{d}{\sim} f_0$$

M-estimator:

$$\hat{\mu} = \underset{\mu}{argmin} \sum_{t=1}^{n} \rho\left(\frac{r_t - \mu}{s}\right)$$

$$\sum_{t=1}^{n} \psi\left(\frac{r_t - \hat{\mu}}{s}\right) = 0, \quad \psi = \rho'$$

MLE: $\rho = -log f_0$