ORIGINAL RESEARCH

Using stochastic frontier analysis instead of data envelopment analysis in modelling investment performance

John D. Lamb[1] · **Kai-Hong Tee[2]**

Abstract

We introduce methods to apply stochastic frontier analysis (SFA) to financial assets as an alternative to data envelopment analysis, because SFA allows us to fit a frontier with noisy data. In contrast to conventional SFA, we wish to deal with estimation risk, heteroscedasticity in noise and inefficiency terms. We investigate measurement error in the risk and return measures using a simulation–extrapolation method and develop residual plots to test model fit. We find that shrinkage estimators for estimation risk makes a striking difference to model fit, dealing with measurement error only improves confidence in the model, and the residual plots are vital for establishing model fit. The methods are important because they allow us to fit a frontier under the assumption that the risks and returns are not known exactly.

1 Introduction

For more than 20 years data envelopment analysis (DEA) has been accepted as the method to fit frontiers and estimate efficiencies of various financial assets (see, for example, Liu et al. (2015)). We know of only two attempts to use stochastic frontier analysis (SFA), its close relative, to fit frontier to a set of assets (Santos et al., 2005; Ferreira & Oliveira, 2016) (though it is used more broadly in finance (Daraio et al., 2020)). This is surprising. DEA and SFA (Bogetoft & Otto, 2011) are the two most important of a number of methods (Wu et al., 2011) for frontier fitting and efficiency estimation. We should choose a method because it is appropriate, not merely because it is convenient.

✉ John D. Lamb
  J.D.Lamb@abdn.ac.uk

  Kai-Hong Tee
  K.Tee@lboro.ac.uk

[1] University of Aberdeen Business School, University of Aberdeen, King Street, Aberdeen AB24 3QY, UK

[2] Loughborough Business School, Loughborough University, Loughborough LE11 3TU, UK

DEA and SFA were not developed for financial assets but for production economics. There the problem is to estimate how much of various outputs is possible with various inputs. We apply them to financial assets using an analogy: return behaves like an output and risk measures behave like inputs. The analogy is imperfect. Some assumptions from production economics, such as free-disposability and convexity (Bogetoft & Otto, 2011), remain reasonable. Others do not. Typically production inputs are measured in comparable units and are not reduced by diversification, but this is not always true of risks (Lamb & Tee, 2012b). And there is invariably uncertainty in the estimates of risk and return measures (Xiao et al., 2021; Lamb & Tee, 2012a).

There are three reasons to think SFA might be a better choice for financial assets than DEA. First, DEA assumes inputs and outputs are fixed and known. But measures of financial assets such as mean and standard deviation are never more than realisations of random variables (from sampling distributions). Thus, a stochastic (involving random variables) model such as SFA is more appropriate than a deterministic model such as DEA. Second, while financial-asset frontier models can have several inputs (risk measures), they nearly always use exactly one output (mean return), the easiest case to fit with SFA. Third, the assumption of nonincreasing returns to scale and a smooth frontier that passes through the origin (Lamb & Tee, 2012a) is precisely what SFA fits.

The objective of this article, then, is to develop methods to allow us to apply SFA in modelling investment fund performance and so help close the gap between the extensive literature on DEA and the surprisingly limited literature on SFA for investment funds. The relevance of SFA is that it is better suited than DEA to situations where there is uncertainty in the input and output measures. So the motivation of this article is to answer the question, given that it is better suited, how can we apply SFA to model investment funds?

In practice it is difficult to fit standard SFA models to financial asset data. When we consider the nature of the data it becomes clear that there are several challenges to overcome and we are unlikely to be able to use the multiplicative translog models (Bogetoft & Otto, 2011) most commonly used in production economics.

The first challenge is estimation risk (Jorion, 1986), which causes sample estimators to overestimate the largest and underestimate the smallest observed statistics. As far as we know it has not been considered before in any SFA, or even DEA, model. We will see that the effect on the estimate of the frontier can be dramatic.

Randomness occurs not just in the position, relative to the efficient frontier, of individual assets, but in the measurement of their inputs and outputs (Xiao et al., 2021). This second challenge is called measurement error (Carroll et al., 1998) and can potentially affect the fit of a model.

The final challenge is to find a suitable model and develop a method to fit it. Multiplicative SFA models are common in production economics. But we cannot use multiplicative assumptions about the distributions of noise and inefficiency terms when we use risk as an input and return as an output. The assumptions of an additive model are more plausible. But if we use an additive model we need to develop new models to deal with heteroscedasticity in the regression residuals. We also develop methods to test the quality of fit.

Our contribution is this. We overcome the three challenges that have limited the use of SFA by developing several new SFA models for financial assets, show how to fit them and how to deal with measurement error. To make it practical for others to fit these models we provide an R (R Core, 2020) package, including a method, similar to residual plots in linear regression analysis, to visualise the residuals of our SFA models and check the model assumptions. This method should be useful for any SFA model, including those with multiple inputs. We test our methods on real data taken from various asset returns.

2 Background

SFA was developed by Meeusen & van den Broeck (1977) and by Aigner et al. (1977). Bogetoft & Otto (2011) provide a good introduction. We fit a frontier

$$y = f(x; \beta) + v - u, \tag{1}$$

where f is a production function, v is a noise and $u > 0$ an inefficiency term. This is an *additive* SFA model. We choose it rather than the *multiplicative* model $y = f(x; \beta) \exp(v) \exp(-u)$ (Bogetoft & Otto, 2011) [Table 7], because it allows us to make more plausible assumptions about the distributions of u and v and deal with heteroscedasticity separately.

SFA was developed for production economics, not financial assets; so we should explore what risk and return measures and what form of f make sense in the context of financial assets. We first note that it is very rare to use a return other than the mean return. So, a single-output model makes sense. It is also convenient, because SFA, like regression, is well-developed for multiple inputs, but so far allows only one output. There are many possible risk measures and it is tempting, for example, to use variance or excess kurtosis. When estimating efficiency it is preferable (Lamb & Tee, 2012a) to use risk measures in the same units as the return measure. This excludes excess kurtosis and variance, but not standard deviation. Coherent measures of risk (Artzner & Delbaen, 1999) such as conditional value-at-risk have become more and more common because they have properties that are desirable when comparing risks of different assets. In particular coherent measures of risk, in common with standard deviation, are convex functions: if the risk of an asset represented by random variable A is given by $x(A)$ then, for $t \in [0, 1]$,

$$x(tA + (1 - t)B) \leq tx(X) + (1 - t)x(B).$$

All the coherent measures of risk we know of are measured in the same units as mean return.

If we allow only mean return as the return measure and coherent measures of risk or standard deviation as the risk measure, what does this tell us about the form of f? First, we can assume $f(\mathbf{0}) = 0$, because we can always obtain zero mean return with zero risk. Plausibly we may also obtain a positive return with zero risk. But we can handle this by subtracting a risk-free rate of return. Second, we assume the efficient frontier is non-decreasing, $\nabla f > \mathbf{0}$ by assuming free disposability of risks: that is, if we can obtain a higher return with lower risk, we assume that we can more efficiently obtain the same higher return by freely taking on extra risk with no return through some financial investment. In practice, we do not know of such an investment, though gambling allows a good approximation. Third, the choice of return and risk measures satisfy

$$y(tA + (1 - t)B) = ty(A) + (1 - t)y(B) \tag{2}$$

$$\mathbf{x}(tA + (1 - t)B) \leq t\mathbf{x}(A) + (1 - t)\mathbf{x}(B) \tag{3}$$

for $t \in [0, 1]$ and random variables A and B representing assets. From $\nabla f > \mathbf{0}$ and Eq. (3) we have

$$f(t\mathbf{x}(A) + (1 - t)\mathbf{x}(B)) \geq f(\mathbf{x}(tA + (1 - t)B))$$
$$= tf(\mathbf{x}) + (1 - t)f(\mathbf{x}(B))$$

by Eq. (2). It follows that f is concave or, equivalently, $\nabla^2 f$ is negative semidefinite.

The properties we have found for f, $f(\mathbf{0}) = 0$, $\nabla f > \mathbf{0}$ and f concave, are satisfied by

$$f(x_1, \ldots, x_p) = \beta_0 x_1^{\beta_1} \cdots x_p^{\beta_p}$$

with $\beta_0 > 0, \beta_1, \ldots, \beta_p \in (0, 1)$ and $\beta_1 + \cdots + \beta_p \le 1$, which is a Cobb–Douglas production function.

Although the methods we develop can be applied with multiple risk measures, we focus on mean–standard-deviation frontiers. That is, we consider the case where y is the mean return and x is the standard deviation in the return because standard deviation is the only risk measure for which we currently have methods to handle estimation risk (Sect. 3.2). In this case the Cobb–Douglas production function is

$$f(x) = \beta_0 x^{\beta_1} \tag{4}$$

with $\beta_0 > 0$ and $\beta_1 \in (0, 1)$. The noise term v and smooth frontier also make SFA a natural choice for fitting a frontier, because we expect substantial random variation and not just inefficiency in the performance of an individual asset.

By contrast, DEA (Bogetoft & Otto, 2011) has been applied much more widely to fitting frontiers and estimating inefficiency of assets or portfolios (see, for example, Gregoriou et al. (2005)). DEA fits a piecewise continuous frontier rather than a smooth one. More plausible frontiers, using diversification-consistent models with nonincreasing returns to scale (Lamb & Tee, 2012a; Liu et al., 2015; Branda, 2015), have recently been developed. The form of these is closer to that more naturally produced by SFA. And recently (see, for example, Lamb & Tee (2012b)) there have been attempts to deal with the stochastic nature of the data.

Although there are stochastic versions of DEA (Olesen & Petersen, 2016), SFA remains the more natural choice when we posit both an inefficiency and a noise in fitting the frontier. But to use it well, we need to consider three issues: the choice of v and u, estimation risk and measurement error. In fact, all but the choice of u should ideally be considered for DEA models too, though we restrict our discussion to SFA.

Typically, v is assumed normal with mean zero and variance σ_v and u half-normal. Normal noise is plausible because it is the error arising from estimating y of Eq. (1) from a sample mean. The half-normal distribution gives positive inefficiency. While its shape has been questioned and others suggested (see Papadopoulos (2021)), the mean–standard-deviation frontier presents a different issue: we expect σ_v to be proportional to x and σ_u to increase, perhaps proportionally, with x or y.

We ought to consider estimation risk when fitting a mean–standard-deviation frontier. When we have n assets or portfolios with similar means and large variances and estimate the means with a multivariate sample mean, the smaller sample means will tend to underestimate and the larger, overestimate the population mean so that we should expect to overestimate the slope of the frontier. Formally, *estimation risk* is the expectation of a risk function (usually mean-squared difference between population and sample), and is something we wish to minimise (Jorion, 1986). It gets worse as n increases (Frost & Savarino, 1986) and affects statistics other than the mean. There are methods to reduce, though not eliminate, estimation risk for both the mean (Jorion, 1986) and covariance matrix (Ledoit & Wolf, 2004, 2017) and these are commonly used in related research areas (Herold & Maurer, 2006; Michaud & Michaud, 2007; Alexander et al., 2009; Antoine, 2012; Davarnia & Cornuéjols, 2017; DeMiguel et al., 2013).

Usually we can assume x and y are measured without error in SFA. But when x and y are the standard deviation and mean of assets, we observe

$$w = x + \psi \quad \text{and} \quad d = y + \kappa$$

where ψ and κ are random variables with mean 0. These are called *measurement errors* (Carroll et al., 1998; Buonaccorsi, 2010) and can influence the coefficients of a regression model such as those we use to fit Eq. (4). As far as we know measurement error has not been considered in any SFA model. One way to deal with it is simulation extrapolation (simex) (Cook & Stefanski, 1994). We can assume that ψ and κ are uncorrelated, but not that they are normal or have constant variance, as required for the original, parametric, simex. But we can generate repeated measurements for w and y, and so adapt the empirical simex method of Devanarayan & Stefanski (2002) to extrapolate model parameters without measurement error.

3 A methodology for mean–standard-deviation SFA

We wish to fit an SFA frontier of the form of Eq. (1) where x is standard deviation and y is mean and so estimate the efficiency of individual assets. To do this we must solve three problems. First, we find a way to deal with heteroscedasticity in both u and v. Section 3.1 does this and then develops a method to test the SFA fit. Second, Sect. 3.2 shows it is straightforward to deal with estimation risk. Third, Sect. 3.3 deals with measurement error.

While estimation risk and measurement error are issues for assets or portfolios, note that the methods of Sect. 3.1 can be applied to any SFA.

3.1 Fitting SFA with heteroscedasticity in *u* and *v*

The frontier we fit should pass through the origin and have nonincreasing returns to scale (Lamb & Tee, 2012a), and while we do not here deal directly with diversification-consistency (Lamb & Tee, 2012a; Liu et al., 2015; Branda, 2015), we expect the SFA frontier shape to be close to that of a diversification-consistent frontier where there are few data points near the frontier.

We assume that f of Eq. (1) has the form of Eq. (4), $v \sim N\left(0, (r_v(x))^2\right)$, u is half-normal with density

$$h_{u|x}(z) = \sqrt{\frac{2}{\pi}} \frac{1}{r_u(x)} \exp\left(-\frac{1}{2}\left(\frac{z}{r_u(x)}\right)^2\right) \qquad (z \geq 0)$$

for some functions r_v and r_u, and u and v are conditionally independent: that is,

$$f_{v,u|x}(w, z \mid x) = f_{v|x}(w \mid x)h_{u|x}(z \mid x).$$

We consider three possibilities:

$$r_t(x) = \sigma_t, \qquad r_t(x) = x\sigma_t, \qquad r_t(x) = \beta_0 x^{\beta_1}\sigma_t, \qquad (t \in \{u, v\}). \qquad (5)$$

That is, we allow the error in v or u to be constant, proportional to x or (cf. Eq. (4)) proportional to y. To fit an SFA frontier we choose r_v and r_u, estimate σ_v and σ_u, and check how well the model fits the assumptions made by the choices of r_v and r_u.

As is usual in SFA, we fit σ_v and σ_u using maximum likelihood estimation. Write $\sigma^2 = (r_v(x))^2 + (r_u(x))^2$, $\lambda = r_u(x)/r_v(x)$ and Φ for the standard normal distribution function. Then we can write $\varepsilon_x = v - u = y - \beta_0 x^{\beta_1}$, and substitute $r_v(x)$ and $r_u(x)$ for σ_v and σ_u

in Eq. (7.10) of Bogetoft & Otto (2011) to get the density of ε_x:

$$\phi_\varepsilon(\varepsilon_x) = 2\Phi\left(-\frac{\lambda\varepsilon_x}{\sigma}\right)\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_x}{\sigma}\right)^2\right), \qquad (6)$$

which is a skew-normal distribution (Azzalini & Capitanio, 2013). Here we just modify the usual definition of the density to allow for different forms of heteroscedasticity in v and u. When we have observations $(x_1, y_1), \ldots, (x_n, y_n)$ we define $\varepsilon_i = y_i - \beta_0 x^{\beta_1}$, $\sigma_i^2 = (r_v(x_i))^2 + (r_u(x_i))^2$ and $\lambda_i = r_u(x_i)/r_v(x_i)$, and compute the log-likelihood function by summing the logarithm of Eq. (6) over all n observations:

$$l(\sigma_v, \sigma_u, \beta_0, \beta_1) = -\frac{n}{2}\log\left(\frac{\pi}{2}\right) + \sum_{i=1}^{n}\left(-\log\sigma_i + \log\Phi\left(-\frac{\lambda_i}{\sigma_i}\varepsilon_i\right) - \frac{1}{2}\left(\frac{\varepsilon_i}{\sigma_i}\right)^2\right). \quad (7)$$

We then obtain estimates of σ_v, σ_u, β_0 and β_1 by maximising the log-likelihood function.

Equation (5) allow us to fit up to nine different SFA models. So we want some way both to compare models and to check that the residuals ε_i of an individual model fit the assumptions we make. In a linear regression model we would plot the standardised regression residuals against the regression predicted values to test these assumptions. We develop a similar method that can be used for SFA.

If the model assumptions are correct then, see Eq. (6), the residual ε_i is a random variate from a distribution with density

$$\phi_\varepsilon(x_i; \lambda_i, \sigma_i) = g(x_i/\sigma_i; \lambda_i)/\sigma_i \quad \text{where } g(x, \lambda) = 2\Phi(-\lambda x)\phi(x).$$

Since g is a skew-normal distribution we can use efficient numerical methods (Azzalini & Capitanio, 2013)[Section 2.1.5] to evaluate its distribution function $G(x; \lambda)$ and thus the distribution function for ε_i,

$$\Phi_\varepsilon(x_i; \lambda_i, \sigma_i) = G(x_i/\sigma_i, \lambda_i)/\sigma_i.$$

Then, if the model assumptions are correct,

$$z_i = \Phi^{-1}(\Phi_\varepsilon(x_i; \lambda_i, \sigma_i)) \sim N(0, 1). \qquad (8)$$

So we can check the model assumptions by plotting z_i against $\beta_0 x_i^{\beta_1}$ $(i = 1, \ldots, n)$ as we do, for example, in Fig. 2.

We use SFA to estimate the technical efficiency of individual asset k in an additive SFA model,

$$\rho_k = 1 - \mathbb{E}[u_k \mid \varepsilon_k]/\beta_0 x_k^{\beta_1}$$

(Bogetoft & Otto, 2011)[Eq. (7.1')]. Since $\mathbb{E}[u_k \mid \varepsilon_k]$ depends only on the distributions of u and v at x_k, we can adopt Eq. (7.17) of Bogetoft & Otto (2011) to estimate

$$\mathbb{E}[u_k \mid \varepsilon_k] = \sigma_{k,*}\left(\frac{\phi(\varepsilon_{k,*})}{1 - \Phi(\varepsilon_{k,*})} - \varepsilon_{k,*}\right) \quad \text{where } \varepsilon_{k,*} = \varepsilon_k\frac{\lambda_k}{\sigma_k}, \ \sigma_{k,*} = \frac{r_v(x_k)r_u(x_k)}{\sigma_k}.$$

3.2 Estimation risk

For fitting an SFA model, we assume we have returns for n assets over T time periods, from which we can estimate sample mean and sample standard deviation vectors
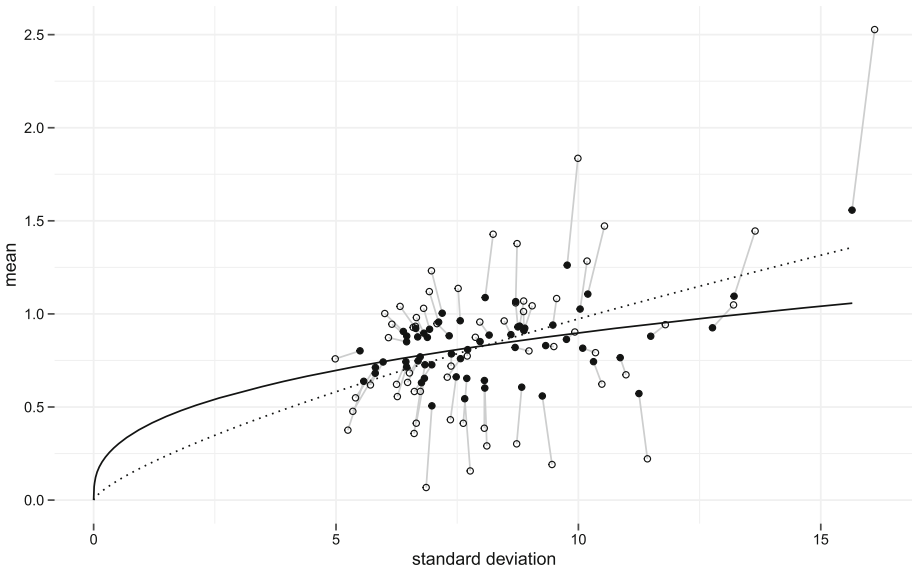
Fig. 1 SFA fit for FTSE 100 stocks

$\mathbf{y} = (y_1, \ldots, y_n)^\top$ and $\mathbf{x} = (x_1, \ldots, x_n)^\top$. Although these are unbiased, setting $\theta = \mathbb{E}[\mathbf{y}]$ does not minimise the risk function

$$(\theta - \mu)^\top \Sigma^{-1} (\theta - \mu) \tag{9}$$

where μ is the vector of true mean returns of the assets and Σ is the covariance matrix (Stein, 1955). The lower means tend to be underestimated and the higher ones overestimated so that a James–Stein estimator (James & Stein, 1961),

$$\hat{\mathbf{y}} = (1 - \alpha)\mathbf{y} + \alpha \overline{y} \mathbf{1}, \tag{10}$$

where $\overline{y}$ is the mean of $y_1, \ldots, y_n$ and $\mathbf{1}$ is a vector of ones of length n, has lower risk for some $\alpha \in [0, 1)$. We set, for some estimator S of the covariance matrix,

$$\alpha = \frac{n + 2}{n + 2 + T(\mathbf{y} - \overline{y}\mathbf{1})^\top S^{-1}(\mathbf{y} - \overline{y}\mathbf{1})} \tag{11}$$

(Jorion, 1986)[Eq. (17)]. This does not necessarily minimise risk function (9) but gives a *shrinkage estimator,* which gives a better estimate than the sample mean for fitting SFA. We use this shrinkage estimator because it is well-established in the finance literature and shrinks towards the mean of all the estimates of means rather than the less plausible target of 0 of many earlier shrinkage estimators.

We can also reduce estimation risk in the sample covariance matrix using recently-developed shrinkage estimators (Ledoit & Wolf, 2004, 2017). We use the nonlinear shrinkage estimator of Ledoit & Wolf (2017) to estimate both S in Eq. (11) and the standard deviations of the asset returns. It has no explicit formula but is implemented in an R-package (Ramprasad, 2016). Like Eq. (9), there is good evidence that this shrinkage estimator works well in practice.

We show the sample and shrinkage estimators for the mean and standard deviation as hollow and filled circles for each of a set of stocks, together with grey lines joining each
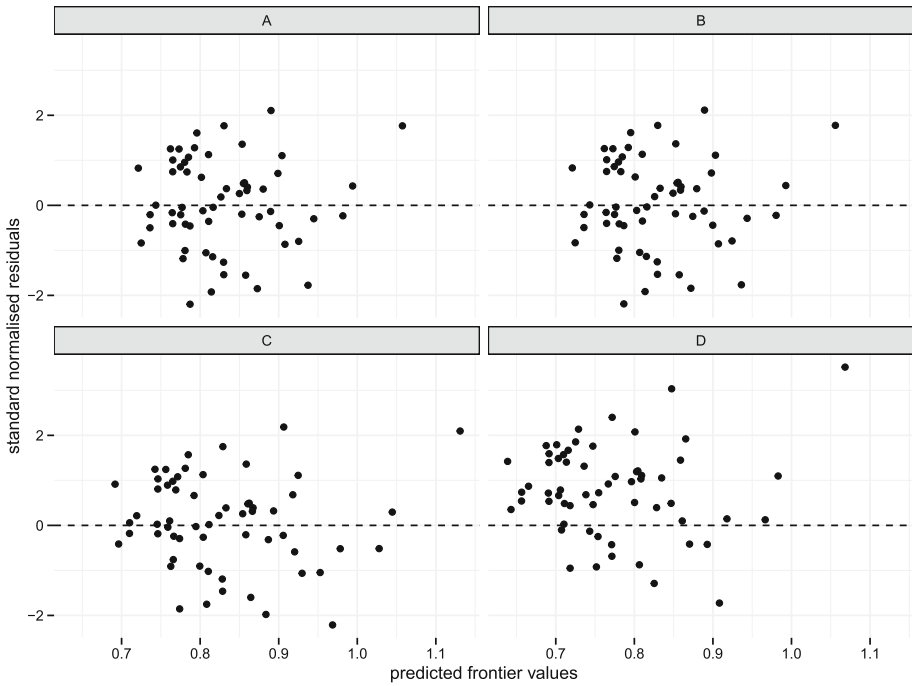
Fig. 2 Standard normalised residuals for four SFA models for FTSE 100 stocks

pair of circles in Figs. 1, 4, 5 and 6. We write x_i and y_i for both the sample and shrinkage estimators of asset i, because the context will indicate which is meant.

3.3 Measurement error

The estimators of the coefficients σ_v, σ_u, β_0 and β_1 of the model of Sect. 3.1 are called *naive* estimators because the model does not account for the measurement errors in the estimates of x and y that arise from sample data. We do not observe x_i and y_i ($i = 1, \ldots, n$) but rather

$$w_i = x_i + \psi_i \quad \text{and} \quad d_i = y_i + \kappa_i,$$

where ψ_i and κ_i are measurement errors satisfying $\mathbb{E}[\psi_i] = \mathbb{E}[\kappa_i] = 0$, $\mathrm{var}(\psi) = \tau_{\psi,i}^2$ and $\mathrm{var}(\kappa_i) = \tau_{\kappa,i}^2$. We can reasonably assume ψ_i and κ_i are independent. Here x_i and y_i are the mean and standard deviation of a distribution. If we assume that the distribution is approximately normal and the samples are independent, then w_i and d_i are observations from the sampling distributions of the sample mean and sample standard deviation, which are independent. Hence ψ_i and κ_i are independent. In practice, we can assume the distribution is approximately normal and the sample values obtained as if from independent samples so that the assumption of independence is reasonable.

We wish to take account of these measurement errors because they may lead us to misestimate the coefficients of the SFA model. One way to do this is to use *simulation extrapolation* (simex) (Cook & Stefanski, 1994) , which give consistent and unbiased estimators Suppose we wish to estimate a coefficient β. Define $\beta(\lambda)$ to be the expected value of the SFA naive estimator of β if $\mathrm{var}(\psi_i) = (1 + \lambda)\tau_{\psi,i}$ and $\mathrm{var}(\kappa_i) = (1 + \lambda)\tau_{\kappa,i}$. Then we can always
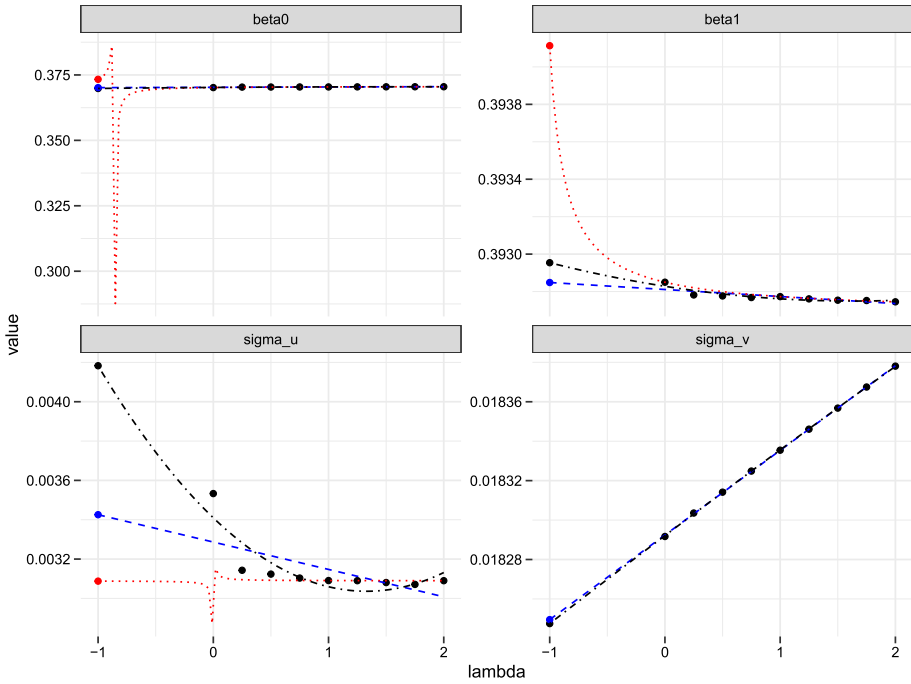
Fig. 3 Simex extrapolation plots for FTSE 100 stocks
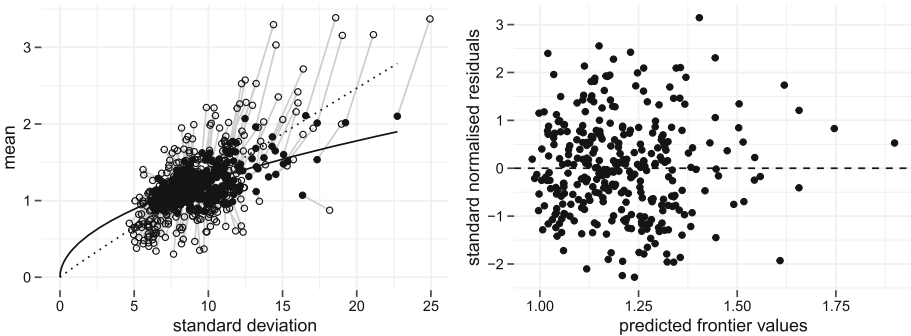

Fig. 4 SFA fit and standard normalised residuals for Standard and Poor stocks

estimate $\beta(0)$. And if we can simulate data to estimate $\beta(\lambda)$ for several values of $\lambda > 0$ we can extrapolate a value for $\beta(-1)$, the coefficient without measurement error.

We estimate standard deviation and mean from T time periods, which means that we can use bootstrap replications of the data to estimate the distributions of ψ_i and κ_i. This means we can adapt the empirical simex method of Devanarayan & Stefanski (2002). If the asset returns are r_{ti} ($t = 1, \dots, T$, $i = 1, \dots, n$) then we generate M bootstrap replications by choosing $u(m, t)$ uniformly with replacement from $1, \dots, T$ and setting

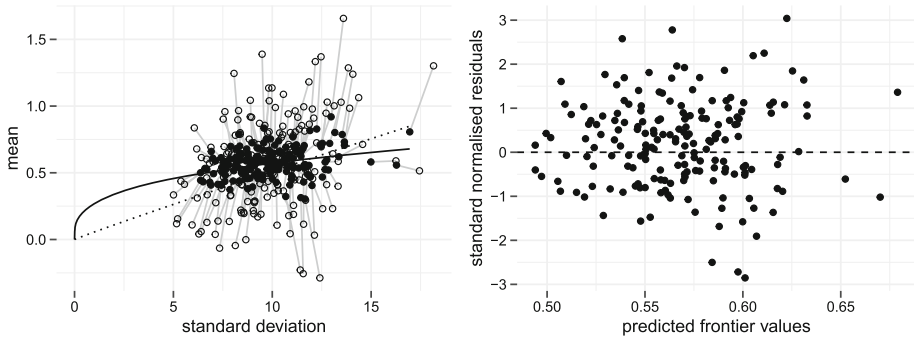$$r_{mti} = r_{u(m,t),i} \qquad (m = 1, \dots, M).$$

Fig. 5 SFA fit and standard normalised residuals for Nikkei stocks

(In practice we choose $M = 2000$, which is typical for the bootstrap.) Write d_i and w_i for the mean and standard deviation of $r_{1i}, \ldots, r_{Ti}$, and d_{mi} and w_{mi} for those of $r_{m1i}, \ldots, r_{mTi}$. Then d_{mi} and w_{mi} are resampled estimates for d_i and w_i. But, we wish to resample shrinkage estimators. If d_i^* and w_i^* are the shrinkage estimators for the means and standard deviations of the asset returns then it is straightforward to check that

$$d_{mi}^* = (d_{mi} - d_i)\frac{w_i^*}{w_i} + d_i^* \quad \text{and} \quad w_{mi}^* = w_{mi}\frac{w_i^*}{w_i}$$

are resampled estimates for the shrinkage estimators.

We can use these estimates in empirical simex. We give the details of the calculations, because we modify the method of Devanarayan & Stefanski (2002) to include an error in y and not just x. First, we choose B (in practice we use $B = 2000$) and let

$$z_{b,i,m} \qquad (i = 1, \ldots, n, \ m = 1, \ldots, M, \ b = 1, \ldots, B)$$

be standard normal variates. Then we calculate

$$\overline{z}_{b,i} = \frac{1}{M}\sum_{m=1}^{M} z_{b,i,m} \quad \text{and} \quad c_{b,i,m} = \frac{z_{b,i,m} - \overline{z}_{b,i}}{\sqrt{\sum_{m=1}^{M}\left(z_{b,i,m} - \overline{z}_{b,i}\right)^2}}.$$

Then we define

$$\overline{w}_i^* = \frac{1}{M}\sum_{m=1}^{M} w_{i,m}^*, \quad \overline{d}_i^* = \frac{1}{M}\sum_{m=1}^{M} d_{i,m}^* \qquad (i = 1, \ldots, n).$$

using the resampled estimates for the shrinkage estimators.

For each value of λ for which we wish to estimate $\beta(\lambda)$ for some coefficient β we compute, for $b = 1, \ldots, B$,

$$w_{b,i,\lambda}^* = \overline{w}_i^* + \sqrt{\frac{\lambda}{M}}\sum_{m=1}^{M} c_{b,i,m}w_{i,m}^* \quad \text{and} \quad d_{b,i,\lambda}^* = \overline{d}_i^* + \sqrt{\frac{\lambda}{M}}\sum_{m=1}^{M} c_{b,i,m}d_{i,m}^*$$

and estimate the coefficients of

$$d_{b,i,\lambda}^* = \beta_{0;b,\lambda} w_{b,i,\lambda}^{* \ \beta_{1;b,\lambda}} + v_{i;b,\lambda}^* - u_{i;b,\lambda}^*$$

by the maximum likelihood method of Sect. 3.1. We then estimate $\beta_{0;\lambda}$, $\beta_{1;\lambda}$ and (usually) $\sigma_{v;\lambda}$ and $\sigma_{u;\lambda}$ as the averages of the B estimates of these coefficients. And we estimate β_0, β_1,
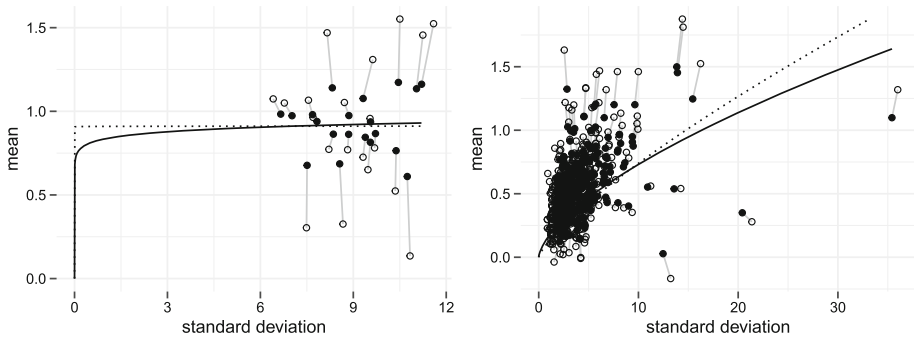
Fig. 6 SFA fits for DAX stocks and hedge funds

σ_v and σ_u using one of the standard extrapolation methods (Cook & Stefanski, 1994; Carroll et al., 1998). These first fit the estimates $\beta(\lambda)$ to a model $\beta(\lambda) = g(\lambda)$, then extrapolate to get $\beta(-1)$ as an estimate of β. There is no one right function g. Typically either the rational or quadratic extrapolant,

$$g(\lambda) = c_1 + \frac{c_2}{c_3 + \lambda} \quad \text{or} \quad g(\lambda) = c_1 + c_2\lambda + c_3\lambda^2, \tag{12}$$

are used. The first has better theoretical justification. The second is more stable. Typically we test both, together with a linear extrapolant, and plot the best fits (see Fig. 3) before choosing a value for β.

3.4 Implementation

We implement our methods in an R (R Core, 2020) package, using C++ (Eddelbuetel, 2013) when R code would be too slow.

We use the nlshrink R-package (Ramprasad, 2016) for covariance and compute mean shrinkage using Eqs. (10)–(11) with the shrinkage covariance estimate. We implement SFA fitting as described in Sect. 3.1, using the C++ with the GNU scientific library (Galassi et al., 2009) to compute the log-likelihood function (7) and also to maximise it (we use the Nelder–Mead simplex method because it is convenient and easy to implement). We also use the GNU scientific library to compute $G(x; \lambda)$ numerically and so compute z_i in Eq. (8) to produce residual plots to test model fit. We implement simex in R as described in Sect. 3.3.

4 Fitting SFA to historic asset returns

We now test the methods of Sect. 3 on some real data. The data we use are stocks in each of the FTSE 100 (61 stocks), Standard and Poor (S&P, 326 stocks) and Nikkei (183 stocks) indices together with DAX and hedge-fund data described below. For each of these stocks we obtain 300 consecutive monthly returns to the beginning of June 2019 from Datastream (2019). As usual in efficiency modelling for asset returns, we ignore time-series effects when estimating mean and standard deviation, though our methods could be adapted to consider them. Since we fit historic asset returns rather than prices, we expect the data to exhibit stationarity. Nonetheless, we use augmented Dickey-Fuller (ADF) tests to test for unit roots. The p-values were negligible except for two hedge funds with p-values of (0.0591) and

Table 1 Summary statistics for data

Statistic	Stocks	Quantile 0	0.25	0.5	0.75	1
Mean	FTSE	0.068	0.583	0.873	1.04	2.527
	S&P	0.301	0.863	1.115	1.482	3.387
	Nikkei	− 0.288	0.36	0.58	0.798	1.657
Standard deviation	FTSE	4.984	6.616	7.707	9.047	16.111
	S&P	4.674	7.192	8.6	10.637	24.937
	Nikkei	5.005	8.343	9.755	11.16	18.157
Skewness	FTSE	− 0.69	− 0.055	0.028	0.23	2.003
	S&P	− 1.709	− 0.132	0.108	0.431	5.318
	Nikkei	− 0.289	0.01	0.245	0.516	7.143
Kurtosis	FTSE	− 0.127	1.082	1.946	3.573	23.518
	S&P	− 0.225	1.252	2.078	3.796	56.28
	Nikkei	− 0.147	0.568	1.28	2.459	89.716

(0.57). For purpose of verification, we re-test for unit roots using the Philips–Perron (PP) method, [1]

Table 1 shows summary statistics for the monthly percentage stock returns. The values of skewness and kurtosis immediately suggest many of the returns are not plausibly normally or lognormally distributed. However, we are not fitting the returns, but rather estimates of their means and standard deviations, which, by the central limit theorem, are much more likely to be approximately normally distributed. So a translog function and multiplicative model, which would assume lognormally distributed mean returns, are not appropriate, even if we could find a way to deal with the negative mean returns in the Nikkei data. On the other hand, an additive model assumes approximately normal mean returns, which are plausible here.

Figure 1 shows the means and standard deviations of the FTSE stocks. For each stock, we show the sample mean and standard deviation by an unfilled circle and the shrinkage mean and standard deviation by a filled one. We join the circles by a grey line to show the effect of using shrinkage estimators. We fit an SFA model

$$y = \beta_0 x^{\beta_1} + v - u$$

using the shrinkage estimators and $r_v(x) = x\sigma_v$ and $r_u(x) = \sigma_u$; see Eq. (5). That is, we assume the standard deviation in the noise is proportional to x while that of the inefficiency is constant and find estimates of the coefficients

$$\beta_0 = 0.387, \quad \beta_1 = 0.366, \quad \sigma_v = 0.018 \quad \text{and} \quad \sigma_u = 0.003. \tag{13}$$

The solid black line shows the SFA efficient frontier. The dotted line shows the SFA frontier for the sample means and standard deviations with the same model assumptions. It also shows that how much estimation risk can distort our estimate of the frontier. The technical

---

[1] We conducted all unit root tests using EViews with its standard set up. Both the ADF and the PP test for unit root are on *level* and include *trend* and *intercept* in the test equation. The set up includes a maximum of 11 lags for the ADF test based on the Schwarz information criterion. For PP we use the Newey–West test bandwidth. Since there are 950 tests, we do not report the details further and infer that the data are plausibly stationary with any evidence of unit roots plausibly attributable to chance which suggest stationarity for all our data.

efficiencies of the stocks range from 0.9969 to 0.9979, reflecting the small σ_u. The model assigns nearly all the variation in stock performance (relative to the frontier) to noise.

The dashed grey line in Fig. 1 shows, for comparison, an estimated diversification-consistent DEA frontier (Lamb & Tee, 2012a; Liu et al., 2015). Like the SFA frontiers it passes through the origin, has nonegative gradient and is concave. However, since we do not know of a way to estimate a DEA frontier that takes account of uncertainty in risk and return measures, no point can be above the DEA frontier.

It is interesting to note the much more plausible frontier shape close to the origin, where we expect nonnegligible return for negligible risk, because there is typically a risk-free rate of return (DeMiguel et al., 2009). We have also fitted SFA models with a risk-free rate of return, which we get by subtracting US treasury bill returns from each period. The results are negligibly different. For example, for the FTSE stocks we get $\beta_0 = 0.394$ and $\beta_1 = 0.361$. So we do not report these results further.

We wish to check both the quality of fit of the SFA model and whether different assumptions of the form of r_v and r_v work better. We do both by plotting charts of the standard normalised model residuals calculated using Eq. (8) against the predicted frontier values calculated using Eq. (4). Panel A of Fig. 2 shows the residual plot for the model of Fig. 1: the fit is good and the model plausible.

The most plausible alternative model uses $r_v = \beta_0 x^{\beta_1} \sigma_v$ instead of $r_v = \sigma_v$ and Fig. 2, Panel B shows its residual plot. The fit is good and the model coefficients hardly different, which are less surprising because the frontier is close to constant for standard deviations in the range we observe. Panel C shows SFA with $r_v = \beta_0 x^{\beta_1} \sigma_v$ and $r_u = \sigma_u$, and panel D shows the standard SFA model, with $r_v = \sigma_v$ and $r_u = \sigma_u$. These fit less well and the standard SFA model has the least good fit.

We consider also measurement error as described in Sect. 3.3, because it can improve model coefficients. Figure 3 shows plots of the model coefficients with various levels λ of noise added. The coefficients of Eq. (13) are approximately the simex ones at $\lambda = 0$, and we extrapolate values without measurement error at $\lambda = -1$. We use three standard extrapolants: the rational linear (when available, dotted lines) and quadratic (dot-dash lines) extrapolants of Eq. (12), and a linear (dashed lines) extrapolant. The changes in β_0, β_1 and σ_v are too small to noticeably affect model fit. The quadratic extrapolant, which we would choose for σ_u, gives greater relative change but not enough to reduce the efficiencies noticeably. While it is important to check the effects of measurement error, they are not guaranteed to impact the model coefficients substantially (van Smeden et al., 2020).

Figure 4 shows the SFA frontiers for the S&P stocks in the same format as Fig. 1 with $r_v(x) = x\sigma_v$ and $r_u(x) = \sigma_u$ as before, together with a standard normalised residual plot for the shrinkage frontier. The coefficients of the SFA fit with shrinkage estimators,

$$\beta_0 = 0.395, \quad \beta_1 = 0.503, \quad \sigma_v = 0.017 \quad \text{and} \quad \sigma_u = 0.002$$

are similar to those of the FTSE stocks in Eq. (13). The residual plot suggests very good model fit. Efficiencies are, again, all close to 1. And simex shows even more negligible effects from measurement errors.

Figure 5 shows the SFA frontiers for the Nikkei stocks, together with a standard normalised residual plot for the SFA frontier using shrinkage estimators. This time the best fit comes from the standard $(r_v(x) = \sigma_v, r_u(x) = \sigma_u)$ SFA model: the residual plot suggests mild heteroscedasticity, but changing the r_v and r_u always produces a much worse residual plot even when the SFA frontier appears reasonable. The shrinkage SFA coefficients are

$$\beta_0 = 0.270, \quad \beta_1 = 0.325, \quad \sigma_v = 0.1 \quad \text{and} \quad \sigma_u = 0.012.$$

And the efficiencies in the range 0.980–0.987, slightly lower than those of the FTSE and S&P stocks, as we expect with a larger value of σ_v. As before, simex hardly changes the coefficient estimates.

We conclude with two cases where our methodology works less well. Figure 6 left shows SFA fits for 280 monthly percentage returns for 20 DAX stocks to end of May 2019 (Datastream, 2019). And Fig. 6 right shows fits for 168 monthly percentage returns for hedge funds to the end of December 2019 (Refinitiv, 2019). The DAX fit and residual plot are good but the sample size limits our confidence in the model. By contrast the large sample of hedge funds do not fit any variant of our model assumptions well. It may be that the mix of hedge fund strategies makes a single fit less sensible. Or it may be that we really need a different form for $f(x; \beta)$ in Eq. (1). In any case, even if is not clear from Figure 6 right, the residual plots show none of our models are plausible in these two cases.

5 Conclusions

SFA can be seen as an alternative to DEA for fitting frontiers and finding efficiencies for asset returns. However, asset return and risk measures do not behave entirely like the inputs and outputs of production economics for which SFA was developed. So we have had to develop new theory and methods. Specifically, we show why a Cobb–Douglas frontier is plausible and how to deal with estimation risk. Then we note that, probably in contrast to production economics, we may expect heteroscedasticy in the SFA regression residuals and measurement error in the risk measures. So we have developed methods to handle these and to test for model fit. In doing so we have had to resolve several issues and have some findings that may be surprising.

One intriguing empirical finding is that there was very little evidence of technical inefficiency in our SFA models. This finding is similar to that of Ferreira & Oliveira (2016) and may be an indication that, for our data, the market is good at estimating both the mean return and the volatility of assets. It may also be a limitation of modelling with one input or a feature of stocks drawn from market indices.

The first issue we resolve is the choice of SFA model. We consider several that allow heteroscedasticity in the noise v or inefficiency u and develop log-likelihood functions to fit them. We find errors of the form $x\sigma_v$ for v and σ_u for u fit best, though the form $\beta_0 x^{\beta_1} \sigma_u$ also fits reasonably for u. This leads us to the second issue: how to test model fit.

We resolve this issue by developing a new method for creating residual plots, like those used in linear regression modelling. These are much better for checking model assumptions and model fit than plotting the data and frontier, and we recommend using them for any SFA model. We anticipate that they will be especially useful in models with more than one input variable.

A third issue we had to investigate was the effect of measurement error, which arises because of uncertainty in the measured values of mean and variance. We resolved this by adapting an empirical simex method to work with SFA. The surprising finding was that the measurement errors had very little influence on the model parameters. This, however, is not a general result and we recommend using simex to check any SFA model that may be influenced by measurement error in the same way as one would routinely check for heteroscedasticity or non-normality. Such checks show the accuracy of the findings even if they hardly improve them.

Finally, we address the issue of estimation risk. Estimation risk is the tendency of statistics like the sample mean and standard deviation to underestimate small values and overestimate large ones when we have several values. As far as we know, no previous frontier-fitting method has dealt with estimation risk, and SFA studies for assets (Santos et al., 2005; Ferreira & Oliveira, 2016) do not show the frontier to let us check plausibility of fit. We resolve the issue by introducing shrinkage estimators instead of the sample ones. The impact is too striking to be ignored. Without shrinkage estimators, our results suggest a frontier whose curvature substantially underestimates that of the true one. With shrinkage estimators, the frontier becomes plausible.

5.1 Limitations and future recommendations

The methods of this article currently have several limitations. They do not work well when there are few assets, as we saw for the DAX stocks. This may be a difficult issue to overcome without adding more assets, because asset returns typically have high volatility relative to return.

The methods can also fail for some assets such as the hedge funds. We can think of three plausible explanations that might be addressed in future studies. First, the hedge funds vary in their choice of investment strategy. So it may be possible to fit separate strategies separately. Second, not all smooth, nondecreasing, concave functions that pass through the origin are well represented by a Cobb–Douglas frontier. So it may be helpful to test other frontier forms both for assets and in production economics. Third, and perhaps most challenging, the assumptions about the distribution of errors and, (since the error in a mean is more plausibly normal) especially, the inefficiencies may be inappropriate in the hedge fund case.

The most obvious limitation is that we only consider one input and one output. The one-output limitation is a well-known issue in SFA. It is less important in fitting assets, where mean return is nearly always the only output wanted. But, future studies could address the issue of multiple inputs. In fact, the theory we have developed can readily be used in the multiple-input case with suitable risk measures. Such, coherent (Artzner & Delbaen, 1999), risk measures are already widely used. The real limitation is that we do not yet have methods to reduce estimation risk for them. So, besides investigating the shape and distribution of inefficiencies, a major challenge for future research is finding estimators of coherent risk measures that substantially reduce estimation risk. This challenge is important both for SFA and DEA, where, to the best of our knowledge, estimation risk has not been considered at all.

Declaration

References

Aigner, D., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics, 6*, 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

Alexander, G. J., Baptista, A. M., & Shu, Y. (2009). Reducing estimation risk in optimal portfolio selection when short sales are allowed. *Managerial and Decision Economics, 30*, 281–305. https://doi.org/10.1002/mde

Antoine, B. (2012). Portfolio selection with estimation risk: A test-based approach. *Journal of Financial Econometrics, 10*, 164–197. https://doi.org/10.1093/jjfinec/nbr008

Artzner, P., & Delbaen, F. (1999). Coherent measures of risk. *Mathematical Finance, 9*, 203–228. https://doi.org/10.1111/1467-9965.00068

Azzalini, A., & Capitanio, A. (2013). *The Skew-Normal and Related Families*. Cambridge University Press.

Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA and R*: Springer.

Branda, M. (2015). Diversification-consistent data envelopment analysis based on directional-distance measures. *Omega, 52*, 65–76. https://doi.org/10.1016/j.omega.2014.11.004

Buonaccorsi, JP. (2010). *Measurement error: models, methods and applications*. Chapman and Hall/CRC

Carroll, R. J., Ruppert, D., & Stefanski, L. A. (1998). *Measurement Error in Nonlinear Models*. CRC Press.

Cook, J., & Stefanski, L. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association*, 89, 1314–1328.10.2307/2290994, https://doi.org/10.2307/2290994, http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1994.10476871

Daraio, C., Kerstens, K., Nepomuceno, T., et al. (2020). Empirical surveys of frontier applications: A meta-review. *International Transactions in Operational Research, 27*, 709–738. https://doi.org/10.1111/itor.12649

Datastream (2019) Thomson Reuters Datastream

Davarnia, D., & Cornuéjols, G. (2017). From estimation to optimization via shrinkage. *Operations Research Letters, 45*, 642–646. https://doi.org/10.1016/j.orl.2017.10.005,http://linkinghub.elsevier.com/retrieve/pii/S0167637717305503

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies, 22*, 1915–1953. https://doi.org/10.1093/rfs/hhm075

DeMiguel, V., Martin-Utrera, A., & Nogales, F. J. (2013). Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance, 37*, 3018–3034. https://doi.org/10.1016/j.jbankfin.2013.04.033, http://dx.doi.org/10.1016/j.jbankfin.2013.04.033

Devanarayan, V., & Stefanski, L. A. (2002). Empirical simulation extrapolation for measurement error models with replicate measurements. *Statistics and Probability Letters, 59*, 219–225. https://doi.org/10.1016/S0167-7152(02)00098-6

Eddelbuetel, D. (2013). *Seamless R and C++ Integration with Rcpp. Springer*. https://doi.org/10.1007/978-1-4614-6867-7

Ferreira, N. B., & Oliveira, M. M. (2016). Portfolio efficiency analysis with SFA: The case of PSI-20 companies. *Applied Economics, 48*, 1–6. https://doi.org/10.1080/00036846.2015.1073837, http://www.tandfonline.com/doi/full/10.1080/00036846.2015.1073837

Frost, P. A., & Savarino, J. E. (1986). Portfolio size and estimation risk. *The Journal of Portfolio Management, 12*, 60–64.

Galassi, M., Davies, J., Theiler, J., & et al. (2009). *GNU Scientific Library Reference Manual*. Network Theory, https://gsl.gnu.org

Gregoriou, G. N., Rouah, F., Satchell, S., et al. (2005). Simple and cross efficiency of CTAs using data envelopment analysis. *The European Journal of Finance, 11*, 393–409. https://doi.org/10.1080/1351847042000286667

Herold, U., & Maurer, R. (2006). Portfolio choice and estimation risk: A comparison of Bayesian to heuristic approaches. *ASTIN Bulletin, 36*, 135–160. https://doi.org/10.2143/AST.36.1.2014147

James, W., & Stein, C. (1961). Estimation with quadratic loss. In: Berkeley Symposium on Mathematical Statistics and Probability, 361–379

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *The Journal of Financial and Quantitative Analysis, 21*, 279–292. https://doi.org/10.2307/2331042, http://www.jstor.org/stable/2331042?origin=crossref

Lamb, J. D., & Tee, K. H. (2012). Data envelopment analysis models of investment funds. *European Journal of Operational Research, 216*, 687–696. https://doi.org/10.1016/j.ejor.2011.08.019

Lamb, J. D., & Tee, K. H. (2012). Resampling DEA estimates of investment fund performance. *European Journal of Operational Research, 223*, 834–841. https://doi.org/10.1016/j.ejor.2012.07.015

Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis, 88*, 365–411. https://doi.org/10.1016/S0047-259X(03)00096-4

Ledoit, O., & Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *Review of Financial Studies, 30*, 4349–4388. https://doi.org/10.1093/rfs/hhx052

Liu, W., Zhou, Z., Liu, D., et al. (2015). Estimation of portfolio efficiency via DEA. *Omega, 52*, 107–118. https://doi.org/10.1016/j.omega.2014.11.006

Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review, 18*, 435–444.

Michaud, R. O., & Michaud, R. (2007). Estimation error and portfolio optimization: A resampling solution. *The Journal of Investment Management, 45*(1), 31–42.

Olesen, O. B., & Petersen, N. C. (2016). Stochastic data envelopment analysis—A review. *European Journal of Operational Research, 251*, 2–21. https://doi.org/10.1016/j.ejor.2015.07.058

Papadopoulos, A. (2021). Stochastic frontier models using the generalized exponential distribution. *Journal of Productivity Analysis, 55*, 15–29. https://doi.org/10.1007/s11123-020-00591

R Core Team (2020) R: A language and environment for statistical computing. https://www.r-project.org/

Ramprasad, P. (2016). nlshrink: Non-linear shrinkage estimation of population eigenvalues and covariance matrices, based on publications by Ledoit and Wolf (2004, 2015, 2016). https://cran.r-project.org/web/packages/nlshrink/index.html

Refinitiv (2019) Lipper TASS database. https://www.refinitiv.com/en/policies/third-party-provider-terms/tass-database

Santos, A., Tusi, J., Da Costa, N., et al. (2005). Evaluating Brazilian mutual funds with stochastic frontiers. *Economics Bulletin, 13*(2), 1–6.

van Smeden, M., Lash, T. L., & Groenwold, R. H. (2020). Reflection on modern methods: Five myths about measurement error in epidemiological research. *International Journal of Epidemiology, 49*, 338–347. https://doi.org/10.1093/ije/dyz251

Stein, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In: Proceedings of the 3rd Berkeley Symposium on Probability and Statistics, pp 197–208

Wu, D. D., Zhou, Z., & Birge, J. R. (2011). Estimation of potential gains from mergers in multiple periods: A comparison of stochastic frontier analysis and data envelopment analysis. *Annals of Operations Research, 186*, 357–381. https://doi.org/10.1007/s10479-011-0903-6

Xiao, H., Ren, T., Zhou, Z., et al. (2021). Parameter uncertainty in estimation of portfolio efficiency: Evidence from an interval diversification-consistent DEA approach. *Omega (United Kingdom), 103*,. https://doi.org/10.1016/j.omega.2020.102357

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.