Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 895

May 2007

A Note on the Coefficient of Determination in Models with Infinite Variance Variables

Jeong-Ryeol Kurz-Kim

Mico Loretan

A Note on the Coefficient of Determination in Models with Infinite Variance Variables*

Jeong-Ryeol Kurz-Kim† Mico Loretan‡

May 2007

Abstract

Since the seminal work of Mandelbrot (1963), α-stable distributions with infinite variance have been regarded as a more realistic distributional assumption than the normal distribution for some economic variables, especially financial data. After providing a brief survey of theoretical results on estimation and hypothesis testing in regression models with infinite-variance variables, we examine the statistical properties of the coefficient of determination in models with α-stable variables. If the regressor and error term share the same index of stability $\alpha < 2$, the coefficient of determination has a nondegenerate asymptotic distribution on the entire $[0, 1]$ interval, and the density of this distribution is unbounded at 0 and 1. We provide closed-form expressions for the cumulative distribution function and probability density function of this limit random variable. In contrast, if the indices of stability of the regressor and error term are unequal, the coefficient of determination converges in probability to either 0 or 1, depending on which variable has the smaller index of stability. In an empirical application, we revisit the Fama-MacBeth two-stage regression and show that in the infinite-variance case the coefficient of determination of the second-stage regression converges to zero in probability even if the slope coefficient is nonzero.

KEYWORDS: Regression models, α-stable distributions, infinite variance, coefficient of determination, Fama-MacBeth regression, Monte Carlo simulation, signal-to-noise ratio, density transformation theorem.

JEL CLASSIFICATIONS: C12, C13, C21, G12

*The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the staff of the Deutsche Bundesbank, the Board of Governors of the Federal Reserve System, or of any other person associated with the Federal Reserve System. We are grateful to Jean-Marie Dufour, Neil R. Ericsson, Peter C.B. Phillips, Werner Ploberger, Jonathan H. Wright, and participants of a workshop at the Federal Reserve Board for valuable comments, and to Zhenyu Wang for the data used in the empirical section of this paper.

†Corresponding author. Research support from the Alexander von Humboldt Foundation is gratefully acknowledged. Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany; Email: jeong-ryeol.kurz-kim@bundesbank.de; Tel: +49-69-9566-4576, Fax: +49-69-9566-2982.

‡Mailstop 18, Board of Governors of the Federal Reserve System, Washington DC 20551, USA; Email: mico.loretan@frb.gov; Tel: +1-202-452-2219; Fax: +1-202-263-4850.

1 Introduction

Granger and Orr (1972) begin their article, "'Infinite variance' and research strategy in time series analysis," by questioning the uncritical use of the normal distribution assumption in economic modelling and estimation:

> It is standard procedure in economic modelling and estimation to assume that random variables are normally distributed. In empirical work, confidence intervals and significance tests are widely used, and these usually hinge on the presumption of a normal population. Lately, there has been a growing awareness that some economic data display distributional characteristics that are flatly inconsistent with the hypothesis of normality.

Due in part to the influential seminal work of Mandelbrot (1963), α-stable distributions are often considered to provide the basis for more realistic distributional assumptions for some economic data, especially for high-frequency financial time series such as those of exchange rate fluctuations and stock returns. Financial time series are typically fat-tailed and excessively peaked around their mean—phenomena that can be better captured by α-stable distributions with $1 < \alpha < 2$ rather than by the normal distribution, for which $\alpha = 2$.[1] The α-stable distributional assumption with $\alpha \leq 2$ is thus a generalization of rather than an alternative to the Gaussian distributional assumption. If an economic series fluctuates according to an α-stable distribution with $\alpha < 2$, it is known that many of the standard methods of statistical analysis, which often rest on the asymptotic properties of sample second moments, do not apply in the conventional way. In particular, as we demonstrate in this paper, the coefficient of determination—a standard criterion for judging goodness of fit in a regression model—has several nonstandard statistical properties if $\alpha < 2$.

The linear regression model is one of the most commonly used and basic econometric tools, not only for the analysis of macroeconomic relationships but also for the study of financial market data. Typical examples for the latter case are estimation of the *ex-post* version of the capital asset pricing model (CAPM) and the two-stage modelling approach of Fama and MacBeth (1973). Because of the prevalence of heavy-tailed distributions in financial time series, it is of interest to study how regression models perform when the data are heavy-tailed rather normally distributed.

The first purpose of the present paper is to survey theoretical results of estimation and hypothesis testing in regression models with infinite-variance distributions, and the second is to establish that infinite variance of the regression variables has important consequences for the statistical properties of the coefficient of determination and tests of the hypothesis that this coefficient is equal to zero. Third, we revisit the Fama-

[1] The normal distribution is the only member of the family of α-stable distributions that has finite second (and higher-order) moments; all other members of this family have infinite variance.

MacBeth two-stage regression approach and demonstrate that infinite variance of the regression variables can affect decisively the interpretation of the empirical results.

The rest of our paper is structured as follows. In Section 2 we provide a brief summary of the properties of α-stable distributions and of aspects of estimation, hypothesis testing, and model diagnostic checking in regression models with infinite-variance regressors and disturbance terms. Section 3 provides a detailed analysis of the asymptotic properties of the coefficient of determination in regression models with infinite-variance variables. In our empirical application, presented in Section 4, we revisit the data used in Fama and French (1992), and we show that the statistical and/or economic interpretation of their findings can be quite different under the maintained assumption of α-stable distributions from an interpretation based on the assumption of normal distributions. Section 5 summarizes the paper and offers some concluding remarks.

2 Framework

2.1 α-stable distributions

A random variable X is said to have a stable distribution if, for any positive integer $n > 2$, there exist constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that $X_1 + \cdots + X_n \stackrel{d}{=} a_n X + b_n$, where $X_1, \ldots, X_n$ are independent copies of X and $\stackrel{d}{=}$ signifies equality in distribution. The coefficient a_n above is necessarily of the form $a_n = n^{1/\alpha}$ for some $\alpha \in (0, 2]$ (see Feller, 1971, Section VI). The parameter α is called the index of stability of the distribution, and a random variable X with index α is called α-stable. An α-stable distribution is described by four parameters and will be denoted by $S(\alpha, \beta, \gamma, \delta)$. Closed-form expressions for the probability density functions of α-stable distributions are known to exist only for three special cases.[2] However, closed-form expressions for the characteristic functions of α-stable distributions are readily available. One parameterization of the logarithm of the characteristic function of $S(\alpha, \beta, \gamma, \delta)$ is

$$\ln\left(\mathrm{E}\, e^{i\tau X}\right) = i\delta\tau - \gamma^\alpha |\tau|^\alpha \left(1 + i\beta \operatorname{sign}(\tau)\, \omega(\tau, \alpha)\right), \tag{1}$$

where $\operatorname{sign}(\tau) = -1$ for $\tau < 0$, $\operatorname{sign}(\tau) = 0$ for $\tau = 0$, and $\operatorname{sign}(\tau) = +1$ for $\tau > 0$; and $\omega(\tau, \alpha) = -\tan(\pi\alpha/2)$ for $\alpha \neq 1$ and $\omega(\tau, \alpha) = (2/\pi) \ln|\tau|$ for $\alpha = 1$.

The tail shape of an α-stable distribution is determined by its index of stability $\alpha \in (0, 2]$. Skewness is governed by $\beta \in [-1, 1]$; the distribution is symmetric about δ if and only if $\beta = 0$. The scale and location parameters of α-stable distributions are denoted by $\gamma > 0$ and $\delta \in \mathbb{R}$, respectively. When $\alpha = 2$, the log characteristic function given by equation (1) reduces to $i\delta\tau - \gamma^2\tau^2$, which is that of a Gaussian random

[2] The three special cases are: (i) the Gaussian distribution $S(2, 0, \gamma, \delta) \equiv N(\delta, 2\gamma^2)$, (ii) the symmetric Cauchy distribution $S(1, 0, \gamma, \delta)$, and (iii) the Lévy distribution $S(0.5, \pm 1, \gamma, \delta)$; see Zolotarev (1986), Section 2, and Rachev et al. (2005), Section 7.

variable with mean δ and variance $2\gamma^2$. For $\alpha < 2$ and $|\beta| < 1$, the tail properties of an α-stable random variable X satisfy

$$\lim_{x\to\infty} \mathrm{P}(X > x) = \big[C(\alpha)\,\gamma^\alpha(1+\beta)/2\big]x^{-\alpha} \quad \text{and} \tag{2}$$

$$\lim_{x\to\infty} \mathrm{P}(X < -x) = \big[C(\alpha)\,\gamma^\alpha(1-\beta)/2\big]x^{-\alpha}\,, \tag{3}$$

i.e., both tails of the probability density function (pdf) of X are asymptotically Paretian. For $\alpha < 2$ and $\beta = +1$ (-1), the distribution is maximally right-skewed (left-skewed) and only the right (left) tail is asymptotically Paretian.[3] The term $C(\alpha)$ in equations (2) and (3) is given by

$$C(\alpha) = \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} \qquad \text{for } \alpha \neq 1 \tag{4}$$

and $2/\pi$ for $\alpha = 1$; see, e.g., Samorodnitsky and Taqqu (1994), p. 17. The function $C(\alpha)$, which is shown in Figure 1, is continuous and strictly decreasing in $\alpha \in (0,2)$, with $\lim_{\alpha\downarrow 0} C(\alpha) = 1$ and $\lim_{\alpha\uparrow 2} C(\alpha) = 0$.[4] In consequence, even though all stable distributions with $\alpha < 2$ have asymptotically Paretian tails, as $\alpha \uparrow 2$ proportionately less and less of the distribution's probability mass is located in the tail region. In addition, the density's tails decline at an increasingly rapid rate as $\alpha \uparrow 2$, thereby limiting the likelihood of observing very large draws conditional on the draw coming from the tail region. These observations explain why potentially very large sample sizes are required if one desires to estimate the index of stability with adequate precision if α is close to but smaller than 2.

Figure 1 somewhere here

Because $\mathrm{E}\,|X|^\xi = \lim_{b\to\infty} \int_0^b \mathrm{P}(|X|^\xi > x)\,\mathrm{d}x$, it follows that $\mathrm{E}\,|X|^\xi < \infty$ for $\xi \in (0,\alpha)$ and $\mathrm{E}\,|X|^\xi = \infty$ for $\xi \geq \alpha$ if X is α-stable with $\alpha \in (0,2)$.[5] Only moments of order up to but not including α are finite if $\alpha < 2$, and a non-Gaussian stable distribution's index of stability is also equal to its maximal moment

[3] For $\alpha < 1$ and $\beta = +1$, $\mathrm{P}(X < \delta) = 0$, i.e., the distribution's support is bounded below by δ. Zolotarev (1986, Theorem 2.5.3) and Samorodnitsky and Taqqu (1994, pp. 17–18) provide expressions for the rate of decline of the non-Paretian tail if $\beta = \pm 1$ and $\alpha \geq 1$.

[4] The function $C(\alpha)$ is smooth on the entire interval $(0,2)$. The numerator and the second term in the denominator of equation (4) both converge to 0 as $\alpha \to 1$; $C(1) = 2/\pi$ therefore follows from an application of L'Hôpital's Rule.

[5] Ibragimov and Linnik (1971, Theorem 2.6.4) show that this result holds not only for α-stable distributions, but that it pertains to *all* distributions that are in the domain of attraction of an α-stable distribution. Ibragimov and Linnik (1971, Theorem 2.6.1) provide necessary and sufficient conditions for a probability distribution to lie in the domain of attraction of an α-stable law.

exponent.[6] In particular, if $\alpha \in (1, 2)$, the variance is infinite but the mean exists. For $\alpha > 1$, it follows that $\mathrm{E}(X) = \delta$; in addition, for $\beta = 0$, δ is equal to the distribution's mode and median irrespective of the value of α, justifying the use of the term "central location parameter" for δ in the finite-mean or symmetric cases. In addition, for $\alpha \neq 1$, one can show that $S(\alpha, \beta, \gamma, \delta) \stackrel{d}{=} \gamma \cdot S(\alpha, \beta, 1, \delta/\gamma)$.[7] We make use of this property below in the derivations of Theorem 1 and Remark 3.

The class of α-stable distributions is an interesting distributional candidate for disturbances in regression models because (i) it is able to capture the relative frequencies of extreme vs. ordinary observations in the economic variables, (ii) it has the convenient statistical property of closure under convolution, and (iii) only α-stable distributions can serve as limiting distributions of sums of independent and identically distributed (iid) random variables, as proven in Zolotarev (1986). The latter two properties are appealing for regression analysis, given that disturbances can be viewed as random variables which represent the sum of all external effects not captured by the regressors. For more details on the properties of α-stable distributions, we refer to Gnedenko and Kolmogorov (1954), Feller (1971), Zolotarev (1986), and Samorodnitsky and Taqqu (1994). The role of the α-stable distribution in financial market and econometric modelling is surveyed in McCulloch (1996) and Rachev et al. (1999).

2.2 Regression models with infinite-variance variables

Let X and Y be two jointly symmetric α-stable (henceforth, $S\alpha S$) random variables with $\alpha > 1$, i.e., we require X and Y to have finite means. Our main reason for concentrating on the case $\alpha > 1$ lies in its empirical relevance. Estimated maximal moment exponents for most empirical financial data, such as exchange rates and stock prices, are generally greater than 1.5; see, for example, de Vries (1991) and Loretan and Phillips (1994). An econometric (purposeful) reason for studying the case $\alpha > 1$ is that, for α-stable distributions with $\alpha > 1$, regression analysis that is based on sample second moments, such as least squares, is still asymptotically consistent for the regression coefficients, even though the limit distributions of these regression coefficients are nonstandard.[8] Suppose that the regression of a random variable Y on a random variable X is linear, i.e., there exists a constant θ such that

$$\mathrm{E}(\,Y \mid X\,) = \theta X \quad a.s., \tag{5}$$

[6] The maximal moment exponent of a distribution is either a finite positive number, or it is infinite if a distribution has finite moments of all orders. For a Student-t distribution, the degrees of freedom parameter is equal to its maximal moment exponent.

[7] This result also holds for the case $\alpha = 1$ *and* $\beta = 0$.

[8] Another reason for this restriction comes from the viewpoint of statistical modelling. The conditional expectation of the bivariate symmetric stable distribution in (5) is, as in the Gaussian case, linear in X only if $\alpha \in (1, 2)$. The regression function is in general nonlinear, or rather only asymptotically linear, under other conditions. For more on bivariate linearity, see Samorodnitsky and Taqqu (1994, Sections 4 and 5).

with

$$\theta = \frac{[Y, X]_\alpha}{{\gamma_x}^\alpha} X,$$

where γ_x is the scale parameter of the $S\alpha S$ random variable X and $[\cdot, \cdot]_\alpha$ in the numerator is covariation (covariance in the Gaussian case), which can be calculated as $\mathrm{E}\big(XY^{<\xi-1>}\big) / \mathrm{E}\big(|Y|^\xi\big)$, for all $\xi \in (1, \alpha)$ with $a^{<\xi>} \equiv |a|^\xi \mathrm{sign}(a)$.

For estimation and diagnostics, the relation (5) can be written as a regression model with a constant term,

$$y_t = c + \theta x_t + u_t, \tag{6}$$

where the maintained hypothesis is that u_t is iid $S\alpha S$, with $\alpha \in (1, 2]$. The econometric issues of interest are to estimate θ properly, to test the hypothesis of significance for the estimated parameter, usually based on the t-statistic, as well as to compute model diagnostics, such as the coefficient of determination, the Durbin-Watson statistic, and the F-test of parameter constancy across subsamples.

The effects of infinite variance in the regressor and disturbance term can be substantial. If the variables share the same index of stability α, the ordinary least squares (OLS) estimate of θ is still consistent, but its asymptotic distribution is α-stable with the same α as the underlying variables. Furthermore, the convergence rate to the true parameter is $T^{(\alpha-1)/\alpha}$, smaller than the rate $T^{1/2}$ which applies in the finite-variance case. If $\alpha < 2$, OLS loses its best linear unbiased estimator (BLUE) property, i.e., it is no longer the minimum-dispersion estimator in the class of linear estimators of θ. In addition, the asymptotic efficiency of the OLS estimator converges to zero as the index of stability α declines to 1. Blattberg and Sargent (1971) (henceforth, BS) derived the BLUE for θ in (6) if the value of α is known. The BS estimator is given by

$$_\alpha\hat{\theta}_{BS} = \frac{\sum_{t=1}^T x_t^{<1/(\alpha-1)>} y_t}{\sum_{t=1}^T |x_t|^{\alpha/(\alpha-1)}}, \qquad 1 < \alpha \leq 2, \tag{7}$$

which coincides with the OLS estimator if $\alpha = 2$. Kim and Rachev (1999) prove that the asymptotic distribution of the BS estimator is also α-stable. Samorodnitsky et al. (2007) consider an optimal power estimate based on the BS estimator for unknown α, and they also provide an optimal linear estimator of the regression coefficients for various configurations of the indices of stability of x_t and u_t. Other efficient estimators of the regression coefficients have been studied as well; Kanter and Steiger (1974) propose an unbiased L_1-estimator, which excludes very large shocks in its estimation to avoid excess sensitivity due to outliers. Using a weighting function, McCulloch (1998) considers a maximum-likelihood estimator which is based on an approximation to a symmetric stable density.

Hypothesis testing is also affected considerably when the regressors and disturbance terms have infinite-variance stable distributions. For example, the t-statistic, commonly used to test the null hypothesis of parameter significance, no longer has a conventional Student-t distribution if $\alpha < 2$. Rather, as established by Logan et al. (1973), its pdf has modes at -1 and $+1$; for $\alpha < 1$ these modes are infinite. Kim (2003) provides empirical distributions of the t-statistic for finite degrees of freedom and various values of α by simulation. The usual applied goodness-of-fit test statistics, such as the likelihood ratio, Lagrange multiplier, and Wald statistics, also no longer have the conventional asymptotic χ^2 distribution, but have a stable χ^2 distribution, a term that was introduced by Mittnik et al. (1998).

In time series regressions with infinite-variance innovations, Phillips (1990) shows that the limit distribution of the augmented Dickey-Fuller tests for a unit root are functionals of Lévy processes, whereas they are functionals of Brownian motion processes in the finite-variance case. The F-test statistic for parameter constancy that is based on the residuals from a sample split test has an F-distribution in the conventional, finite-variance case. Kurz-Kim et al. (2005) obtain the limiting distribution of the F-test if the random variables have infinite variance. As shown by the authors, as well as by Runde (1993), the limiting distribution of the F-statistic for $\alpha < 2$ behaves completely differently from the Gaussian case: whereas in the latter case the statistic converges to 1 under the null as the degrees of freedom for both numerator and denominator of the statistic approach infinity, in the former case the statistic converges to a ratio of two independent, positive, and maximally right-skewed $\alpha/2$-stable distributions. This result is used below to derive closed-form expressions for the pdf and cumulative distribution function (cdf) of the limiting distribution of the R^2 statistic if the regressor and disturbance term share the same index of stability $\alpha < 2$.

Moreover, commonly used criteria for judging the validity of some of the maintained hypotheses of a regression model, such as the Durbin-Watson statistic and the Box-Pierce Q-statistic, would be inappropriate if one were to rely on conventional critical values. Phillips and Loretan (1991) study the properties of the Durbin-Watson statistic for regression residuals with infinite variance, and Runde (1997) examines the properties of the Box-Pierce Q-statistic for random variables with infinite variance. Loretan and Phillips (1994) and Phillips and Loretan (1994) establish that both the size of tests of covariance stationarity under the null and their rate of divergence of these tests under the alternative are strongly affected by failure of standard moment conditions; indeed, standard tests of covariance stationarity are *inconsistent* if population second moments do not exist.

3 Asymptotic properties of the coefficient of determination in models with α-stable regressors and error terms

3.1 Basic results

For the general asymptotic theory of stochastic processes with stable random variables, we refer to Resnick (1986) and Davis and Resnick (1985a, 1985b, 1986). Our results in this section are, in large part, an application of their work to the regression diagnostic context.

The maintained assumptions are:

1. The relationship between the dependent and independent variable conforms to the classical bivariate linear regression model,
$$y_t = c + \theta x_t + u_t, \quad t = 1, \ldots, T\,. \tag{8}$$
2. u_t is iid $S\alpha S(\alpha_u, 0, \gamma_u, 0)$, with $\alpha_u \in (1, 2)$.
3. x_t is exogenous and is also iid $S\alpha S(\alpha_x, 0, \gamma_x, 0)$, with $\alpha_x \in (1, 2)$.
4. The regressor and the error term have the same index of stability, i.e., $\alpha_x = \alpha_u = \alpha$.
5. The coefficients c and θ are consistently estimated by $\hat{c}$ and $\hat{\theta}$.[9]

The fourth assumption, that the regressor and the error term have the same index of stability, is rather strong, and its validity may be difficult to ascertain in empirical applications. In Corollary 2 below, we examine the consequences of having unequal values for the indices of stability for x_t and u_t for the asymptotic properties of the coefficient of determination.

The coefficient of determination measures the proportion of the total squared variation in the dependent variable that is explained by the regression:

$$R^2 = \frac{\text{Explained Sum of Squares}}{\text{Total Sum of Squares}} = \frac{\sum_{t=1}^{T}(\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2}\,.$$

Because $\hat{y}_t - \bar{y} = \hat{\theta}(x_t - \bar{x})$ and $y_t - \bar{y} = \hat{\theta}(x_t - \bar{x}) + \hat{u}_t$, where $\bar{y}$ and $\bar{x}$ are the respective sample averages of y_t and x_t, and because $\sum_{t=1}^{T}(x_t - \bar{x})\hat{u}_t$=0 by construction, the coefficient of determination may be written as

$$R^2 = \frac{\hat{\theta}^2 \sum_{t=1}^{T}(x_t - \bar{x})^2}{\hat{\theta}^2 \sum_{t=1}^{T}(x_t - \bar{x})^2 + \sum_{t=1}^{T}\hat{u}_t^2}\,. \tag{9}$$

[9] If $\alpha_x = \alpha_u$, OLS is known to generate consistent estimates of c and θ. See Samorodnitsky et al. (2007) for an overview and discussion of estimation methods that are consistent for various combinations of α_u and α_x.

Since x_t^2 and u_t^2 are in the normal domain of attraction of a stable distribution with index of stability $\alpha/2$, norming by $T^{-2/\alpha}$ rather than by T^{-1} is required to obtain non-degenerate limits for the sums of the squared variables. Because $\hat{\theta} \to_p \theta$ by the assumption of consistent estimation, an application of the law of large numbers to $\bar{x}$, the continuous mapping theorem, and the results of Davis and Resnick (1985b) yield the following expression for the joint limiting distribution of the elements in equation (9):

$$\begin{aligned}\left(T^{-2/\alpha}\gamma_u^{-2}\sum_{t=1}^{T}\hat{u}_t^2\,,\,\hat{\theta}^2T^{-2/\alpha}\gamma_x^{-2}\sum_{t=1}^{T}(x_t-\bar{x})^2\right) &\sim \left(T^{-2/\alpha}\gamma_u^{-2}\sum_{t=1}^{T}u_t^2\,,\,\theta^2T^{-2/\alpha}\gamma_x^{-2}\sum_{t=1}^{T}x_t^2\right)\\ &= \left(T^{-2/\alpha}\sum_{t=1}^{T}(u_t/\gamma_u)^2\,,\,\theta^2T^{-2/\alpha}\sum_{t=1}^{T}(x_t/\gamma_x)^2\right)\\ &\to_d \left(S_u\,,\,\theta^2S_x\right)\,. \end{aligned} \tag{10}$$

For $\alpha < 2$, the random variables S_u and S_x are independent, maximally right-skewed, and positive stable random variables with index of stability $\alpha/2 < 1$, $\beta = +1$, $\gamma = 1$,[10] $\delta = 0$, and log characteristic function

$$\ln \mathrm{E}\left(e^{i\tau S_x}\right) = \ln \mathrm{E}\left(e^{i\tau S_u}\right) = -|\tau|^{\alpha/2}\left(1 - i\,\mathrm{sign}(\tau)\,\tan(\pi\alpha/4)\right). \tag{11}$$

We therefore conclude that, under the five maintained assumptions of this section, the R^2 statistic of the regression model (8) has the following asymptotic distribution.

Theorem 1 *Under the maintained assumptions of the regression model in equation* (8)*, the coefficient of determination is distributed asymptotically as*

$$R^2 \to_d \frac{\theta^2\gamma_x^2S_x}{\theta^2\gamma_x^2S_x+\gamma_u^2S_u} = \frac{\eta S_x}{\eta S_x + S_u} = \frac{\eta Z}{\eta Z + 1} = \widetilde{R}(\alpha,\eta)\,,\ \textit{say,} \tag{12}$$

where $\eta = (\theta\gamma_x/\gamma_u)^2 \geq 0$[11] *and* $Z = S_x/S_u$*. For* $\alpha < 2$*,* S_x *and* S_u *are independent and are identically distributed with log characteristic functions given by equation* (11)*.*

Thus, for $\alpha < 2$ and $\eta > 0$, the coefficient of determination does *not* converge to a constant but has a nondegenerate asymptotic distribution on the interval $[0,1]$. This contrasts starkly with the standard, finite-variance result, which is stated here for completeness.

[10] To prove that $\gamma = 1$, see equation (13.3.14) on p. 529 of Brockwell and Davis (1991). In that equation, put $C = C(\alpha/2)$, where $C(\cdot)$ is given by equation (4), and employ the recursive relationship $\Gamma(2-\alpha/2) = (1-\alpha/2)\cdot\Gamma(1-\alpha/2)$.

[11] Observe that $\eta = 0$ if and only if $\theta = 0$, as the dispersion parameters γ_x and γ_u are necessarily positive.

Corollary 1 *If $\alpha = 2$, and hence if x_t and u_t have finite variance, the limit variables S_x and S_u in Theorem 1 are non-random constants and are, in fact, equal to 2.*[12] *In the finite-variance case, then, the limit of R^2 as $T \to \infty$ is given by*

$$R^2 \to_p \frac{\theta^2 \sigma_x^2}{\theta^2 \sigma_x^2 + \sigma_u^2} = \frac{\eta}{\eta + 1} ,$$

where now $\eta = (\theta \sigma_x / \sigma_u)^2$.

In the finite-variance case, the model's asymptotic signal-to-noise ratio, $\eta = (\theta \sigma_x / \sigma_u)^2$, is constant, as is therefore the limit of the coefficient of determination. In contrast, in the infinite-variance case the model's limiting signal-to-noise ratio is given by ηZ, where $\eta = (\theta \gamma_x / \gamma_u)^2$ and $Z = S_x / S_u$, and is therefore a random variable even asymptotically; it is this feature that causes the randomness of $\widetilde{R}(\alpha, \eta)$. We postpone a fuller discussion of the intuition that underlies this result to the end of this section, after we provide a detailed analysis of the statistical properties of $\widetilde{R}$.

Before doing so, however, we note that the fourth maintained assumption, i.e., that the indices of stability of the regressor and error term in (8) be the same, is crucial for obtaining the result that the asymptotic distribution of $\widetilde{R}$ is nondegenerate. Indeed, if the two indices of stability differ, the asymptotic properties of the R^2 statistic are as follows.

Corollary 2 *Suppose that the maintained assumptions of Theorem 1 apply except that $\alpha_x \neq \alpha_u$, i.e., suppose that the indices of stability of the regressor and error term are unequal. Let $\theta \neq 0$ to rule out the trivial case from further consideration. Then,*

- *if $\alpha_x < \alpha_u$, $1 - R^2 = o_p\left(T^{2/\alpha_u - 2/\alpha_x}\right)$; and*
- *if $\alpha_u < \alpha_x$, $R^2 = o_p\left(T^{2/\alpha_x - 2/\alpha_u}\right)$.*

Thus, R^2 converges to 1 in probability if $\alpha_x < \alpha_u$, and it converges to 0 in probability if $\alpha_u < \alpha_x$.

Proof. These results follow immediately from the fact that if $\alpha_x \neq \alpha_u$, different norming factors, *viz.*, T^{2/α_x} and T^{2/α_u}, are needed in equation (10) to achieve joint convergence of the terms $\hat{\theta} \sum_{t=1}^T (x_t - \bar{x})^2$ and $\sum_{t=1}^T \hat{u}_t^2$ to the limiting random variables S_x and S_u. Whenever the two norming factors differ, the larger of the two factors dominates the ratio that defines R^2 as $T \to \infty$, and this statistic must therefore converge either to 0 or 1 in probability.

Suppose first that $\alpha_x < \alpha_u$; since $T^{2/\alpha_x} > T^{2/\alpha_u}$, we find $T^{-2/\alpha_x} \sum \hat{u}_t^2 = T^{-2/\alpha_u}\left(T^{2/\alpha_u - 2/\alpha_x}\right) \sum \hat{u}_t^2 = o_p\left(T^{2/\alpha_u - 2/\alpha_x}\right)$. Therefore,

[12] Recall that in the finite-variance case, $\gamma^2 = \sigma^2/2$; therefore, norming by $T^{-1}\gamma_x^{-2}$ and $T^{-1}\gamma_u^{-2}$ in equation (10) produces a constant of 2.

$$
\begin{aligned}
R^2 &= \frac{\hat{\theta}^2 T^{-2/\alpha_x} \sum (x_t - \bar{x})^2}{\hat{\theta}^2 T^{-2/\alpha_x} \sum (x_t - \bar{x})^2 + T^{-2/\alpha_x} \sum \hat{u}_t^2} \\
&\to_d \frac{\theta^2 \gamma_x^2 S_x}{\theta^2 \gamma_x^2 S_x + o_p\left(T^{2/\alpha_u - 2/\alpha_x}\right)} \\
&\to_p 1 \, .
\end{aligned}
$$

Similarly, if $\alpha_u < \alpha_x$, $T^{-2/\alpha_u} \sum (x_t - \bar{x})^2 = o_p\left(T^{2/\alpha_x - 2/\alpha_u}\right)$, and $R^2 \to_p 0$. ■

Heuristically, if $\alpha_x \neq \alpha_u$ and $\theta \neq 0$, the limiting distribution of the R^2 statistic is degenerate at 0 or 1 because the model's asymptotic signal-to-noise ratio is either zero (if $\alpha_u < \alpha_x$) or infinite (if $\alpha_x < \alpha_u$). From an examination of the proof of this corollary, we can also deduce that if $\alpha_x \neq \alpha_u$, the fifth maintained assumption—that the regression coefficients are estimated consistently—could be relaxed, to require merely that an estimation method be employed that guarantees $\hat{\theta} \neq o_p(1)$; the result that R^2 converges either to 0 or 1 would continue to hold in this case.

3.2 Qualitative properties of $\widetilde{R}$

Returning to the main case of $\alpha_x = \alpha_u = \alpha$, we note that the random variable $\widetilde{R}$ is defined for all values of $\alpha \in (0, 2)$, even though in a regression context one would typically assume that $\alpha \in (1, 2)$. We now establish some important qualitative properties of $\widetilde{R}$.

Remark 1 *For $\eta > 0$, the median of $\widetilde{R}$, m, equals $\eta/(\eta + 1)$.*

Proof. For $\eta > 0$, observe that

$$
\begin{aligned}
\mathrm{P}\left(\widetilde{R} \leq \frac{\eta}{\eta + 1}\right) &= \mathrm{P}\left(\frac{\eta S_x}{\eta S_x + S_u} \leq \frac{\eta}{\eta + 1}\right) \\
&= \mathrm{P}\left(S_x \leq \frac{1}{\eta + 1}\left(\eta S_x + S_u\right)\right) \\
&= \mathrm{P}\big((\eta + 1) S_x - \eta S_x \leq S_u\big) \\
&= \mathrm{P}\left(S_x \leq S_u\right) \, .
\end{aligned}
$$

Because S_x and S_u are iid and have continuous cdfs, $\mathrm{P}(S_x \leq S_u) = 0.5$ by an application of Fubini's Theorem.[13] ■

[13] See, e.g., Resnick (1999, p. 155).

Thus, m is equal to the non-random limit of R^2 in the finite-variance case. Since S_x and S_u are positive *a.s.*, we also have $\mathrm{P}(S_x/S_u \le 1) \equiv \mathrm{P}(Z \le 1) = 0.5$, i.e., the median of Z is equal to 1, regardless of the value of α. As we will demonstrate rigorously later in this paper, the probability mass of Z is highly concentrated around 1 for values of α close to 2. Conversely, for small values of α, Z is *unlikely* to be close to 1; instead, it is very likely that one will obtain a draw of Z that is either very small, i.e., close to 0, or very large. A small or large draw of Z has a crucial effect on the model's signal-to-noise ratio, ηZ, and therefore also on R^2. This suggests that an informal measure of the effect of infinite variance in the regression variables on the value of R^2 in a given sample may be based on the *difference* between the model's coefficient of determination and a consistent estimate of its median m, say $\hat{m} = \hat{\eta}/(\hat{\eta}+1)$, where $\hat{\eta} = (\hat{\theta}\hat{\gamma}_x/\hat{\gamma}_u)^2$. The larger the difference between R^2 and $\hat{m}$, the more important the effect is of having obtained a small (or large) value of Z.

The following remark shows that a finite-variance property of $R^2(\eta)$ for $\eta > 0$, *viz.*, $R^2(1/\eta) = 1 - R^2(\eta)$, carries over in a natural way to $\widetilde{R}$.

Remark 2 *For $\eta > 0$, the distribution of $\widetilde{R}(\alpha, \eta)$ is skew-symmetric, viz.,*

$$\widetilde{R}(\alpha, \eta) \stackrel{d}{=} 1 - \widetilde{R}(\alpha, 1/\eta)\,,$$

or, equivalently, $\widetilde{R}(\alpha, m) \stackrel{d}{=} 1 - \widetilde{R}(\alpha, 1-m)$. The pdf of $\widetilde{R}$ therefore satisfies

$$f_{\widetilde{R}(\alpha,m)}(r) = f_{\widetilde{R}(\alpha,1-m)}(1-r) \quad \forall\, r \in [0,1]\,.$$

The distribution of $\widetilde{R}$ is symmetric about 0.5 *for $\eta = 1$.*

Proof. Recall that S_x and S_u are iid. Thus, for $\eta > 0$

$$\begin{aligned} 1 - \widetilde{R}(\alpha, 1/\eta) &= 1 - \frac{(1/\eta)S_x}{(1/\eta)S_x + S_u} \\ &= \frac{S_u}{(1/\eta)S_x + S_u} \\ &= \frac{\eta S_u}{\eta S_u + S_x} \\ &\stackrel{d}{=} \frac{\eta S_x}{\eta S_x + S_u} = \widetilde{R}(\alpha, \eta)\,. \end{aligned}$$

The symmetry of $\widetilde{R}$ about 0.5 for $\eta = 1$ follows immediately from this result and the fact that the distribution's support is the interval $[0,1]$. ■

Next, as the following remark shows, the pdf of $\widetilde{R}$ has *infinite* modes at 0 and 1, i.e., at the *endpoints* of its support.

Remark 3 *(i) For $\eta > 0$, the pdf of $\widetilde{R}$ is unbounded at 0 and 1, i.e., $f_{\widetilde{R}}(0) = f_{\widetilde{R}}(1) = \infty$. (ii) The cdf of $\widetilde{R}$ is continuous on $[0,1]$, and the distribution does not have atoms at 0 and 1.*

Proof. To demonstrate the validity of the first part of this remark, we apply a standard result for the pdf of the ratio of two random variables,[14] adapted to the present case where the random variables in the numerator and denominator are both strictly positive. For $\eta > 0$, set $V = \eta S_x$ and $W = \eta S_x + S_u$. We have

$$f_{\widetilde{R}}(r) = \int_0^\infty w f_{V,W}(rw, w)\,\mathrm{d}w, \qquad 0 \le r \le 1\,,$$

where the joint pdf $f_{V,W}(\cdot,\cdot)$ is nonzero on $\mathbb{R}^+ \times \mathbb{R}^+$. The case $r = 1$ can occur only if $S_u = 0$; if $S_u = 0$, however, the random variables V and W are perfectly dependent, their joint pdf is nonzero only on the positive 45°-halfline, and the joint pdf $f_{V,V}(w,w)$ reduces to $(1/\sqrt{2}) f_V(w)$, $w \ge 0$. Hence, for $r = 1$ we find

$$f_{\widetilde{R}}(1) = \int_0^\infty w f_{V,V}(1 \cdot w, w)\,\mathrm{d}w = \frac{1}{\sqrt{2}} \int_0^\infty w f_V(w)\,\mathrm{d}w = \frac{1}{\sqrt{2}}\,\mathrm{E}(\eta S_x) = \infty\,.$$

By Remark 2, we have $f_{\widetilde{R}}(0) = \infty$ as well.

The continuity of the cdf of $\widetilde{R}$ on $[0,1]$ for $\eta > 0$ follows from the continuity of the cdfs of S_x and S_u on $\mathbb{R}^+$ and the fact that their pdfs are equal to zero at the origin. For example, one finds that $\mathrm{P}(\widetilde{R} = 1) = \mathrm{P}(S_u = 0) = 0$; the result $\mathrm{P}(\widetilde{R} = 0) = 0$ then follows from Remark 2. ■

The fact that the probability density function of $\widetilde{R}$ has infinite singularities may seem unusual. However, the presence of singularities is a regular feature of pdfs that are based on *ratios* of stable random variables. For example, Logan et al. (1973) and Phillips and Hajivassiliou (1987) showed that if $\alpha < 1$, the density of the t-statistic has infinite modes at -1 and $+1$; similarly, Phillips and Loretan (1991) demonstrated that if $\alpha < 2$, this feature is also present in the asymptotic distributions of the von Neuman ratio and the normalized Durbin-Watson test statistic.

3.3 The cdf and pdf of $\widetilde{R}$

The remarks in the preceding subsection provide important qualitative information about some of the distributional properties of $\widetilde{R}$. However, they do not address issues such as whether the distribution has modes beyond those at 0 and 1, whether the discontinuity of the pdf at the endpoints is simple or if $f_{\widetilde{R}}(r)$ diverges—and, if so, at which rate—as $r \downarrow 0$ or $r \uparrow 1$, or how much of the distribution's mass is concentrated near the endpoints of the support. To examine these issues, we provide expressions for the cdf and pdf of $f_{\widetilde{R}}(r)$ in this subsection. It is possible to do so because $\widetilde{R}$ is a continuously differentiable and invertible function of

[14] See, e.g., Mood, Graybill, and Boes (1974), p. 187.

the *ratio* of two independent, maximally right-skewed, and positive α-stable random variables, and because closed-form expressions for the cdf and pdf of this ratio are known. The latter expressions are provided in the following proposition.

Proposition 1 (Zolotarev 1986, p. 205; Runde 1993, p. 11) *Let S_1 and S_2 be two iid positive α-stable random variables with common parameters $\alpha/2 \in (0,1)$, $\beta = +1$, $\gamma = 1$, and $\delta = 0$. Set $Z = S_1/S_2$. For $z \geq 0$, the cdf of Z is given by*

$$F_Z(z) = \mathrm{P}(Z \leq z) = \frac{1}{\pi\alpha/2} \arctan\left(\frac{z^{\alpha/2} + \cos(\pi\alpha/2)}{\sin(\pi\alpha/2)}\right) - \frac{1}{\alpha} + 1\,. \tag{13}$$

Differentiating this expression with respect to z, the pdf of Z for $z > 0$ is obtained as

$$f_Z(z) = \frac{\mathrm{d}}{\mathrm{d}z} F_Z(z) = \frac{\sin(\pi\alpha/2)}{\pi z\big[z^{-\alpha/2} + z^{\alpha/2} + 2\cos(\pi\alpha/2)\big]}\,. \tag{14}$$

As Z is a positive random variable, $F_Z(z) = f_Z(z) = 0$ for $z < 0$.

Figure 2 somewhere here

The cdf of the random variable Z is shown in Figure 2 for various values of α between 1.98 and 0.25.[15] The random variable Z has several interesting properties. First, note that $\lim_{z\downarrow 0} f_Z(z) = \infty$ and that the rate of divergence to infinity of $f_Z(z)$ as $z \downarrow 0$ is given by $(1/z)^{1-\alpha/2}$; thus, the pdf of Z has a one-sided infinite singularity at 0. Second, as $z \to \infty$, $f_Z(z) \approx \kappa \cdot z^{-\alpha/2-1}$ for a suitable constant $\kappa > 0$. This result, along with $\mathrm{P}(Z > 0) = 1$, implies that Z lies in the normal domain of attraction of a positive stable distribution, say Z', with index of stability $\alpha/2$ and $\beta = +1$, the same parameters as that of the variables S_1 and S_2.[16] Hence, the mean of Z is infinite for all values of $\alpha < 2$. Third, in the special case of $\alpha = 1$, S_1 and S_2 are each distributed as a Lévy α-stable random variable, which is well known to be equivalent to the inverse of a $\chi^2(1)$ random variable. For $\alpha = 1$, then, the pdf of Z reduces to $\big(\pi z^{1/2}(1+z)\big)^{-1}$, which is also the pdf of an $F_{1,1}$ distribution; see Runde (1993).

As was noted earlier, the median of Z is equal to 1 for all values of $\alpha \in (0, 2)$. The regression model's signal-to-noise ratio is given by the random variable ηZ if $\alpha < 2$, whereas it is given by the constant η in the standard, i.e., finite-variance case. The fact that the random variable which multiplies η has a

[15] Runde (1993) graphs pdfs of Z for values of α between 1.0 and 1.9.
[16] See Mittnik et al. (1998) for a discussion of some of the properties of the stable law Z'.

median of 1 helps to develop further the intuition that underlies the result of Remark 1, *viz.*, that the median of $\widetilde{R}$, $\eta/(\eta+1)$, is the same in both the finite-variance and the infinite-variance cases. Finally, an inspection of equation (13) reveals that $\lim_{\alpha\uparrow 2} \mathrm{P}(Z<1)=0$ and $\lim_{\alpha\uparrow 2} \mathrm{P}(Z>1)=0$; put differently, $\lim_{\alpha\uparrow 2} \mathrm{P}(Z=1)=1$. The probability mass of Z therefore becomes perfectly concentrated at 1 as $\alpha\uparrow 2$, even though, of course, its mean remains infinite as long as $\alpha<2$.

From Theorem 1, we have $\widetilde{R}=\eta Z/(\eta Z+1)=g(Z)$, say. Note that $Z\equiv S_x/S_u$ satisfies the conditions of Proposition 1 and that the function $Z=g^{-1}(\widetilde{R})=(1/\eta)\big(\widetilde{R}/(1-\widetilde{R})\big)$ is continuously differentiable and strictly increasing in the interior of its domain. We are therefore able to provide the following expressions for the cdf and pdf of $\widetilde{R}$ by an application of the density transformation theorem.[17]

Theorem 2 *For $r\in(0,1)$ and $\eta>0$, set $z=g^{-1}(r)=(1/\eta)\big(r/(1-r)\big)$, and let the cdf and pdf of Z be given by equations (13) and (14). The cdf of $\widetilde{R}$ for $r\in(0,1)$ is given by*

$$F_{\widetilde{R}}(r)=F_Z\left[g^{-1}(r)\right]. \tag{15}$$

Furthermore, $F_{\widetilde{R}}(0)=0$ and $F_{\widetilde{R}}(1)=1$.

The pdf of $\widetilde{R}$ for $r\in(0,1)$ is given by

$$\begin{aligned} f_{\widetilde{R}}(r) &= \left|\frac{\mathrm{d}}{\mathrm{d}r}g^{-1}(r)\right| f_Z\left[g^{-1}(r)\right] \\ &= \frac{1}{\eta(1-r)^2}\cdot\frac{\sin(\pi\alpha/2)}{\pi g^{-1}(r)\Big([g^{-1}(r)]^{-\alpha/2}+[g^{-1}(r)]^{\alpha/2}+2\cos(\pi\alpha/2)\Big)} \\ &= \frac{\sin(\pi\alpha/2)}{\pi\, r(1-r)}\cdot\left[z^{-\alpha/2}+z^{\alpha/2}+2\cos(\pi\alpha/2)\right]^{-1}, \quad \textit{where } z=r/\big(\eta(1-r)\big). \end{aligned} \tag{16}$$

As $r\downarrow 0$ or $r\uparrow 1$, $f_{\widetilde{R}}(r)$ diverges to infinity at a rate proportional to $(1/r)^{1-\alpha/2}$ and $\big(1/(1-r)\big)^{1-\alpha/2}$, respectively.

Proof. The results stated in equations (15) and (16) follow immediately from Proposition 1 and the density transformation theorem. Because $\lim_{r\downarrow 0} \mathrm{d}g^{-1}(r)/\mathrm{d}r=\eta^{-1}$, the rate of divergence of $f_{\widetilde{R}}(r)$ as $r\downarrow 0$ is equal to—apart from the multiplicative constant η^{-1}—that of $f_Z(z)$ as $z\downarrow 0$, which is $(1/z)^{1-\alpha/2}$. Finally, it follows from Remark 2 that as $r\uparrow 1$ the pdf of $\widetilde{R}$ also diverges to infinity at this rate. ■

The probability density functions and cumulative distribution functions of $\widetilde{R}(\alpha,\eta)$ for values of α between 0.25 and 1.98 are graphed in Figures 3 and 4. (In all cases, we have set $\eta=1$.) The pdfs in Figure 3

[17] See, e.g., Mood, Graybill, and Boes (1974, p. 200).

are shown with a logarithmic scale on the ordinate. Since we know that $f_{\widetilde{R}}(0) = f_{\widetilde{R}}(1) = \infty$, we graph the functions only for $r \in (10^{-13}, 1 - 10^{-13})$. The graphs show that

- If α is close to but less than 2, e.g., if $\alpha = 1.98$ or $\alpha = 1.90$, the pdf has an interior mode, and most of the probability mass of $\widetilde{R}$ is concentrated near its median. Conversely, only very little mass is located near 0 and 1, and the pdfs register only mild increases as r approaches either edge of the distribution's support.

- For $\alpha = 1.75$ and $\alpha = 1.50$, the distribution of $\widetilde{R}$ continues to have an interior mode (as well as, of course, the two unbounded modes at 0 and 1). However, the distribution is noticeably less concentrated around the interior mode than if α is closer to 2.

- By $\alpha = 1.20$, the interior mode has disappeared and the distribution is nearly uniform over the entire interval $[0, 1]$.

- If α takes on even smaller values, less and less of the probability mass of $\widetilde{R}$ is located near the median, and more and more of it is concentrated close to 0 and 1.

- If $\alpha = 0.25$, about 75 percent of the probability mass lies within 0.001 of the two endpoints of the distribution, while the probability of observing a realization of $\widetilde{R}$ for $r \in [0.25, 0.75]$ is less than 5 percent.

Figures 3 and 4 somewhere here

A heuristic summary of these properties of $\widetilde{R}$ is straightforward. We begin by recalling that the multiplicative term $C(\alpha)$, shown in equation (4) and Figure 1, affects the probability of tail-region values of the random variables in question, and that the rate of decline in the tail areas of density of α-stable random variables increases as $\alpha \uparrow 2$. Suppose first that α is very close to 2; then, $C(\alpha)$ is close to 0, and the fraction of observations of x_t and u_t that fall into the respective Paretian-tail regions is therefore very low; moreover, given the fairly rapid decay of the density's tails for α close to 2, the likelihood of obtaining a very large draw, conditional on obtaining a draw from the Paretian tail area, is also low. As a result, the probability of observing large observations of x_t and u_t is quite low. This, in turn, makes it unlikely to observe a very large draw of either S_x or S_u and thus of observing a value of Z that is either close to 0 or very large. Therefore, if α is very close to 2, Z is likely close to its median of 1, and most of the mass of $\widetilde{R}$ is concentrated near its median, $\eta/(\eta+1)$. Next, as α moves down and away from 2, say to around 1.5, $C(\alpha)$ increases rapidly, leading to a higher frequency of observing tail-region draws for x_t and u_t. In addition, as the density in the tail region declines more slowly for smaller values of α, it is much more likely of obtaining very large draws of the regressor and error term than if α is close to 2. In consequence, if α is around 1.5, it is quite likely to obtain draws of Z that are either very close to zero or very large, and thus more of the probability mass of $\widetilde{R}$ is located near the edges of its support. Conversely the interior mode of $\widetilde{R}$ is considerably less pronounced than if α is close to 2. Finally, as α decreases further, $C(\alpha)$ rises further, and both the frequency of tail observations and the likelihood that any draws from the tail areas will be very large increase. Therefore, it is very likely that the largest few observations of x_t or u_t will dominate the realization of Z and therefore the realization of $\widetilde{R}$. As a result, if α is small the central mode of $\widetilde{R}$ vanishes entirely and almost all of its probability mass is located very close to the endpoints of the distribution's support. In the limit, as $\alpha \downarrow 0$, $\widetilde{R}$ converges to a Bernoulli random variable, for which all of the probability mass is located at 0 and 1.

4 An empirical application

Fama and MacBeth (1973) proposed the so-called Fama-MacBeth regression to test the hypothesis of a linear relationship between risk and risk premium in stock returns in a cross-sectional setting. Let r_{it} be the return on market portfolio i at time t, where $i = 1, \dots, N$ and $t = 1, \dots, T$; denote the average return of portfolio i as $\bar{r}_i = T^{-1} \sum_{t=1}^{T} r_{it}$; denote the average portfolio return at time t as $R_t = N^{-1} \sum_{i=1}^{N} r_{it}$; and denote the average portfolio return across all time periods by $\mu_R = T^{-1} \sum_{t=1}^{T} R_t$. The first-stage Fama-MacBeth regression is an *ex post* CAPM,

$$r_{it} = \theta_{0i} + \theta_i R_t + u_t, \quad t = 1, \dots, T, \tag{17}$$

where $\mathrm{E}(u_t) = 0$, $\mathrm{E}(u_t R_t) = 0$, and u_t is iid $S\alpha S$ with the same index $\alpha \in (1, 2]$ as r_{it}. We may assume that the distribution of θ_i has a finite mean and variance, say, $\mathrm{E}(\theta)$ and $\mathrm{Var}(\theta)$. Denote the OLS estimates of the regression coefficients in equation (17) by $\hat{\theta}_{0i}$ and $\hat{\theta}_i$. The second-stage Fama-MacBeth regression is given by

$$\bar{r}_i = \lambda_0 + \lambda_1 \hat{\theta}_i + \varepsilon_i, \qquad i = 1, \dots, N\,, \tag{18}$$

where ε_i is iid $S\alpha S$ with the same index α as r_{it}, $\mathrm{E}(\varepsilon_i) = 0$, and $\mathrm{E}(\varepsilon_i \hat{\theta}_i) = 0$.

The R^2 statistic of the second-stage Fama-MacBeth regression is given by

$$R^2 = \frac{N^{-1}\hat{\lambda}_1^2 \sum_{i=1}^N \big(\hat{\theta}_i - \bar{\hat{\theta}}_i\big)^2}{N^{-1}\hat{\lambda}_1^2 \sum_{i=1}^N \big(\hat{\theta}_i - \bar{\hat{\theta}}_i\big)^2 + N^{-1}\sum_{i=1}^N \hat{\varepsilon}_i^2}\,. \tag{19}$$

This statistic has the following asymptotic properties.

Theorem 3 *If the individual portfolio returns r_{it} follow an iid $S\alpha S$ distribution with $\alpha \in (1,2]$ and if $\mu_R > 0$, the coefficient of determination in* (19) *has the following limits as $T \to \infty$ and $N \to \infty$:*

- *If $\alpha = 2$, $R^2 \to_p \eta/(\eta+1)$, where $\eta = \lambda_1^2 \mathrm{Var}(\theta)/\mathrm{Var}(\varepsilon)$; and*
- *If $\alpha < 2$, $R^2 = o_p\big(N^{1-2/\alpha}\big)$.*

Thus, if $\alpha < 2$, $R^2 \to_p 0$, at a rate that is proportional to $N^{1-2/\alpha}$.

Proof. The result for the finite-variance case follows immediately from Corollary 1. For $\alpha < 2$, observe that the normalized estimator of θ_i, $T^{(\alpha-1)/\alpha}\big(\hat{\theta}_i - \mathrm{E}(\theta)\big)$, is in the domain of attraction of an α-stable distribution for *fixed* values of T. As $T \to \infty$, the dispersion of $\hat{\theta}_i$ about $\mathrm{E}(\theta)$ converges to 0, and the distributional properties of the estimated regressors $\hat{\theta}_i$ converge to those of θ_i; by assumption, the variance of θ_i is finite. Thus, as $N \to \infty$ and $T \to \infty$, the numerator in equation (19) converges to $\lambda_1^2\mathrm{Var}(\theta)$. In contrast, the second summand in the denominator of (19) requires norming by $N^{2/\alpha} > N$ in order to attain a proper limit. The coefficient of determination therefore converges to 0 in probability as $N \to \infty$ and $T \to \infty$, at a rate of $N^{1-2/\alpha}$. ■

This result does not conflict with the one provided in Theorem 1, as the present case is one of an unbalanced regression design: the regressor has an asymptotically finite variance, whereas the error term has infinite variance, implying that the asymptotic signal-to-noise ratio is zero. Instead, this result is closely related to the one provided in Corollary 2, which examined the asymptotic limit of R^2 if $\alpha_x \neq \alpha_u$. We note that even if T is fixed (as is generally taken to be the case in Fama-MacBeth regressions), the dispersion of $\hat{\theta}_i$ will likely be quite a bit smaller than that of ε_i, indicating that the model's signal-to-noise ratio, η, and hence the median of R^2, in the second-stage regression will be quite small unless λ_1 is sufficiently large in absolute value.

These qualitative observations are confirmed by a small-scale Monte Carlo simulation, shown in Table 1, in which we report the median value of R^2 as a function of two values of α and selected values of T, N, and μ_R.[18] It is evident for both $\alpha = 1.5$ and $\alpha = 1.75$ that the median value of R^2 declines as N increases

[18] The design of the simulation and the choices of values for α, T, N and μ_R were influenced by a desire to maximize the empirical relevance of the simulation exercise. We chose $\alpha = 1.5$ and $\alpha = 1.75$ because $\hat{\alpha} \geq 1.5$ for most empirical economic

if T is fixed, that this effect is particularly strong if T is large, and that this effect is more pronounced for $\alpha = 1.5$ than it is for $\alpha = 1.75$. The final result is as one would expect, given that Theorem 3 states that the rate of convergence of R^2 to zero increases as α moves down further from 2.

Table 1 somewhere here

On the basis of the small value of coefficient of determination from the Fama-MacBeth regression, Jagannathan and Wang (1996) confirm the finding of Fama and Macbeth (1973) of a "flat" relation between average return and market beta. They report a very low coefficient of determination of 1.35%=0.0135 for the Sharpe-Lintner-Black (SLB) static CAPM. Regarding "thick-tailed" phenomena in empirical data, Fama and French (1992) conjectured that neglecting the heavy-tails phenomenon of the data does not lead to serious errors in the interpretation of empirical results. In the following, we use the same CRSP dataset as was used by Jagannathan and Wang (1996); the data are very similar to those that were used in the study of Fama and French (1992). The data consist of stock returns of nonfinancial firms listed on the NYSE and AMEX from July 1963 until December 1990 covered by CRSP alone; the frequency of observation is monthly. In the preceding notation, we have $T = 330$ and $N = 100$. Figure 5 displays the time series of these monthly returns.

Figure 5 somewhere here

For our analysis we need to obtain point estimates the index of stability of the stock returns and determine whether the estimates are less than 2. Under the assumption of symmetry, which implies that the left and right tails of the returns distribution possess the same maximal moment exponent and dispersion coefficient, the point estimate of α for monthly stock returns in the CRSP dataset using the Hill method (Hill, 1975)

data. We study the cases of $T = 100$, 250, 1,000, and 2,500 because $T = 250$ corresponds approximately to the number of business days in a calendar year. The values of $N = 30$, 100, 500, and 1,000 correspond to the numbers of stocks contained in certain well-known stock price indices, such as the U.S. Dow-Jones Industial and German DAX indices, the U.K. FTSE-100 index, the U.S. S&P-500 index, etc. The choice of $\mu_R = 0$ provides a reference to contrast the cases of $\theta_i = 0$ and $\theta_i \neq 0$; $\mu_R = 0.1$ is particularly relevant for the empirical study provided below.

is 1.77, with a standard deviation of 0.15.[19] On the basis of these estimates, normality ($\alpha = 2$) can be excluded only at a confidence level of approximately 87.5 percent. However, inference about the width of the confidence interval for the Hill estimator is valid only asymptotically; in finite samples, the Hill-method estimates are known to be quite sensitive to even minor departures from exactly Paretian tail behavior.[20] In contrast, the method of Dufour and Kurz-Kim (2007) provides exact confidence intervals for finite samples. By their method, the point estimate of α for the monthly stock returns data is 1.78, and the exact finite-sample 90% confidence interval for this point estimate is [1.64, 1.99]. This result also does not offer very strong evidence against the hypothesis $\alpha = 2$. Nevertheless, because of estimation uncertainty in small samples, and because this uncertainty is especially severe if α is close to 2, the data can be regarded as being in the domain of attraction of a stable distribution with $\alpha < 2$.[21] We therefore proceed to investigate the consequences of this finding for the proper interpretation of the low R^2 statistic reported by Jegannathan and Wang (1996).

We designed Monte Carlo simulations to obtain the cdf of R^2 for our empirical data, first under the assumption that the returns data are in the domain of attraction of an α-stable distribution with $\alpha < 2$, and second under the assumption of normality ($\alpha = 2$). The simulation was calibrated to the main characteristics the empirical data; we set $\alpha = 1.78$, $T = 330$, $N = 100$, and we set the expected return equal to the average annual return in the full sample, i.e., $\mu_R = 0.1088$. The number of replications of the first-stage and second-stage Fama-MacBeth regressions is 100,000, for the both values of α. The simulated cdfs of the R^2-statistic are shown in Figure 6, where a vertical line is drawn at $R^2 = 0.0135$ to indicate the in-sample value of the coefficient of determination. The shapes of the two curves are rather different, with the one for $\alpha = 1.78$ rising much more quickly for small values of R^2.

Figure 6 somewhere here

[19] In this estimation, we used 0.0031 as the centering offset for the empirical data; this adjustment is necessary because the Hill estimator is not location-invariant. The offset is equal to the estimated location parameter obtained by the quantile estimation method of McCulloch (1986). The choice of the number of order statistics to include in the Hill method used was determined by the Monte Carlo method of Dufour and Kurz-Kim (2007). For the present dataset, this method indicated the use of 43% of all observations.

The Hill estimator uses extreme observations from both tails of the empirical distribution under the assumption of symmetry, but it uses only observations from the right (left) tail under the assumption of right-skewed (left-skewed) asymmetry. In the case of the monthly stock returns, the distribution is clearly left-skewed, i.e., the largest negative returns are larger in the sample than the largest positive returns; see Figure 5. Under the assumption of left-skewed asymmetry, the point estimate of α for the left tail using the Hill method is 1.47, with one standard deviation of 0.18.

[20] Stable distributions have tails that are *asymptotically* Paretian. In finite samples, and especially if the index of stability is not far below 2, it is known that the tails of stable distributions are not approximated particularly well by Pareto distributions with the same value of *alpha*. See Resnick (2006, pp. 86–9) for a discussion of the consequences of these finite-sample features for the reliability of the Hill estimator.

[21] For a broader discussion of how to decide if $\alpha < 2$, see McCulloch (1997).

The simulated median R^2 of the second-stage Fama-MacBeth regression is 0.384 for $\alpha = 2$, but it is only 0.072 for $\alpha = 1.78$. The simulated probability of obtaining $R^2 \leq 0.0135$ is a minuscule 1.55 percent for $\alpha = 2$, but it is a much more sizable 21.88 percent for $\alpha = 1.78$; thus, if $\alpha = 1.78$ the event $R^2 \leq 0.0135$ is about 14 times more probable than if $\alpha = 2$. On the basis of these findings, we conclude that the inference drawn from the low value of R^2 by Fama and French (1992)—that the empirical usefulness of the SLB CAPM is refuted—does not seem to be robust once proper allowance is made for the distributional properties of the data that give rise to this statistic.

5 Concluding remarks

After providing a brief overview of some of the properties of α-stable distributions, this paper surveys the literature on the estimation of linear regression models with infinite-variance variables and associated methods of conducting hypothesis and specification tests. Our paper adds to the already-wide body of knowledge that there are substantial differences between regression models with infinite-variance and finite-variance regressors and error terms by examining the properties of the coefficient of determination. In the infinite-variance case with iid regressors and error terms that share the same index of stability α, we find that the R^2 statistic does not converge to a constant but instead that it has a nondegenerate asymptotic distribution on the $[0, 1]$ interval, with a pdf that has infinite singularities at 0 and 1. We provide closed-form expressions for the cdf and pdf of this limit random variable. If the regressors and error term do not have the same index of stability, we show that the coefficient of determination collapses either to 0 or to 1, depending on whether the model's signal-to-noise ratio converges asymptotically to zero or infinity. Finally, we provide an empirical application of our methods to the Fama-MacBeth two-stage regression setup, and we show that the coefficient of determination asymptotically converges to 0 in probability if the regression variables have infinite variance. This, in turn, strongly suggests that low values of the R^2 statistic should not, by themselves, be taken as proof of a "flat" relationship between the dependent variable and the regressor.

In view of the random nature of the limit law $\widetilde{R}$ if the regressors and error terms share the same index of stability, and given our related finding that the coefficient of determination converges to zero in probability if the tail index of the disturbance term is smaller than that of the regressor, a case that may be difficult to rule out in empirical practice unless the sample size is very large, we view our results as establishing that one should *not* rely on R^2 as a measure of the goodness of fit of a regression model whenever the regressors and disturbance terms are sufficiently heavy-tailed to call into question the existence of second (population) moments. At the very least, if one chooses to report the coefficient of determination in regressions with infinite-variance variables at all, one should also report a point estimate of the median of $\widetilde{R}$, $\hat{m} = \hat{\eta}/(\hat{\eta} + 1)$, where η is as in Theorem 1. In addition, one should indicate whether the error terms and regressors may reasonably be assumed to share the same index of stability. If the validity of that assumption is in doubt, the authors should also indicate which of the two parameters is likely to be smaller and how far apart the two parameters may plausibly be.

It is widely known, and it is certainly stressed in all introductory econometrics textbooks, that a *high* value of R^2 does not provide a sufficient basis for concluding that an empirical regression model is a "good" explanation of the dependent variable, or even that the regression is correctly specified. Nevertheless, one suspects, researchers may view *low* values of R^2 in an empirical regression as an indication that the (linear) relationship is either weak or unreliable. A direct implication of the work presented in this paper is that whenever the data are characterized by significant outlier activity, a low value of R^2 should not, by itself, be used to disqualify the model from further consideration.

Several extensions to the work presented here are possible. First, the regression F-statistic is a simple function of the coefficient of determination; e.g., $F = (T-2) \cdot R^2/(1-R^2)$ in the bivariate regression case. Given the close connection between the two statistics, it seems useful to study if and how the distributional properties of the regression F-statistic are affected by the presence of α-stable regressors and error terms under both the null hypothesis, $\theta = 0$, and the alternative hypothesis, $\theta \neq 0$. It would also be useful to elaborate on our idea, offered after Remark 1 in subsection 3.2, that the difference between the estimate of R^2 and a consistent estimate of its median may serve as a diagnostic check of the size of the effect of infinite variance on R^2. For example, it may be feasible to develop an asymptotic theory of the distributional properties of this difference.

It also seems desirable to study how well the distribution of $\widetilde{R}$ approximates the empirical distribution of R^2 in finite samples, for various types of heavy-tailed distributions that are in the domain of attraction of $S\alpha S$ distributions, and for various types of estimators (such as OLS, Blattberg-Sargent's BLUE, and the least-absolute deviation estimator). In addition, an extension to a multiple-regression framework may produce additional insights into the properties of the coefficient of determination in the infinite-variance case. Finally, the theoretical results presented in our paper depend crucially on the assumption that the random variables are iid. Relaxing this assumption would seem to be useful, as many economic and financial time series—especially if they are sampled at very high frequencies—are characterized by interesting dependence and heterogeneity features. Introducing serial dependence and heterogeneity, especially conditional heterogeneity, would serve the purpose of studying how the properties of $\widetilde{R}$ may be affected by such departures from the basic case of iid variables. The authors are considering conducting research to extend the work presented in this paper along these lines.

References

Blattberg, R. and T. Sargent (1971): Regression with non-Gaussian stable disturbances: Some sampling results, *Econometrica* 39, 501–510.

Brockwell, P.J. and R.A. Davis (1991): *Time Series: Theory and Models* (2nd ed). Springer: New York.

Davis, R.A. and S.I. Resnick (1985a): Limit theory for moving averages of random variables with regularly varying tail probabilities, *Annals of Probability* 13, 179–195.

Davis, R.A. and S.I. Resnick (1985b): More limit theory for sample correlation functions of moving averages, *Stochastic Processes and Their Applications* 20, 257–279.

Davis, R.A. and S.I. Resnick (1986): Limit theory for sample covariance and correlation functions of moving averages, *Annals of Statistics* 14, 533–558.

De Vries, C. (1991): On the relation between GARCH and stable processes, *Journal of Econometrics* 48, 313–324.

Dufour, J.-M. and J.-R. Kurz-Kim (2007): Exact inference and optimal invariant estimation for the stability parameter of symmetric α-stable distributions, *Journal of Empirical Finance*, forthcoming.

Fama, E.F. and K.R. French (1992): The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, E.F. and J.D. MacBeth (1973): Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607–636.

Feller, W. (1971): *An Introduction to Probability Theory and its Applications, Vol. 2* (2nd ed). Wiley: New York.

Gnedenko, B.V. and A.N. Kolmogorov (1954): *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley: Reading, Mass.

Granger, C.W.J. and D. Orr (1972): "Infinite variance" and research strategy in time series analysis, *Journal of the American Statistical Association* 67, 275–285.

Ibragimov, I.A. and Yu.V. Linnik (1971): *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff: Groningen.

Hill, B.M. (1975): A simple general approach to inference about the tail of a distribution, *Annals of Statistics* 3, 1163–1174.

Jagannathan, R. und Z. Wang (1996): The conditional CAPM and cross-section of expected returns, *Journal of Finance* 51, 3–53.

Kanter, M. and W.L. Steiger (1974): Regression and autoregression with infinite variance, *Advances in Applied Probability* 6, 768–783.

Kim, J.-R. (2003): Finite-sample distributions of self-normalized sums, *Computational Statistics* 18, 493–504.

Kim, J.-R. and S.T. Rachev (1999): Asymptotic distributions of BLUE in the presence of infinite variance residuals, Unpublished manuscript.

Kurz-Kim, J.-R., S.T. Rachev and G. Samorodnitsky (2005): F-type Distributions for Heavy-tailed Variates, Unpublished manuscript.

Logan, B.F., C.L. Mallows, S.O. Rice and L.A. Shepp (1973): Limit distributions of self-normalized sums, *Annals of Probability* 1, 788–809.

Loretan, M. and P.C.B. Phillips (1994): Testing the covariance stationarity of heavy-tailed time series, *Journal of Empirical Finance* 1, 211–248.

Mood, A.M., F.A. Graybill and D.C. Boes (1974): *Introduction to the Theory of Statistics* (3rd ed). McGraw-Hill: New York.

Mandelbrot, B. (1963): The variation of certain speculative prices, *Journal of Business* 36, 394–419.

McCulloch, J.H. (1986): Simple consistent estimators of stable distribution parameters, *Communications in Statistics, Computation and Simulation* 15, 1109–1136.

McCulloch, J.H. (1996): Financial applications of stable distributions, in: *Handbook of Statistics—Statistical Methods in Finance, Vol. 14*, 393–425, eds. G.S. Maddala and C.R. Rao. Elsevier Science: Amsterdam.

McCulloch, J.H. (1997): Measuring tail thickness to estimate the stable index α: A critique, *Journal of Business and Economic Statistics* 15, 74–81.

McCulloch, J.H. (1998): Linear regression with stable disturbances, in: *A Practical Guide to Heavy Tails*, 359–376, eds. R. Adler, R. Feldman, and M.S. Taqqu. Birkhäuser: Berlin and Boston.

Mittnik, S., S.T. Rachev and J.-R. Kim (1998): Chi-square-type distributions for heavy-tailed variates, *Econometric Theory* 14, 339–354.

Phillips, P.C.B. (1990): Time series regression with a unit root and infinite-variance errors, *Econometric Theory* 6, 44–62.

Phillips, P.C.B. and V.A. Hajivassiliou (1987): Bimodal t-ratios, *Cowles Foundation Discussion Paper* No. 842.

Phillips, P.C.B. and M. Loretan (1991): The Durbin-Watson ratio under infinite-variance errors, *Journal of Econometrics* 47, 85–114.

Phillips, P.C.B. and M. Loretan (1994): On the theory of testing covariance stationarity under moment condition failure, in: *Advances in Econometrics and Quantiative Economics. Essays in Honor of Professor C.R. Rao*, 198–233, eds. G.S. Maddala, P.C.B. Phillips, and T.N. Srinivasan. Blackwell: Oxford.

Rachev, S.T., J.-R. Kim and S. Mittnik (1999): Stable Paretian econometrics, Part I and II, *The Mathematical Scientists* 24, 24–55 and 113–127.

Rachev, S.T., C. Menn and F.J. Fabozzi (2005): *Fat-Tailed and Skewed Asset Return Distributions*. Wiley: New York.

Resnick, S.I. (1986): Point processes regular variation and weak convergence, *Advances in Applied Probability* 18, 66–138.

Resnick, S.I. (1999): *A Probability Path*. Birkhäuser: Berlin and Boston.

Resnick, S.I. (2006), *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer: New York.

Runde, R. (1993): A note on the asymptotic distribution of the F-statistic for random variables with infinite variance, *Statistics & Probability Letters* 18, 9–12.

Runde, R. (1997): The asymptotic null distribution of the Box-Pierce Q-statistic for random variables with infinite variance, with an application to German stock returns, *Journal of Econometrics* 78, 205–216.

Samorodnitsky, G., S.T. Rachev, J.-R. Kurz-Kim and S. Stoyanov (2007): Asymptotic distribution of linear unbiased estimators in the presence of heavy-tailed stochastic regressors and residuals, *Probability and Mathematical Statistics*, forthcoming.

Samorodnitsky, G. and M.S. Taqqu (1994): *Stable Non-Gaussian Random Processes*. Chapman & Hall: New York.

Zolotarev, V.M. (1986): *One-Dimensional Stable Distributions*, Translations of Mathematical Monographs, Vol. 65. American Mathematical Society: Providence.

Table 1: Median value of R^2 as a function of α, T, N, and μ_R

α	T	N	μ_R				
			0.0	0.1	0.3	0.5	1.0
1.50	100	30	0.0404	0.0425	0.0576	0.1009	0.2963
		100	0.0162	0.0172	0.0292	0.0596	0.2019
		500	0.0068	0.0075	0.0147	0.0325	0.1206
		1000	0.0046	0.0055	0.0114	0.0264	0.0973
	250	30	0.0402	0.0426	0.0779	0.1598	0.4417
		100	0.0161	0.0190	0.0448	0.1058	0.3304
		500	0.0064	0.0075	0.0220	0.0565	0.2020
		1000	0.0047	0.0058	0.0172	0.0452	0.1667
	1000	30	0.0387	0.0484	0.1499	0.3320	0.6748
		100	0.0162	0.0223	0.0940	0.2272	0.5558
		500	0.0065	0.0104	0.0521	0.1341	0.3994
		1000	0.0046	0.0072	0.0399	0.1079	0.3443
	2500	30	0.0403	0.0580	0.2478	0.4806	0.7962
		100	0.0155	0.0294	0.1621	0.3581	0.6973
		500	0.0066	0.0130	0.0883	0.2243	0.5507
		1000	0.0047	0.0103	0.0737	0.1896	0.4970
1.75	100	30	0.0488	0.0543	0.1332	0.2944	0.6410
		100	0.0260	0.0328	0.1032	0.2413	0.5756
		500	0.0177	0.0222	0.0779	0.1941	0.5055
		1000	0.0149	0.0199	0.0720	0.1778	0.4792
	250	30	0.0474	0.0642	0.2509	0.4899	0.7993
		100	0.0265	0.0430	0.2066	0.4264	0.7560
		500	0.0169	0.0290	0.1571	0.3500	0.6950
		1000	0.0143	0.0251	0.1440	0.3273	0.6730
	1000	30	0.0470	0.1193	0.5351	0.7665	0.9309
		100	0.0265	0.0871	0.4612	0.7124	0.9115
		500	0.0169	0.0635	0.3910	0.6507	0.8865
		1000	0.0144	0.0579	0.3663	0.6257	0.8744
	2500	30	0.0480	0.2185	0.7214	0.8804	0.9677
		100	0.0255	0.1704	0.6599	0.8474	0.9578
		500	0.0169	0.1251	0.5900	0.8066	0.9452
		1000	0.0149	0.1202	0.5674	0.7902	0.9394

The numbers in the body of the table are the medians from simulated distributions with 100,000 replications.

Figure 1. The function $C(\alpha),\ 0<\alpha<2$

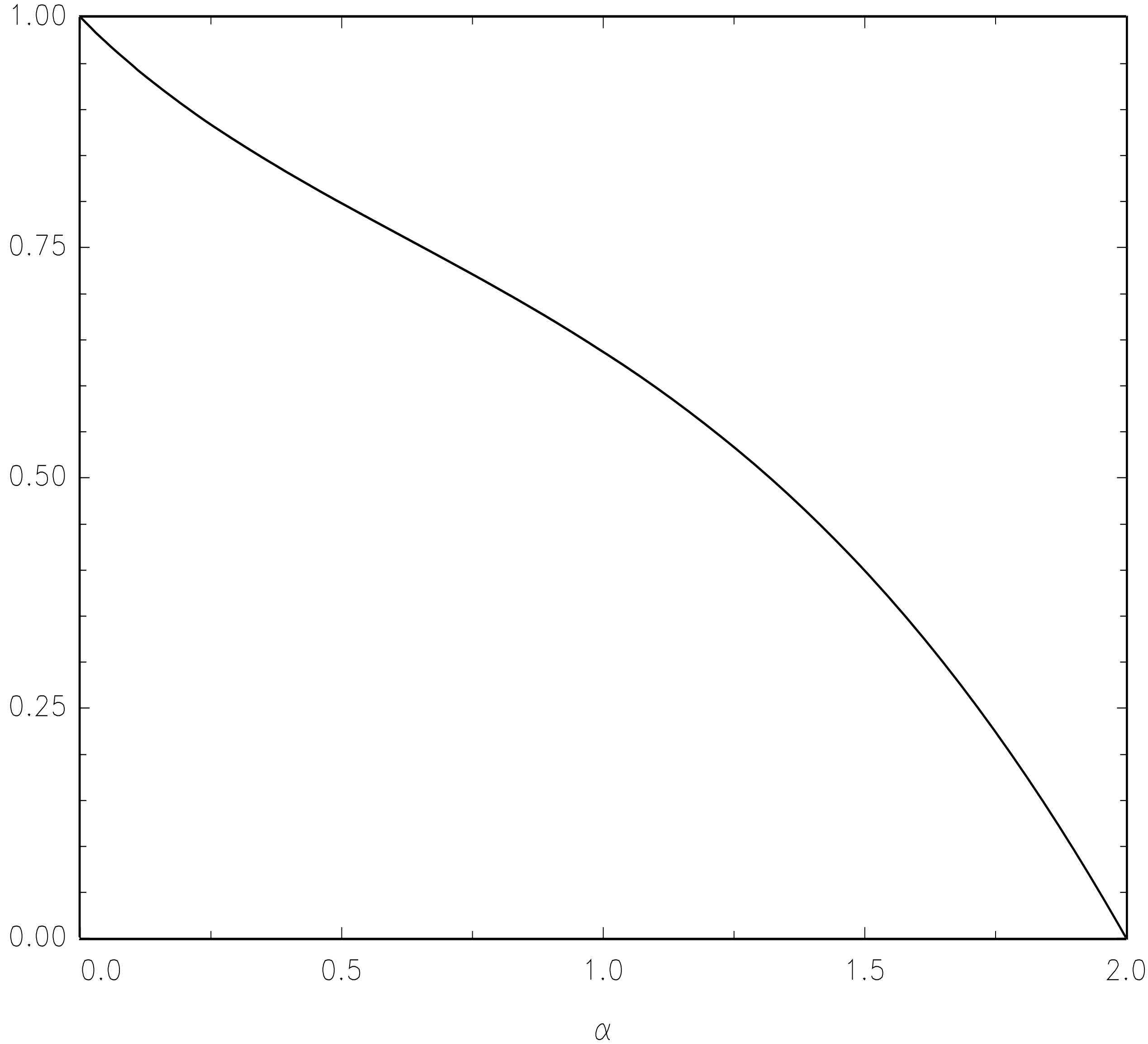

Figure 2. Cumulative distribution functions of $Z \equiv S_x / S_u$

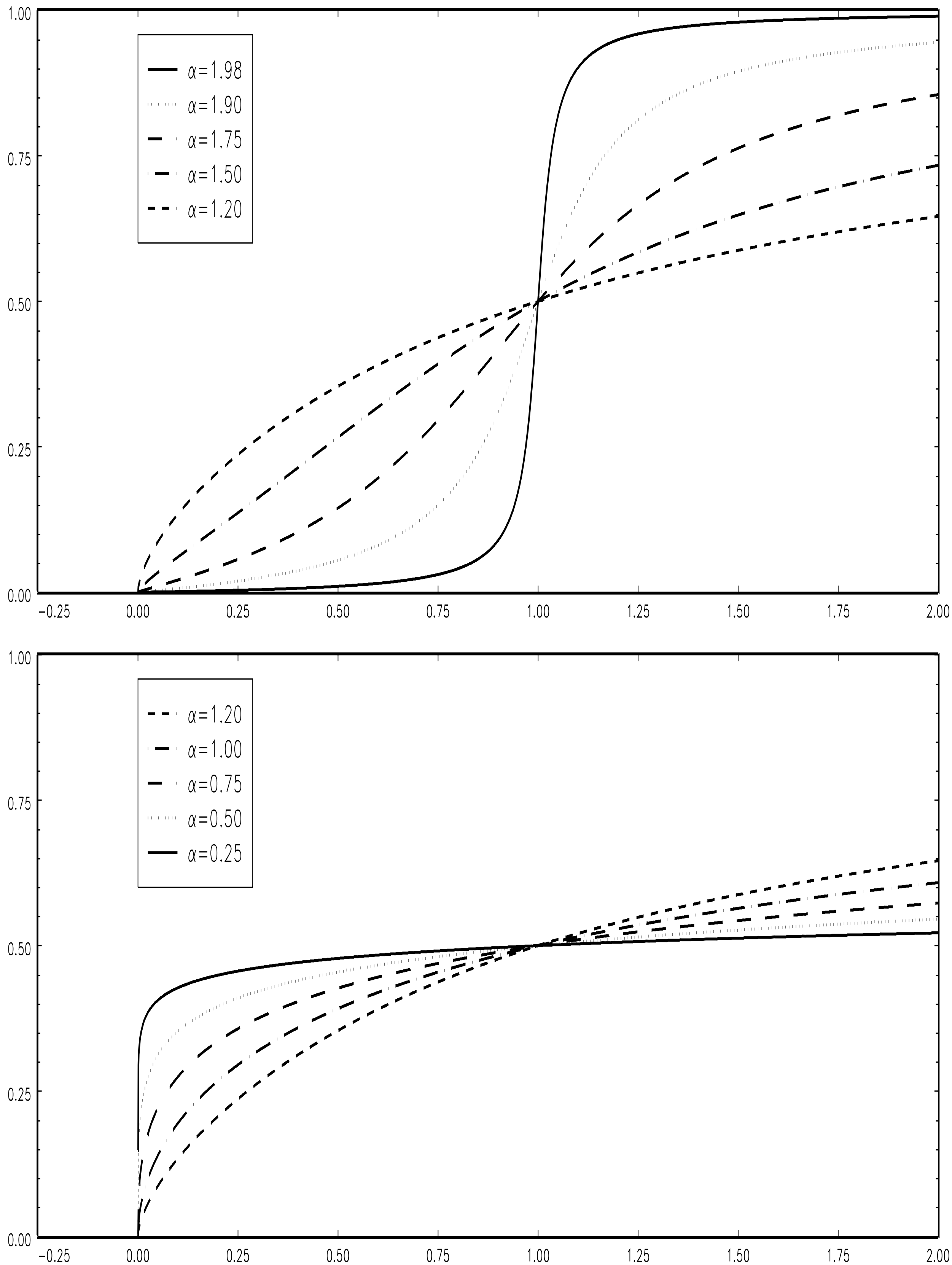

Figure 3. Probability density functions of $\tilde{R}(\alpha,\eta)$, $\eta = 1$

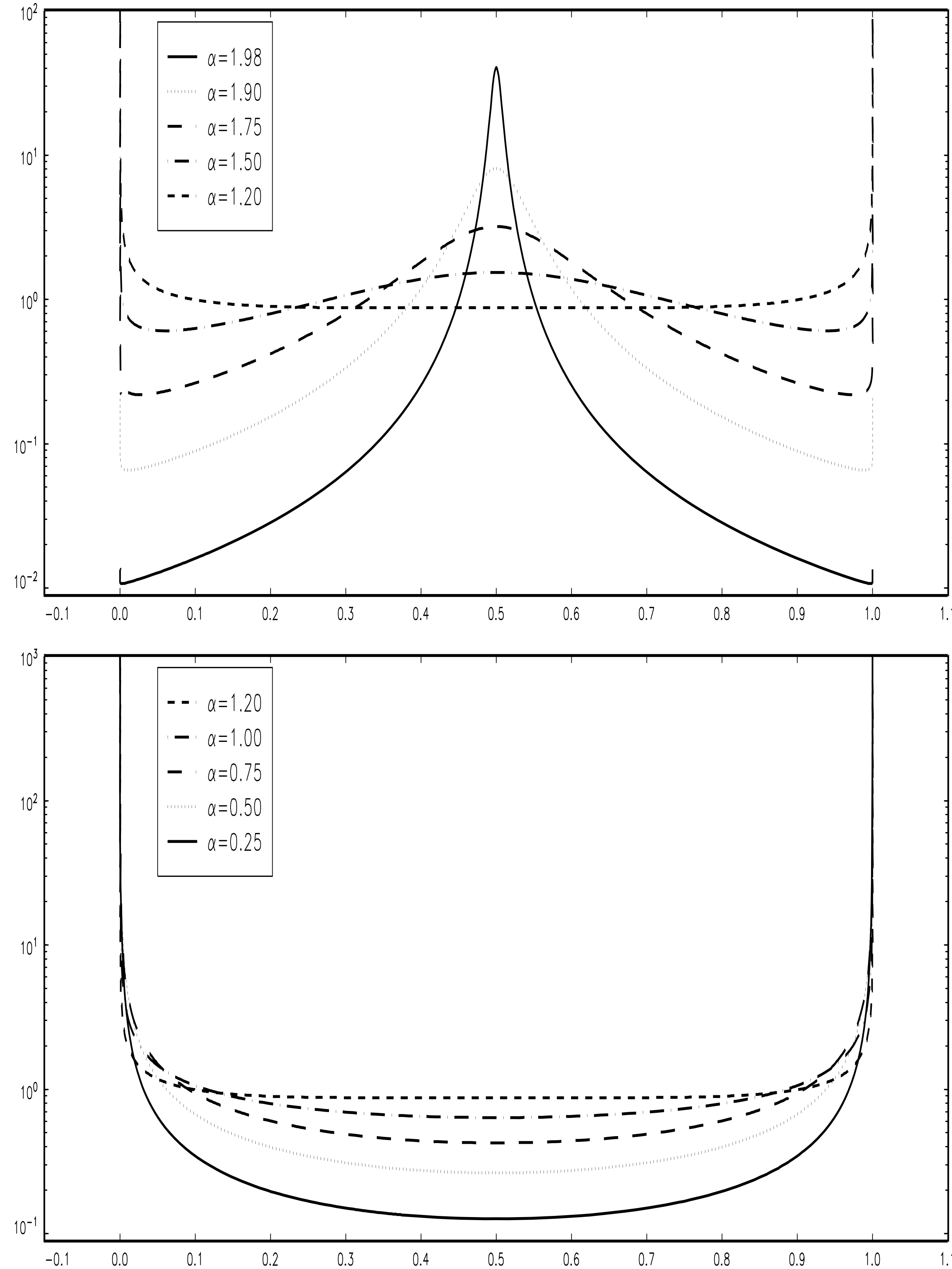

Figure 4. Cumulative distribution functions of $\tilde{R}(\alpha, \eta)$, $\eta = 1$

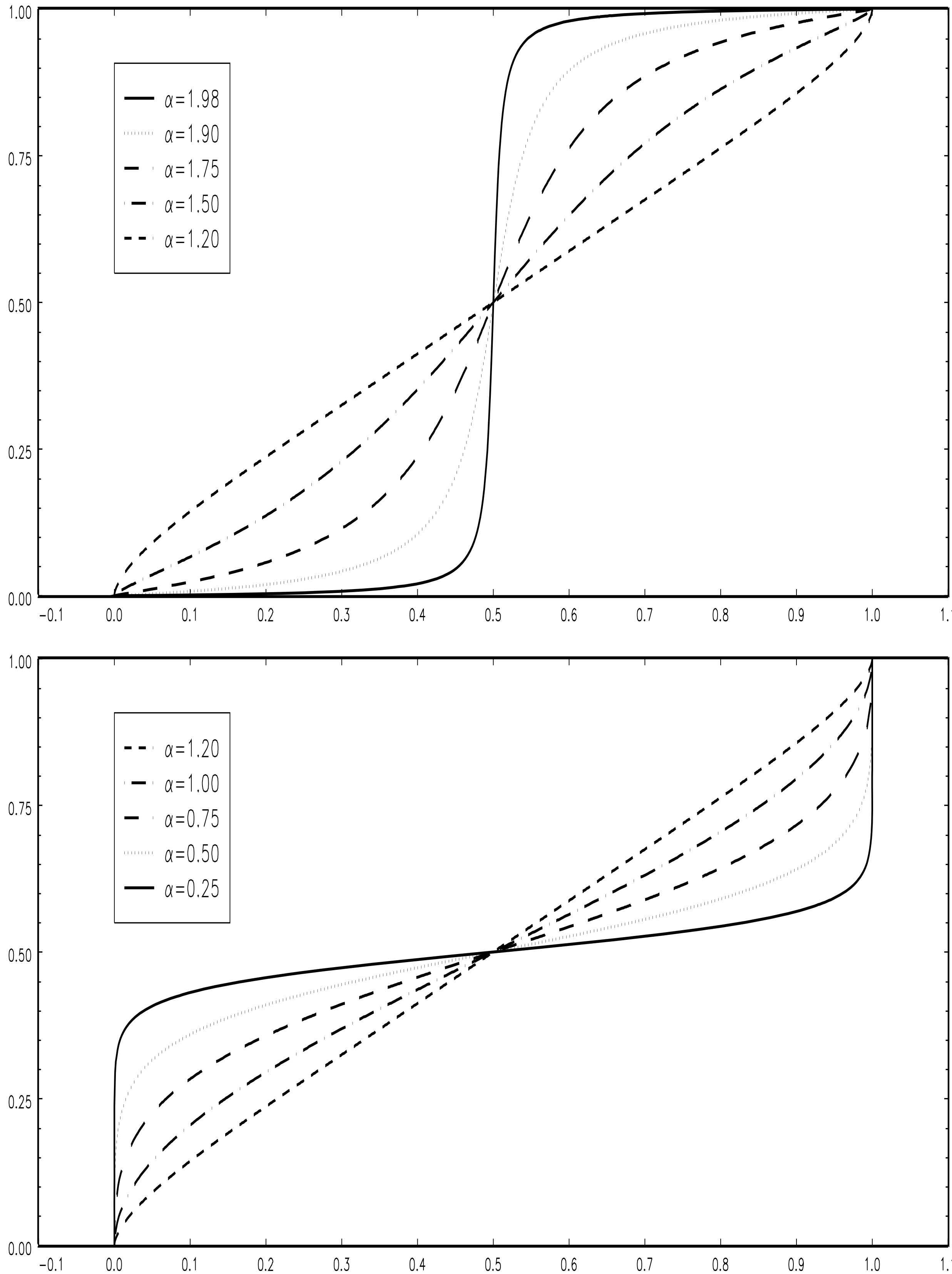

Figure 5. CRSP Returns, July 1963 to December 1992

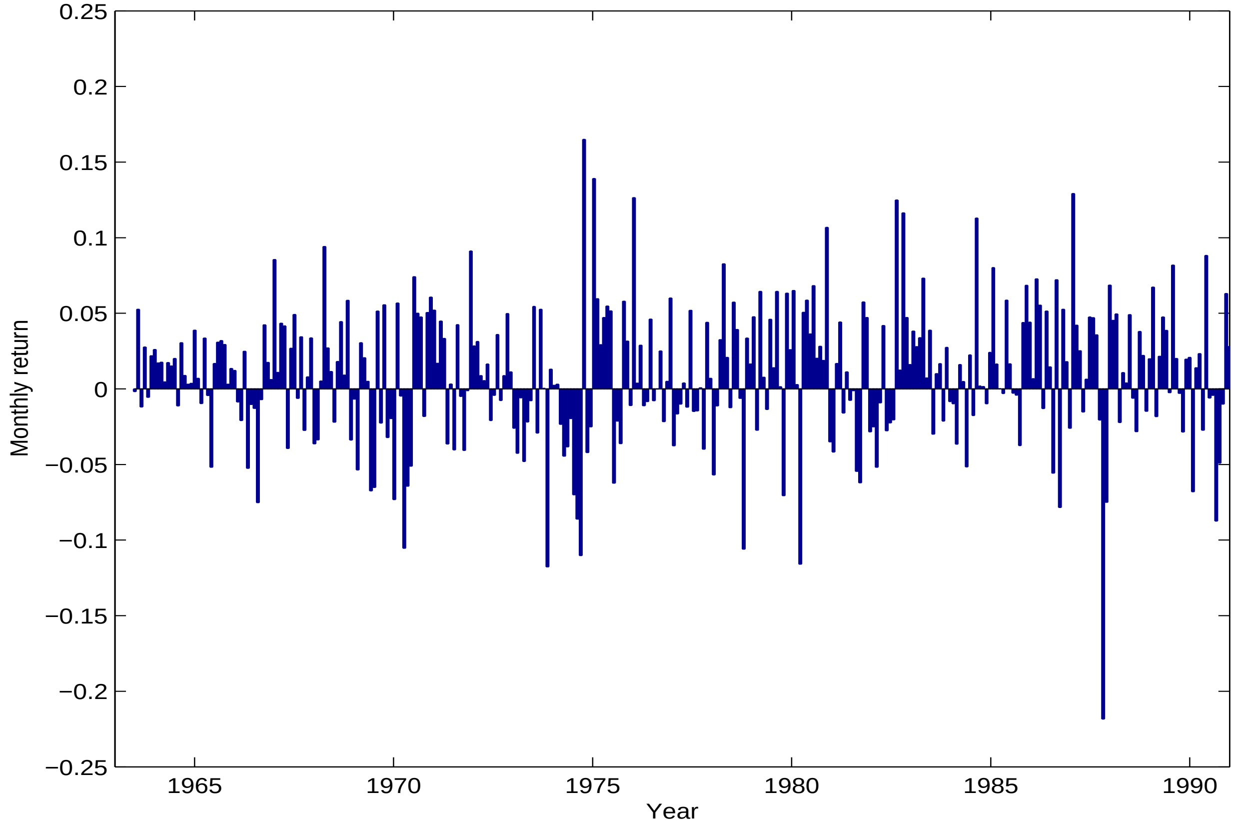

Figure 6. Simulated cdf of R^2, Second-stage Fama-MacBeth regressions