The Stata Journal (2023)
23, Number 1, pp. 197–229 DOI: 10.1177/1536867X231162033

Two-tier stochastic frontier analysis using Stata

Yujun Lian
Lingnan College
Sun Yat-sen University
Guangzhou, China
arlionn@163.com

Chang Liu
Lingnan College
Sun Yat-sen University
Guangzhou, China
liuch288@mail2.sysu.edu.cn

Christopher F. Parmeter
Miami Herbert Business School
University of Miami
Miami, FL
cparmeter@bus.miami.edu

Abstract. In this article, we introduce the `sftt` command, which fits two-tier stochastic frontier (2TSF) models with cross-sectional data. Like most frontier models, a 2TSF model consists of a linear frontier model and a composite error term. The error term is assumed to be a mixture of three components: two one-sided inefficiency terms—strictly nonnegative and nonpositive, respectively—and a symmetric noise term. When providing appropriate distributional assumptions, `sftt` can fit models with exponential and half-normal specifications. `sftt` also fits 2TSF models with the scaling property to mitigate concerns over distributional specifications. In addition, we provide two subcommands, `sftt sigs` and `sftt eff`, to assist in postestimation efficiency analysis. We provide an overview of the 2TSF literature, a description of the `sftt` command and its options, and examples using simulated and actual data.

Keywords: st0705, sftt, two-tier stochastic frontier model, inefficiency, information asymmetry

1 Introduction

In this article, we introduce the `sftt` command, which fits parametric two-tier stochastic frontier (2TSF) models using cross-sectional data. Since Polachek and Yoon (1987), who introduced this model to reflect a boundary between observed wages and what workers were willing to accept simultaneously with what firms were willing to pay, this class of models has become a popular tool for studying price bargaining (Kumbhakar and Parmeter 2009, 2010; Blanco 2017; Fried and Tauer 2019), information asymmetry (Lu, Lian, and Lu 2011; Liu, Yao, and Wei 2019), and corporate governance (Lin, Liu, and Sun 2017; Lyu, Decker, and Ni 2018; Ge et al. 2020), amongst other application domains. A detailed review of these models can be found in Papadopoulos (2021).

The appeal of 2TSF models is that they allow for measurement of the impact of asymmetries in markets where economic agents are operating in opposite directions, such as workers and firms, buyers and sellers, and countries giving or receiving aid. As alternatives to the classic stochastic frontier literature, which has a single, explicit upper or lower boundary, 2TSF models have both upper and lower boundaries. For example, a home seller has the lowest price they would sell their house for, while simultaneously a home buyer has a maximum price they would pay for a house. It is likely that these two prices will differ and that the final observed price carries information on the relative positions of both agents.

To our knowledge, the estimation of 2TSF models is currently unavailable in distributed Stata commands (or in any other statistical languages). Nevertheless, these models remain popular in various application domains. As such, we developed the `sftt` command and designed the syntax following the popular `sfcross` command by Belotti et al. (2013).

To fit 2TSF models, researchers usually use two kinds of model assumptions and estimation techniques. The first is to impose distributional assumptions and estimate via maximum likelihood. For example, the one-sided terms in the composite error are assumed to follow either the exponential distribution (Polachek and Yoon 1987) or the half-normal distribution (Papadopoulos 2015), from which the joint distribution function of the composite error term and the corresponding likelihood function can be derived.

The second approach to fit 2TSF models is to use nonlinear least squares (NLS). Parmeter (2018) proposed using NLS when observable characteristics are presumed to impact the deviations from the boundaries. In this case, if the scaling property is assumed (Wang and Schmidt 2002), NLS can be used to fit the model with no distributional requirements.

Within Stata, multiple ways exist to fit a (single-tier) stochastic frontier model. The built-in commands `frontier` and `xtfrontier` fit stochastic frontier models well, and `sfcross` and `sfpanel` by Belotti et al. (2013) are quite popular among researchers. With the development of stochastic frontier models, many new commands have been published in recent years. `sfkk` and `xtsfkk` by Karakaplan (2017, 2022) can control endogeneity in stochastic frontier models, while Fé and Hofler (2020) developed `sfcount` to fit count-data stochastic frontier models. However, these commands fit only production or cost-frontier models. If we are interested in topics like price bargaining and information asymmetry, 2TSF models, which contain two one-sided terms to capture the inefficiencies in different directions, are required.

`sftt` provides a user-friendly way to fit 2TSF models with either distributional assumptions or the scaling property. When estimating with distributional assumptions, `sftt` fits a 2TSF model with either the exponential or the half-normal specification. `sftt` can also fit a 2TSF model with the scaling property to address concerns over distributional specifications. `sftt sigs` and `sftt eff` are two subcommands to decompose error terms and calculate measures of inefficiency, which are helpful in the postestimation efficiency analysis.

The remainder of the article is organized as follows. In sections 2 and 3, we briefly review the 2TSF model and discuss the rudiments of estimation using maximum likelihood and NLS, respectively. Section 4 describes the syntax of `sftt`, focusing on the main options. Sections 5 and 6 illustrate the use of the `sftt` command using simulated data and three empirical applications from the 2TSF literature. Finally, section 7 offers some conclusions.

2 2TSF models with distributional assumptions

We begin our discussion with the benchmark specification of the 2TSF model proposed by Polachek and Yoon (1987) and discuss estimation via maximum likelihood and many postestimation objects that researchers may find interesting. We then turn our attention to the half-normal specification of Papadopoulos (2015). We intentionally do not discuss the derivation of the corresponding criterion functions but refer the reader to the cited literature for details on the estimation of each model.

2.1 The exponential specification

2.1.1 Model estimation

Following Kumbhakar and Parmeter (2009), consider the 2TSF model

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \mathbf{v} - \mathbf{u} + \mathbf{w} \end{aligned} \tag{1}$$

where $\mathbf{y}$ is an $n \times 1$ vector containing observations of the outcome variable, $\mathbf{X}$ is an $n \times K$ matrix of covariates, $\boldsymbol{\delta}$ is a $K \times 1$ vector of the coefficients, and $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of the composite error term, with $\mathbf{u}$ and $\mathbf{w}$ being two one-sided inefficiency terms and $\mathbf{v}$ capturing stochastic noise. These three components are assumed to be jointly independent. For each i, we have

$$\begin{aligned} v_i &\sim \text{i.i.d. } \mathcal{N}(0, \sigma_v^2) \\ u_i &\sim \text{i.i.d. Exp}(\sigma_u) \\ w_i &\sim \text{i.i.d. Exp}(\sigma_w) \end{aligned} \tag{2}$$

where i.i.d. stands for independent and identically distributed, $\mathcal{N}(0, \sigma_v^2)$ denotes a normal distribution with mean 0 and variance σ_v^2, and $\text{Exp}(\sigma_z)$ denotes a random variable z that is exponentially distributed with mean σ_z and variance σ_z^2.

Using the assumptions in (2), we can derive the probability density function of ε_i,

$$f(\varepsilon_i) = \frac{e^{a_{1i}}}{\sigma_u + \sigma_w}\Phi(b_{1i}) + \frac{e^{a_{2i}}}{\sigma_u + \sigma_w}\Phi(b_{2i})$$

where $a_{1i} = (\varepsilon_i/\sigma_u) + \{\sigma_v^2/(2\sigma_u^2)\}$, $b_{1i} = -\{(\varepsilon_i/\sigma_v) + (\sigma_v/\sigma_u)\}$, $a_{2i} = \{\sigma_v^2/(2\sigma_w^2)\} - (\varepsilon_i/\sigma_w)$, and $b_{2i} = (\varepsilon_i/\sigma_v) - (\sigma_v/\sigma_w)$. $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF).

The log-likelihood function for a sample of n observations is

$$\ln L(\boldsymbol{\varepsilon}|\mathbf{y},\mathbf{X},\boldsymbol{\theta}) = -n\ln(\sigma_u+\sigma_w) + \sum_{i=1}^{n} \ln\left\{e^{a_{1i}}\Phi(b_{1i}) + e^{a_{2i}}\Phi(b_{2i})\right\}$$

where $\boldsymbol{\theta} = \left(\boldsymbol{\delta}', \sigma_v, \sigma_u, \sigma_w\right)'$. Estimates can be obtained by directly maximizing the above log-likelihood function.

2.1.2 Measuring one-sided terms in levels

Using the maximum likelihood estimates, the conditional distributions of u_i and w_i can be written as

$$f(u_i|\varepsilon_i) = \frac{\lambda e^{-\lambda u_i}\Phi(u_i/\sigma_v + b_{2i})}{\chi_{1i}}$$
$$f(w_i|\varepsilon_i) = \frac{\lambda e^{-\lambda w_i}\Phi(w_i/\sigma_v + b_{1i})}{\chi_{2i}}$$

where $\lambda = (1/\sigma_u) + (1/\sigma_w)$, $\chi_{1i} = \Phi(b_{2i}) + e^{a_{1i}-a_{2i}}\Phi(b_{1i})$, and $\chi_{2i} = \Phi(b_{1i}) + e^{a_{2i}-a_{1i}}\Phi(b_{2i}) = e^{a_{2i}-a_{1i}}\chi_{1i}$.

The observation-specific conditional expectations of u_i and w_i are

$$E(u_i|\varepsilon_i) = \frac{1}{\lambda} + \frac{e^{a_{1i}-a_{2i}}\sigma_v\left\{\phi(-b_{1i}) + b_{1i}\Phi(b_{1i})\right\}}{\chi_{1i}} \tag{3}$$
$$E(w_i|\varepsilon_i) = \frac{1}{\lambda} + \frac{\sigma_v\left\{\phi(-b_{2i}) + b_{2i}\Phi(b_{2i})\right\}}{\chi_{1i}} \tag{4}$$

2.1.3 Measuring one-sided terms in logarithmic specification

Following Papadopoulos (2018), if the dependent variable enters the regression in logarithmic form, we need to consider the expected values of the exponentiated variables.

We derive the following conditional expectations to obtain the logarithmic one-sided terms, which are e^{-u} and e^{-w}. The interpretation of these one-sided terms will be discussed later in section 2.3.

$$E(e^{-u_i}|\varepsilon_i) = \frac{\lambda}{1+\lambda}\frac{1}{\chi_{1i}}\left\{\Phi(b_{2i}) + e^{a_{1i}-a_{2i}} \times e^{\sigma_v^2/2-\sigma_v b_{1i}}\Phi(b_{1i}-\sigma_v)\right\} \tag{5}$$
$$E(e^{-w_i}|\varepsilon_i) = \frac{\lambda}{1+\lambda}\frac{1}{\chi_{2i}}\left\{\Phi(b_{1i}) + e^{a_{2i}-a_{1i}} \times e^{\sigma_v^2/2-\sigma_v b_{2i}}\Phi(b_{2i}-\sigma_v)\right\} \tag{6}$$

The relative measure of one-sided terms $E(e^{w_i}e^{-u_i}|\varepsilon_i)$ is

$$E(e^{w_i}e^{-u_i}|\varepsilon_i) = \frac{e^{(1+\sigma_u)\left(a_{1i}+\frac{\sigma_v^2}{2\sigma_u}\right)}\Phi(b_{1i}-\sigma_v) + e^{(1-\sigma_w)\left(a_{2i}-\frac{\sigma_v^2}{2\sigma_w}\right)}\Phi(b_{2i}+\sigma_v)}{e^{a_{1i}}\Phi(b_{1i}) + e^{a_{2i}}\Phi(b_{2i})} \tag{7}$$

2.2 The half-normal specification

2.2.1 Model estimation

As in Papadopoulos (2015), we now assume that the inefficiency terms u_i and w_i follow half-normal distributions,

$$\begin{aligned} v_i &\sim \text{i.i.d. } \mathcal{N}(0, \sigma_v^2) \\ u_i &\sim \text{i.i.d. } \mathcal{N}_+(0, \sigma_u^2) \\ w_i &\sim \text{i.i.d. } \mathcal{N}_+(0, \sigma_w^2) \end{aligned}$$

where $\mathcal{N}_+(0, \sigma^2)$ represents a half-normal distribution, $\mathcal{N}_+(0, \sigma^2) = |\mathcal{N}(0, \sigma^2)|$.

For compactness, we use the notations

$$\omega_1 \equiv \frac{s\sqrt{1+\theta_2^2}}{\theta_1}, \quad \omega_2 \equiv \frac{s\sqrt{1+\theta_1^2}}{\theta_2}, \quad \lambda_1 \equiv \frac{\theta_2}{\theta_1}\sqrt{1+\theta_1^2+\theta_2^2}, \quad \lambda_2 \equiv \frac{\theta_1}{\theta_2}\sqrt{1+\theta_1^2+\theta_2^2}$$

where $\theta_1 \equiv (\sigma_w/\sigma_v)$, $\theta_2 \equiv (\sigma_u/\sigma_v)$, and $s \equiv \sqrt{\sigma_v^2 + \sigma_u^2 + \sigma_w^2} = \sigma_v\sqrt{1+\theta_1^2+\theta_2^2}$.

With these notations, the density of ε_i is

$$f_\varepsilon(\varepsilon_i) = \frac{2}{s}\phi\left(\varepsilon_i/s\right)\left\{G(\varepsilon_i; 0, \omega_1, -\lambda_1) - G(\varepsilon_i; 0, \omega_2, \lambda_2)\right\}$$

where $G(z; \text{location}, \text{scale}, \text{skew})$ is the CDF of a univariate skew-normal random variable. We further use

$$G_{1i} \equiv G(\varepsilon_i; 0, \omega_1, -\lambda_1), \quad G_{2i} \equiv G(\varepsilon_i; 0, \omega_2, \lambda_2)$$

For convenience of empirical implementation, we evaluate the CDF of the skew-normal distribution with the correlated bivariate standard normal CDF, Φ_2, following Papadopoulos (2018):

$$G(\varepsilon_i; \xi, \omega, \lambda) = 2\Phi_2\left(\frac{\varepsilon_i - \xi}{\omega}, 0; \rho = \frac{-\lambda}{\sqrt{1+\lambda^2}}\right)$$

The corresponding log-likelihood function is

$$\ln L(\boldsymbol{\varepsilon}|\mathbf{y}, \mathbf{X}, \mathbf{q}) = n\ln\left(\frac{2}{\sqrt{2\pi}}\right) - n\ln s - \frac{1}{2s^2}\sum_{i=1}^{n}(y_i - \mathbf{x}_i'\boldsymbol{\delta})^2 + \sum_{i=1}^{n}\ln(G_{1i} - G_{2i})$$

where $\mathbf{q} = \left(\boldsymbol{\delta}', s, \theta_1, \theta_2\right)'$. $\mathbf{x}_i$ is the column vector taken from the ith row of $\mathbf{X}$.

2.2.2 Measuring one-sided terms in levels

As for the measurement of inefficiency, the conditional expected values of u_i and w_i are

$$E(u_i|\varepsilon_i) = s^2\psi_{2i} - \frac{\sigma_u^2}{s^2}(\varepsilon_i - s^2\psi_i) \tag{8}$$

$$E(w_i|\varepsilon_i) = s^2\psi_{1i} + \frac{\sigma_w^2}{s^2}(\varepsilon_i - s^2\psi_i) \tag{9}$$

where $\psi_{1i} \equiv \{g_{1i}/(G_{1i} - G_{2i})\}$, $\psi_{2i} \equiv \{g_{2i}/(G_{1i} - G_{2i})\}$, $\psi_i \equiv \psi_{1i} - \psi_{2i}$, and $g_{\cdot i}$ is the probability density function of the corresponding skew-normal distribution.

2.2.3 Measuring one-sided terms in logarithmic specification

The logarithmic inefficiency can be estimated by $1 - E(e^{-u_i}|\varepsilon_i)$ and $1 - E(e^{-w_i}|\varepsilon_i)$, where

$$E(e^{-u}|\varepsilon_i) = 2\left(G_{1i} - G_{2i}\right)^{-1} \exp\left(\frac{1}{2}\omega_u^2 + \frac{\omega_u}{\omega_2}\varepsilon_i\right)\left\{\Phi\left(\frac{\varepsilon_i - \sigma_u^2}{\omega_1}\right)\right.$$
$$\left. - \Phi_2\left(\frac{\varepsilon_i - \sigma_u^2}{\omega_1}, \omega_u + \frac{\varepsilon_i}{\omega_2}; \rho = \frac{-\sigma_w\sigma_u}{s_1 s_2}\right)\right\} \tag{10}$$

$$E(e^{-w}|\varepsilon_i) = 2\left(G_{1i} - G_{2i}\right)^{-1} \exp\left(\frac{1}{2}\omega_w^2 + \frac{\omega_w}{\omega_1}\varepsilon_i\right)\left\{\Phi\left(-\frac{\varepsilon_i + \sigma_w^2}{\omega_2}\right)\right.$$
$$\left. - \Phi_2\left(-\frac{\varepsilon_i + \sigma_w^2}{\omega_2}, \omega_w - \frac{\varepsilon_i}{\omega_1}; \rho = \frac{-\sigma_w\sigma_u}{s_1 s_2}\right)\right\} \tag{11}$$

in which $s_1 \equiv \sqrt{\sigma_w^2 + \sigma_v^2}$, $s_2 \equiv \sqrt{\sigma_u^2 + \sigma_v^2}$, $\omega_w \equiv \{(\sigma_w s_2)/s\}$, $\omega_u \equiv \{(\sigma_u s_1)/s\}$, and $\Phi_2(\cdot)$ is the correlated bivariate standard normal CDF.

$E(e^{w_i}e^{-u_i}|\varepsilon_i)$ can be obtained as

$$E(e^{w_i}e^{-u_i}|\varepsilon_i) = \exp\left\{\frac{\sigma_w^2 + \sigma_u^2}{s^2}\left(\frac{\sigma_v^2}{2} + \varepsilon_i\right)\right\}$$
$$\times \frac{\Phi_2\left(\frac{\varepsilon_i + \sigma_v^2}{\omega_1}, 0; \rho = \frac{\lambda_1}{\sqrt{1+\lambda_1^2}}\right) - \Phi_2\left(\frac{\varepsilon_i + \sigma_v^2}{\omega_2}, 0; \rho = \frac{-\lambda_2}{\sqrt{1+\lambda_2^2}}\right)}{\Phi_2\left(\frac{\varepsilon_i}{\omega_1}, 0; \rho = \frac{\lambda_1}{\sqrt{1+\lambda_1^2}}\right) - \Phi_2\left(\frac{\varepsilon_i}{\omega_2}, 0; \rho = \frac{-\lambda_2}{\sqrt{1+\lambda_2^2}}\right)} \tag{12}$$

2.3 Interpreting the one-sided terms

Interpreting the one-sided terms is an essential part of 2TSF analysis. To study welfare allocation, researchers usually pay more attention to the relative measures of one-sided terms, which measure the result of price bargaining or information asymmetry.

For example, we introduce the interpretation of one-sided terms in a logarithmic price bargaining model. The actual price is assumed to be $F(\mathbf{x})e^{v-u+w}$, where $F(\mathbf{x}) = e^{\mathbf{x}'\boldsymbol{\delta}}$ is the optimal price. Correspondingly, the benchmark price $F(\mathbf{x})e^{v}$ is the optimal price plus a stochastic markup, while the maximum price is $F(\mathbf{x})e^{v+w}$ and the minimum price is $F(\mathbf{x})e^{v-u}$. The relative measure of surplus can be evaluated as follows:

- $1 - e^{-u}$ measures two things that appear different but are equal in magnitude:
  - Seller surplus with respect to the benchmark price, relative to the benchmark price.

$$\frac{\text{Benchmark price} - \text{Minimum price}}{\text{Benchmark price}} = \frac{F(\mathbf{x})e^{v} - F(\mathbf{x})e^{v-u}}{F(\mathbf{x})e^{v}} = 1 - e^{-u}$$

  - Consumer surplus with respect to the actual price, relative to the maximum price.

$$\frac{\text{Maximum price} - \text{Actual price}}{\text{Maximum price}} = \frac{F(\mathbf{x})e^{v+w} - F(\mathbf{x})e^{v-u+w}}{F(\mathbf{x})e^{v+w}} = 1 - e^{-u}$$

- $1 - e^{-w}$ measures two things that appear different but are equal in magnitude:
  - Seller surplus with respect to the actual price, relative to the actual price.

$$\frac{\text{Actual price} - \text{Minimum price}}{\text{Actual price}} = \frac{F(\mathbf{x})e^{v-u+w} - F(\mathbf{x})e^{v-u}}{F(\mathbf{x})e^{v-u+w}} = 1 - e^{-w}$$

  - Consumer surplus with respect to the benchmark price, relative to the maximum price.

$$\frac{\text{Maximum price} - \text{Benchmark price}}{\text{Maximum price}} = \frac{F(\mathbf{x})e^{v+w} - F(\mathbf{x})e^{v}}{F(\mathbf{x})e^{v+w}} = 1 - e^{-w}$$

- $e^{-w} - e^{-u}$ measures the net gain in consumer surplus, which is actually the deviation of actual price from benchmark price as a percentage of maximum price.

$$\frac{\text{Benchmark price} - \text{Actual price}}{\text{Maximum price}} = \frac{F(\mathbf{x})e^{v} - F(\mathbf{x})e^{v-u+w}}{F(\mathbf{x})e^{v+w}} = e^{-w} - e^{-u}$$

- $e^{w}e^{-u} - 1$ measures the difference between actual price from benchmark price, as a percentage of benchmark price.

$$\frac{\text{Actual price} - \text{Benchmark price}}{\text{Benchmark price}} = \frac{F(\mathbf{x})e^{v-u+w} - F(\mathbf{x})e^{v}}{F(\mathbf{x})e^{v}} = e^{w}e^{-u} - 1$$

3 2TSF models with the scaling property

Assuming that the one-sided error component is from a one-parameter distribution, it possesses the scaling property (Parmeter 2018), which has many benefits. The scaling property enables the estimation of 2TSF models without distributional assumptions (Wang and Schmidt 2002). Further, as noted by Alvarez et al. (2006), ease of interpretation of estimates also stems from use of the scaling property. Finally, as noted in Parmeter (2018), independence among the separate error terms is no longer necessary for identification and estimation when the scaling property is invoked. This point is essential because, in many empirical settings, it is likely that u and w will have some type of dependence at a minimum.

The ability to avoid distributional assumptions is naturally favorable, although the imposition of correct distributional assumptions will produce statistically efficient estimators via maximum likelihood. Empiricists, however, may find the simplicity of NLS more palatable and, as such, argue for the imposition of the scaling property. `sftt` also offers users an option to fit the 2TSF model with the scaling property.

Starting with (1),

$$\mathbf{y} = \mathbf{X}\boldsymbol{\delta} - \mathbf{u} + \mathbf{w} + \mathbf{v} \tag{13}$$

assuming that the distributions of u_i and w_i depend upon the level of observable characteristics $\mathbf{z_u}_i$ and $\mathbf{z_w}_i$, which are vectors of characteristics for the ith observation. We then introduce the scaling property into the 2TSF model:

$$u_i = u(\mathbf{z_u}_i, \boldsymbol{\delta}_\mathbf{u}) = g_u(\mathbf{z_u}_i, \boldsymbol{\delta}_\mathbf{u}) \times u_i^*$$
$$w_i = u(\mathbf{z_w}_i, \boldsymbol{\delta}_\mathbf{w}) = g_w(\mathbf{z_w}_i, \boldsymbol{\delta}_\mathbf{w}) \times w_i^*$$

u_i^* and w_i^* are from what are termed basic distributions and are independent from $\mathbf{x}_i$, $\mathbf{z_u}_i$, and $\mathbf{z_w}_i$. $g_u(\cdot)$ and $g_w(\cdot)$ are scaling functions: $g_u(\cdot) \geq 0$ and $g_w(\cdot) \geq 0$. Following Parmeter (2018), we assume $g(\mathbf{z}_i, \boldsymbol{\delta}) = e^{\mathbf{z}_i'\boldsymbol{\delta}}$.

To fit the model, rewrite (13) as

$$y_i = \mathbf{x}_i'\boldsymbol{\delta} - e^{\mathbf{z_u}_i'\boldsymbol{\delta}_\mathbf{u}} u_i^* + e^{\mathbf{z_w}_i'\boldsymbol{\delta}_\mathbf{w}} w_i^* + v_i \tag{14}$$

Taking the expectation of (14),

$$E\left(y_i|\mathbf{x}_i, \mathbf{z_u}_i, \mathbf{z_w}_i\right) = \mathbf{x}_i'\boldsymbol{\delta} - \mu_u^* e^{\mathbf{z_u}_i'\boldsymbol{\delta}_\mathbf{u}} + \mu_w^* e^{\mathbf{z_w}_i'\boldsymbol{\delta}_\mathbf{w}} \tag{15}$$

where δ, δ_u, δ_w, μ_u, and μ_w are parameters; μ_u^* and μ_w^* are the expectations of u_i^* and w_i^*; $\mu_u^* = E(u_i^*)$; and $\mu_w^* = E(w_i^*)$. Because u_i^* and w_i^* are independent from $\mathbf{z_u}_i$ and $\mathbf{z_w}_i$, we can extract the mean of u_i^* and w_i^* from $e^{\mathbf{z_u}_i'\boldsymbol{\delta}_\mathbf{u}} u_i^*$ and $e^{\mathbf{z_w}_i'\boldsymbol{\delta}_\mathbf{w}} w_i^*$.

To fit this model, we use NLS estimation:

$$\left(\widehat{\boldsymbol{\delta}}, \widehat{\boldsymbol{\delta}}_\mathbf{u}, \widehat{\boldsymbol{\delta}}_\mathbf{w}, \widehat{\mu}_u^*, \widehat{\mu}_w^*\right) = \min_{\boldsymbol{\delta}, \boldsymbol{\delta}_\mathbf{u}, \boldsymbol{\delta}_\mathbf{w}, \mu_u^*, \mu_w^*} n^{-1} \sum_{i=1}^{n} \left(y_i - \mathbf{x}_i'\boldsymbol{\delta} + \mu_u^* e^{\mathbf{z_u}_i'\boldsymbol{\delta}_\mathbf{u}} - \mu_w^* e^{\mathbf{z_w}_i'\boldsymbol{\delta}_\mathbf{w}}\right)^2$$

4 The sftt command

This section introduces the syntax of `sftt`.

4.1 Estimation syntax

The syntax to fit a 2TSF model with distributional assumptions following Kumbhakar and Parmeter (2009) and Papadopoulos (2015) is

`sftt` *depvar* [*indepvars*] [*if*] [*in*] [`,` `hnormal` `noconstant` `robust` `vce(`*vcetype*`)` `findseed` `seed(`*#*`)` `sigmau(`*varlist*`)` `sigmaw(`*varlist*`)` `iterate(`*#*`)`]

The syntax to fit a 2TSF model with the scaling property following Parmeter (2018) is

`sftt` *depvar* *indepvars* [*if*] [*in*] [`,` `noconstant` `robust` `vce(`*vcetype*`)` `sigmau(`*varlist*`)` `sigmaw(`*varlist*`)` `iterate(`*#*`)` `scaling` `initial(`*initial_values*`)`]

indepvars may contain factor variables; see [U] **11.4.3 Factor variables**.

4.2 Options for sftt

`hnormal` uses the half-normal/half-normal/normal specification rather than the benchmark exponential/exponential/normal specification. Note that, when using this option, the estimation might not converge because of flat derivatives or missing values; setting another random seed by using the option `seed()` might help. Users may also specify the `findseed` option to find a usable random seed.

`noconstant` suppresses the constant term (intercept) in the linear model.

`robust` is the synonym for `vce(robust)`.

`vce(`*vcetype*`)` specifies the type of standard error reported, which includes types that are robust to some kinds of misspecification (`robust`), allow for intragroup correlation (`cluster` *clustvar*), or use bootstrap or jackknife methods (`bootstrap`, `jackknife`) for estimation with `scaling`.

`findseed` loops through 100 estimations, during which the random seed was set from 1 to 100. We iterate at most 200 times for each seed to accelerate the process.

`seed(`*#*`)` sets a random seed before estimating to ensure that the results are reproducible.

`sigmau(`*varlist*`)` specifies heteroskedasticity in the negative inefficiency component, with the variance expressed as a linear model of the covariates defined in *varlist*.

`sigmaw(`*varlist*`)` specifies heteroskedasticity in the positive inefficiency component, with the variance expressed as a linear model of the covariates defined in *varlist*.

`iterate(`*#*`)` specifies the maximum iterations. The default is `iterate(1000)`. In most cases, the optimization should converge in fewer than 1,000 iterations.

`scaling` fits the 2TSF model with the scaling property by NLS. The results might be very sensitive to the initial values if the models are complex.

`initial(`*initial_values*`)` specifies the initial values to begin the NLS estimation. This option is used only when estimating with `scaling`. When NLS runs slowly or cannot converge, assigning initial values might help. If the independent variable is named `x` and the covariates for the two one-sided error terms are `zu` and `zw`, then the initial values should be assigned with the syntax `initial(delta_x 1 du_zu 0.6 mu_u 1 dw_zw 0.8 mu_w 1)`, where `mu_u` and `mu_w` represent μ_u^* and μ_w^* in (15) and the numbers correspond to initial values. By default, the estimation results of ordinary least squares (OLS) would be used as initial values for dependent variables, and other parameters would be initialized to 1.

4.3 Subcommands

After model estimation using `sftt`, the subcommands `sftt sigs` and `sftt eff` may be used for variance decomposition and calculation of inefficiency measures, respectively.

Error term decomposition

`sftt sigs`

This command calculates the parameters of each component's distribution (u_i, w_i, and v_i) in the composite error term.

Inefficiency measurements

`sftt eff` [`,` <u>`lev`</u>`el exp` <u>`abs`</u>`olute` <u>`rel`</u>`ative u_hat(`*newvar*`) w_hat(`*newvar*`)`
`wu_diff(`*newvar*`) u_hat_exp(`*newvar*`) w_hat_exp(`*newvar*`)`
`wu_diff_exp(`*newvar*`) wu_net_effect(`*newvar*`) replace`]

This subcommand encapsulates several of the most commonly used algorithms of inefficiency in both level and logarithmic specification. By default, this command will generate all measures of inefficiency using the default variable name. This subcommand is unavailable when the `scaling` option is invoked. Users may use the following options to select the type of measurement to be performed:

`level` generates inefficiency terms only in the level specification. `level` may not be used with `exp`.

`exp` generates inefficiency terms only in the logarithmic specification. `exp` may not be used with `level`.

`absolute` generates only absolute measures of inefficiency. `absolute` may not be used with `relative`.

`relative` generates only relative measures of inefficiency. `relative` may not be used with `absolute`.

The variable names of inefficiency measures may be customized using the following options:

`u_hat(`*newvar*`)` sets the variable name of $E(u_i|\varepsilon_i)$, which is the conditional expectation of u_i, calculated by (3) and (8). The default is `u_hat(_u_hat)`.

`w_hat(`*newvar*`)` sets the variable name of $E(w_i|\varepsilon_i)$, which is the conditional expectation of w_i, calculated by (4) and (9). The default is `w_hat(_w_hat)`.

`wu_diff(`*newvar*`)` sets the variable name of $E(w_i|\varepsilon_i) - E(u_i|\varepsilon_i)$, which is the net surplus in the level specification. The default is `wu_diff(_wu_diff)`.

`u_hat_exp(`*newvar*`)` sets the variable name of $E(e^{-u_i}|\varepsilon_i)$, the conditional expectation of e^{-u_i}, calculated by (5) and (10). The default is `u_hat_exp(_u_hat_exp)`.

`w_hat_exp(`*newvar*`)` sets the variable name of $E(e^{-w_i}|\varepsilon_i)$, the conditional expectation of e^{-w_i}, calculated by (6) and (11). The default is `w_hat_exp(_w_hat_exp)`.

`wu_diff_exp(`*newvar*`)` sets the variable name of $E(e^{-w_i}|\varepsilon_i) - E(e^{-u_i}|\varepsilon_i)$, the net surplus in the logarithmic specification. The default is `wu_diff_exp(_wu_diff_exp)`.

`wu_net_effect(`*newvar*`)` sets the variable name of $E(e^{w_i}e^{-u_i}|\varepsilon_i) - 1$, which is the net effect in the logarithmic specification, calculated by (7) and (12). The default is `wu_net_effect(_wu_net_effect)`.

The command will check if there are any previous variables before generating a new variable. To overwrite the existing variable, use the following option:

`replace` permits `sftt` to overwrite existing variables.

5 Examples with simulated data

This section provides several examples with simulated data to illustrate the features of `sftt`.

5.1 The benchmark 2TSF model

We first illustrate the `sftt` command by fitting a 2TSF model with simulated data. As in Kumbhakar and Parmeter (2009), the one-sided error terms u_i and w_i follow exponential distributions, and the stochastic noise v_i is assumed to come from a normal distribution.

The data-generating process is

$$\begin{aligned} y_i &= x_{1i} + 2x_{2i} - u_i + w_i + v_i \\ u_i &\sim \text{i.i.d. Exp}(0.6) \\ w_i &\sim \text{i.i.d. Exp}(1.4) \\ v_i &\sim \text{i.i.d. } \mathcal{N}(0,1) \end{aligned}$$

where y_i is the outcome variable and the covariates x_{1i} and x_{2i} are normally distributed with 0 means and variances equal to 1.

We first fit a random sample with 1,600 observations by using the `sftt` command:

```
. set seed 999
. quietly set obs 1600
. generate x1 = invnormal(runiform())
. generate x2 = invnormal(runiform())
. generate ue = invexponential(0.6, runiform())
. generate we = invexponential(1.4, runiform())
. generate v = invnormal(runiform())
. generate y = x1 + 2 * x2 - ue + we + v
```

```
. sftt y x1 x2, noconstant
initial:       log likelihood = -3233.1231
rescale:       log likelihood = -3233.1231
rescale eq:    log likelihood = -3229.7914
Iteration 0:   log likelihood = -3229.7914
Iteration 1:   log likelihood = -3170.7503  (not concave)
Iteration 2:   log likelihood = -3161.2027
Iteration 3:   log likelihood = -3161.0853
Iteration 4:   log likelihood = -3161.0851
Two-tier stochastic frontier model with exponential specification
                                                       Number of obs =   1,600
                                                       Wald chi2(2)  = 3186.44
Log likelihood = -3161.0851                            Prob > chi2   =  0.0000
```

y	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
frontier_y						
x1	.9915255	.0419678	23.63	0.000	.9092702	1.073781
x2	2.032645	.0402523	50.50	0.000	1.953752	2.111538
ln_sig_v						
_cons	-.0712458	.0951567	-0.75	0.454	-.2577494	.1152579
ln_sig_u						
_cons	-.4742256	.0972085	-4.88	0.000	-.6647506	-.2837005
ln_sig_w						
_cons	.3918662	.0468007	8.37	0.000	.3001386	.4835939

From the estimation results, the coefficients of x_{1i} and x_{2i} are 0.9915 and 2.0326, which are very close to their true values (1 and 2).

To make sure the distributional parameters are strictly positive, the command takes the exponential form during estimation. Thus, the `_cons` in `sigma_v`, `sigma_u`, and `sigma_w` are actually $\ln(\sigma_v)$, $\ln(\sigma_u)$, and $\ln(\sigma_w)$. We then run the following postestimation commands to interpret the actual distributional parameters and decompose the residuals:

```
. sftt sigs
                Variance Estimation
--------------------------------------------------
sigma_v    :      0.9312
sigma_u    :      0.6224
sigma_w    :      1.4797
sigma_v_sq :      0.8672
sigma_u_sq :      0.3873
sigma_w_sq :      2.1896
--------------------------------------------------
                Variance Analysis
--------------------------------------------------
Total sigma_sqs     :  3.4442
(sigu2+sigw2)/Total :  0.7482
sigu2/(sigu2+sigw2) :  0.1503
sigw2/(sigu2+sigw2) :  0.8497
sig_w - sig_u       :  0.8574
--------------------------------------------------

. sftt eff
The following variables have been generated:
_u_hat
_w_hat
_wu_diff
_u_hat_exp
_w_hat_exp
_wu_diff_exp
_wu_net_effect

. summarize _u_hat_exp _w_hat_exp _wu_diff_exp

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
  _u_hat_exp |      1,600    .3818473    .1049245   .3046399   .9900407
  _w_hat_exp |      1,600    .6003683     .198864   .3046399   .9999939
_wu_diff_exp |      1,600     .218521    .2852138  -.6854007    .695354

. summarize _wu_diff_exp, detail

                        _wu_diff_exp
-------------------------------------------------------------
      Percentiles      Smallest
 1%    -.4977102      -.6854007
 5%    -.2708443      -.6775006
10%    -.1505466       -.644028       Obs               1,600
25%     .0229923      -.6344826       Sum of wgt.       1,600

50%     .2233657                      Mean            .218521
                        Largest       Std. dev.      .2852138
75%     .4496792       .6951633
90%     .6112484       .6952174       Variance       .0813469
95%     .6619091        .695277       Skewness       -.304613
99%     .6920454        .695354       Kurtosis       2.605813
```

The command `sftt sigs` identifies the variance of each component in the composite error term. The estimated standard errors of u_i, w_i, and v_i are 0.6224, 1.4797, and 0.9312, and the actual standard errors are 0.6, 1.4, and 1, respectively. `sftt eff` decomposes the residual and calculates the inefficiency measures.

5.2 Estimation with the scaling property

`sftt` also fits 2TSF models with the scaling property (Parmeter 2018). The most attractive feature of the scaling property is that it is free from distributional and independence assumptions. Here we assume that the basic distribution of the one-sided error terms is the exponential distribution. Consider the data-generating process

$$\begin{aligned} y_i &= x_i - e^{0.6 z_{ui}} \times u_i^* + e^{0.8 z_{wi}} \times w_i^* + v_i \\ u_i^* &\sim \text{i.i.d. Exp}(1) \\ w_i^* &\sim \text{i.i.d. Exp}(1) \\ v_i &\sim \text{i.i.d. } \mathcal{N}(0,1) \end{aligned}$$

where x_i, z_{ui}, and z_{wi} are the covariates. As in Parmeter (2018), x_i, z_{ui}, and z_{wi} follow a trivariate normal distribution with correlation 0.1. u_i^*, w_i^*, and v_i are mutually independent from one another and from (x_i, z_{ui}, z_{wi}). We do not explicitly include an intercept in this model, because the simulations in Parmeter (2018) suggested it was quite difficult to separately identify an intercept and the constant terms for both u and w.

The following commands demonstrate the generation of the simulated data and the way to fit the model. We first estimate without setting initial values.

```
. clear
. set seed 999
. quietly set obs 10000
. matrix C = (1, 0.1, 0.1 \ 0.1, 1, 0.1 \ 0.1, 0.1, 1)
. drawnorm x zu zw, corr(C)
. generate ui = invexponential(1, runiform())
. generate wi = invexponential(1, runiform())
. generate vi = invnormal(runiform())
. generate y = x - exp(0.6 * zu) * ui + exp(0.8 * zw) * wi + vi
```

```
. sftt y x, scaling sigmau(zu) sigmaw(zw) robust noconstant
initial value: delta_x 1.056823117155457 du_zu 1 mu_u 1 dw_zw 1 mu_w 1
Iteration 0:  residual SS =  70216.84
Iteration 1:  residual SS =  68805.46
Iteration 2:  residual SS =  68791.96
Iteration 3:  residual SS =  68791.83
Iteration 4:  residual SS =  68791.83
Iteration 5:  residual SS =  68791.83

Nonlinear regression                              Number of obs =     10,000
                                                  R-squared     =     0.3221
                                                  Adj R-squared =     0.3217
                                                  Root MSE      =   2.623476
                                                  Res. dev.     =   47663.77
Two-tier stochastic frontier model with scaling property
------------------------------------------------------------------------------
             |               Robust
           y | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    /delta_x |   1.004185    .025706    39.06   0.000     .9537958    1.054574
      /du_zu |   .5945575   .0549976    10.81   0.000     .4867511    .7023639
       /mu_u |   1.013727   .1194435     8.49   0.000     .7795938     1.24786
      /dw_zw |    .777372   .0684228    11.36   0.000     .6432495    .9114945
       /mu_w |   1.030119   .1314908     7.83   0.000     .7723707    1.287868
------------------------------------------------------------------------------
```

As expected, the coefficients are precisely estimated.

In the above example, we did not provide any initial value. The `sftt` command will run a `regress` command to specify a set of initial values for `delta_x`, while the other parameters are initialized to 1.

A suitable set of initial values can speed up the optimization process. Here we use the actual values of these parameters to emphasize the comparison.

```
. sftt y x, scaling sigmau(zu) sigmaw(zw) robust noconstant
>           initial(delta_x 1 du_zu 0.6 mu_u 1 dw_zw 0.8 mu_w 1)
initial value: delta_x 1 du_zu 0.6 mu_u 1 dw_zw 0.8 mu_w 1

Iteration 0:  residual SS =  68791.84
Iteration 1:  residual SS =  68791.83
Iteration 2:  residual SS =  68791.83
Iteration 3:  residual SS =  68791.83

Nonlinear regression                                Number of obs =     10,000
                                                    R-squared     =     0.3221
                                                    Adj R-squared =     0.3217
                                                    Root MSE      =   2.623476
                                                    Res. dev.     =   47663.77

Two-tier stochastic frontier model with scaling property

------------------------------------------------------------------------------
             |               Robust
           y | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    /delta_x |   1.004185    .025706    39.06   0.000     .9537958    1.054574
       /du_zu |   .5945575   .0549976    10.81   0.000      .486751    .7023639
       /mu_u |   1.013727   .1194433     8.49   0.000     .7795941     1.24786
       /dw_zw |    .777372   .0684228    11.36   0.000     .6432495    .9114944
       /mu_w |   1.030119   .1314907     7.83   0.000      .772371    1.287867
------------------------------------------------------------------------------
```

From the results, we can see that the number of iterations is reduced, but the estimation results are identical.

We demonstrate the performance of `sftt` with a series of Monte Carlo simulations, and the results show a similar pattern to table 1 in Parmeter (2018).

6 Empirical applications

This section provides three empirical applications to illustrate the use of `sftt`. We first examine the wage bargaining between firms and workers by using the benchmark exponential specification following Kumbhakar and Parmeter (2009). We then estimate the hedonic price function in Kumbhakar and Parmeter (2010) by using a heterogeneous 2TSF model with half-normal specification. Finally, we compare different specifications' impacts on inefficiency by using a dataset from the medical service market following Lu, Lian, and Lu (2011).

6.1 Match uncertainty and wage bargaining

In labor markets, workers' willingness to accept a job and firms' willingness to offer a job are private information; workers and firms are incentivized to extract more surplus from each other. We can demonstrate the optimal wage and surplus extraction in each firm–worker pair in a two-tier stochastic frontier model:

$$\text{wage}_i = \mu(\mathbf{x}_i) + \varepsilon_i$$

wage_i is the actual wage, and $\mu(\mathbf{x}_i) = \mathbf{x}_i'\boldsymbol{\delta}$ is a linear model that represents the optimal wage in the ith firm–worker pair. $\varepsilon_i = v_i - u_i + w_i$ is the composite error term. The lower frontier of wage (wage_i) is $\underline{\text{wage}}_i = \mu(\mathbf{x}_i) - u_i$, which represents the minimum wage that the ith worker is willing to accept. The upper frontier indicates the maximum wage that a firm would pay to hire a worker and is given by $\overline{\text{wage}}_i = \mu(\mathbf{x}_i) + w_i$. The worker's surplus is $\mu(\mathbf{x}_i) - \underline{\text{wage}}_i$, and the firm's surplus is $\overline{\text{wage}}_i - \mu(\mathbf{x}_i)$.

Distributional assumptions are then proposed to make it possible to estimate the surplus extraction components. Following Kumbhakar and Parmeter (2009), we assume that $v_i \sim$ i.i.d. $\mathcal{N}(0, 1)$, $u_i \sim$ i.i.d. $\text{Exp}(\sigma_u)$, and $w_i \sim$ i.i.d. $\text{Exp}(\sigma_w)$.

Run the following syntax:

```
. set seed 20220612
. use https://sftt.oss-cn-hangzhou.aliyuncs.com/kp09.dta, clear
. sftt lwage iq educ educ2 exper exper2 tenure tenure2 age married south
>          urban black sibs brthord meduc feduc
initial:       log likelihood = -821.98656
rescale:       log likelihood = -821.98656
rescale eq:    log likelihood = -821.98656
Iteration 0:   log likelihood = -821.98656  (not concave)
Iteration 1:   log likelihood =   -790.073  (not concave)
  (output omitted)
Iteration 15:  log likelihood = -226.06913
Two-tier stochastic frontier model with exponential specification
                                                Number of obs =    663
                                                Wald chi2(16) = 319.41
Log likelihood = -226.06913                     Prob > chi2   = 0.0000

              lwage   Coefficient  Std. err.      z    P>|z|     [95% conf. interval]

frontier_lw~e
                 iq     .0043955   .0011061     3.97   0.000     .0022277    .0065634
               educ     2.036803   1.128628     1.80   0.071    -.1752675    4.248873
              educ2    -.7563295   .5533698    -1.37   0.172    -1.840914    .3282555
              exper       .28234   .1561886     1.81   0.071     -.023784     .588464
             exper2    -.1005302   .0921451    -1.09   0.275    -.2811314    .0800709
             tenure     .1535866   .0635696     2.42   0.016     .0289925    .2781808
            tenure2    -.0568321   .0380829    -1.49   0.136    -.1314732    .0178091
                age     .4950143   .1867546     2.65   0.008      .128982    .8610465
            married     .2060007   .0443427     4.65   0.000     .1190905    .2929109
              south    -.0348172   .0297127    -1.17   0.241    -.0930529    .0234186
              urban     .2208797   .0290555     7.60   0.000     .1639319    .2778275
              black    -.1022164   .0526949    -1.94   0.052    -.2054965    .0010637
               sibs     .0088167    .007232     1.22   0.223    -.0053579    .0229912
            brthord     -.015372   .0106093    -1.45   0.147    -.0361658    .0054218
              meduc     .0081452   .0056126     1.45   0.147    -.0028553    .0191456
              feduc     .0076534   .0050712     1.51   0.131    -.0022858    .0175927
              _cons     3.858254   .6128944     6.30   0.000     2.657004    5.059505

ln_sig_v
              _cons    -1.659804   .1570302   -10.57   0.000    -1.967578   -1.352031

ln_sig_u
              _cons    -1.510707   .1146933   -13.17   0.000    -1.735502   -1.285912

ln_sig_w
              _cons    -1.665971   .1230849   -13.54   0.000    -1.907213   -1.424729
```

```
. sftt sigs
                Variance Estimation
--------------------------------------------------
sigma_v    :       0.1902
sigma_u    :       0.2208
sigma_w    :       0.1890
sigma_v_sq :       0.0362
sigma_u_sq :       0.0487
sigma_w_sq :       0.0357
--------------------------------------------------
                Variance Analysis
--------------------------------------------------
Total sigma_sqs      :  0.1206
(sigu2+sigw2)/Total :  0.7002
sigu2/(sigu2+sigw2) :  0.5770
sigw2/(sigu2+sigw2) :  0.4230
sig_w - sig_u       : -0.0317
--------------------------------------------------
```

From the results in `Variance Analysis`, the unexplained variation in log wage ($\sigma_v^2 + \sigma_u^2 + \sigma_e^2$) is 0.1206, while 70.02% of the unexplained variation is due to bargaining. From the estimate of $E(w_i - u_i) = \sigma_w - \sigma_u$, we can tell whether bargaining affects wages on average, and if so, in which direction. In this application, $E(w_i - u_i) = -3.17\% < 0$ means bargaining may lead to lower wages on average.

However, if the interest is to obtain the exact impact of bargaining on wages, we should analyze observation-specific estimates of $E(u_i|\boldsymbol{\varepsilon})$ and $E(w_i|\boldsymbol{\varepsilon})$. As an example, we analyze the surplus extraction between races (table 4 of Kumbhakar and Parmeter [2009]) with the `sftt eff` command.

```
. sftt eff
The following variables have been generated:
_u_hat
_w_hat
_wu_diff
_u_hat_exp
_w_hat_exp
_wu_diff_exp
_wu_net_effect
```

```
. tabstat _w_hat _u_hat _wu_diff, by(black) statistics(mean p25 p50 p75)
> format(%6.3f) columns(statistics)

Summary for variables: _w_hat _u_hat _wu_diff
Group variable: black

   black |      Mean       p25       p50       p75
---------+----------------------------------------
       0 |     0.189     0.114     0.143     0.212
         |     0.221     0.122     0.165     0.249
         |    -0.032    -0.135    -0.022     0.091
---------+----------------------------------------
       1 |     0.184     0.110     0.150     0.210
         |     0.213     0.122     0.156     0.279
         |    -0.029    -0.169    -0.006     0.088
---------+----------------------------------------
   Total |     0.189     0.114     0.144     0.212
         |     0.221     0.122     0.164     0.251
         |    -0.032    -0.137    -0.020     0.091
--------------------------------------------------

. tabstat _w_hat_exp _u_hat_exp _wu_diff_exp, by(black)
> statistics(mean p25 p50 p75) format(%6.3f) columns(statistics)

Summary for variables: _w_hat_exp _u_hat_exp _wu_diff_exp
Group variable: black

   black |      Mean       p25       p50       p75
---------+----------------------------------------
       0 |     0.159     0.103     0.127     0.182
         |     0.181     0.109     0.144     0.210
         |    -0.022    -0.107    -0.017     0.073
---------+----------------------------------------
       1 |     0.157     0.100     0.133     0.180
         |     0.178     0.110     0.138     0.232
         |    -0.021    -0.132    -0.005     0.071
---------+----------------------------------------
   Total |     0.159     0.103     0.128     0.182
         |     0.181     0.109     0.144     0.211
         |    -0.022    -0.109    -0.016     0.073
--------------------------------------------------
```

From the results, we can conclude that the difference between the average surplus extractions of Black and White workers is insignificant. However, from the tails of the extraction distributions, the lower quartile suggests that Black workers have 2–3% more extraction from the benchmark. In contrast, White workers can extract about 1% more than Black workers in the upper quartile.

6.2 Estimation of the hedonic price function

With the development of technology, information conduits such as advertisements and the Internet are increasingly ubiquitous in today's marketplace. Both buyers and sellers can gain information through search. However, given that search costs exist, market participants most likely will not become fully informed and price variations due to ignorance will exist even after controlling for product characteristics (Kumbhakar and Parmeter 2010).

As an example, the housing market might be inefficient for certain types of buyers and sellers. We can use the 2TSF model to study the advantage a buyer may possess over a seller or vice versa, simultaneously. Following Kumbhakar and Parmeter (2010), we express prices observed in the market as

$$\mathbf{P}_m = h(\mathbf{X}) + \mathbf{v} - \mathbf{u} + \mathbf{w}$$

where $h(\mathbf{X})$ is the implied price of the characteristics $\mathbf{X}$, $\mathbf{u}$ and $\mathbf{w}$ are the costs of incomplete information to the sellers and buyers, and $\mathbf{v}$ is a vector of the random noise.

To capture the heterogeneous cost of incomplete information, we allow the distributions of $\mathbf{u}$ and $\mathbf{w}$ to be functions of buyers' and sellers' characteristics, $\mathbf{Z_u}$ and $\mathbf{Z_w}$, respectively. Thus, we specify the vectors of standard errors of $\mathbf{u}$ and $\mathbf{w}$, which are $\boldsymbol{\sigma}_\mathbf{u}$ and $\boldsymbol{\sigma}_\mathbf{w}$, as

$$\boldsymbol{\sigma}_\mathbf{u} = e^{\mathbf{Z_u}\boldsymbol{\delta}_\mathbf{u}}, \quad \boldsymbol{\sigma}_w = e^{\mathbf{Z_w}\boldsymbol{\delta}_\mathbf{w}}$$

We introduce buyers' and sellers' attributes through the options `sigmau()` and `sigmaw()`. Here we assume the ignorance of information follows a half-normal distribution, so we add the option `hnormal`. The syntax and results are as follows:

```
. use https://sftt.oss-cn-hangzhou.aliyuncs.com/kp10.dta, clear
. sftt lprn lsf unitsftc bathstot roomsn sfan sfdn
> agelt5 age510 age1015 agegte30
> cencityn urbsubn urbann riuraln inadeq degreen
> s87 s88 s89 s90 s91 s92 s93
> verylg large siz1to3 small,
> sigmaw(outbuy firstbuy incbuy busbuy agebuy blkbuy marbuy sfbuy edubuy kidbuy)
> sigmau(incsell bussell agesell blksell marsell sfsell edusell kidsell)
> hnormal seed(6)
initial:       log likelihood = -5527.2566
rescale:       log likelihood = -5527.2566
rescale eq:    log likelihood = -5411.5646
Iteration 0:   log likelihood = -5411.5646  (not concave)
Iteration 1:   log likelihood = -4377.3679
  (output omitted)
Iteration 26:  log likelihood = -2846.9465
```

```
Two-tier stochastic frontier model with half-normal specification
                                                    Number of obs =   4,962
                                                    Wald chi2(27) = 2215.50
Log likelihood = -2846.9465                         Prob > chi2   =  0.0000

------------------------------------------------------------------------------
        lprn | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
frontier_lprn|
         lsf |   .2974248   .0246532    12.06   0.000     .2491053    .3457442
    unitsftc |  -.1209992   .0424964    -2.85   0.004    -.2042905   -.0377078
    bathstot |   .2234376   .0154124    14.50   0.000     .1932298    .2536454
      roomsn |   .0100356   .0062459     1.61   0.108    -.0022062    .0222774
        sfan |   .5028812   .0473966    10.61   0.000     .4099856    .5957769
        sfdn |   .5890742   .0349242    16.87   0.000     .5206241    .6575243
      agelt5 |    .196477   .0299935     6.55   0.000     .1376909    .2552631
      age510 |   .0916725   .0242236     3.78   0.000     .0441951    .1391499
     age1015 |   .0365939   .0237961     1.54   0.124    -.0100456    .0832333
    agegte30 |  -.0010998   .0201294    -0.05   0.956    -.0405527    .0383532
    cencityn |  -.1176012   .0298582    -3.94   0.000    -.1761222   -.0590803
     urbsubn |  -.0234166   .0275902    -0.85   0.396    -.0774924    .0306592
      urbann |  -.3236883   .0336743    -9.61   0.000    -.3896887   -.2576879
     riuraln |  -.3196333   .0312238   -10.24   0.000    -.3808308   -.2584358
      inadeq |   .0445825   .0669426     0.67   0.505    -.0866226    .1757877
     degreen |  -.0093918   .0057882    -1.62   0.105    -.0207364    .0019528
         s87 |   .0408672   .0285615     1.43   0.152    -.0151123    .0968467
         s88 |   .0559373   .0281587     1.99   0.047     .0007473    .1111273
         s89 |   .1140626   .0288774     3.95   0.000     .0574638    .1706613
         s90 |   .1885622   .0294923     6.39   0.000     .1307583    .2463661
         s91 |   .1295755   .0296901     4.36   0.000     .0713839    .1877671
         s92 |   .1351328   .0285183     4.74   0.000      .079238    .1910276
         s93 |   .1557371   .0304917     5.11   0.000     .0959744    .2154998
      verylg |   .5600216   .0376576    14.87   0.000     .4862141    .6338291
       large |   .1516597   .0362234     4.19   0.000     .0806632    .2226563
     siz1to3 |   .1975842   .0259854     7.60   0.000     .1466538    .2485146
       small |  -.0954866   .0270787    -3.53   0.000    -.1485599   -.0424133
       _cons |   8.171205   .1780178    45.90   0.000     7.822296    8.520113
-------------+----------------------------------------------------------------
ln_sig_v     |
       _cons |  -1.142592   .0527501   -21.66   0.000     -1.24598   -1.039203
-------------+----------------------------------------------------------------
ln_sig_u     |
     incsell |  -.3762457   .0403635    -9.32   0.000    -.4553567   -.2971348
     bussell |  -.2718317   .0475565    -5.72   0.000    -.3650408   -.1786227
     agesell |  -.2809178   .0516282    -5.44   0.000    -.3821072   -.1797284
     blksell |   .2727522   .0809753     3.37   0.001     .1140435    .4314608
     marsell |  -.1369204   .0412778    -3.32   0.001    -.2178234   -.0560174
      sfsell |  -.0725139   .0466659    -1.55   0.120    -.1639775    .0189497
     edusell |  -.3189302   .0356865    -8.94   0.000    -.3888744   -.2489859
     kidsell |  -.0180103    .034288    -0.53   0.599    -.0852135     .049193
       _cons |   .4297773   .0753251     5.71   0.000     .2821429    .5774118
------------------------------------------------------------------------------
```

ln_sig_w						
outbuy	.079076	.0725611	1.09	0.276	-.063141	.2212931
firstbuy	.0079438	.0672442	0.12	0.906	-.1238524	.13974
incbuy	.3901659	.0388278	10.05	0.000	.3140647	.4662671
busbuy	.3152677	.0641257	4.92	0.000	.1895835	.4409518
agebuy	.6749622	.108253	6.24	0.000	.4627902	.8871342
blkbuy	-1.366972	.6209467	-2.20	0.028	-2.584005	-.1499389
marbuy	-.0038389	.0801271	-0.05	0.962	-.1608851	.1532073
sfbuy	.1127059	.0998305	1.13	0.259	-.0829583	.3083702
edubuy	.4212932	.0724415	5.82	0.000	.2793104	.563276
kidbuy	-.1204749	.0587324	-2.05	0.040	-.2355884	-.0053615
_cons	-2.485998	.2466709	-10.08	0.000	-2.969464	-2.002532

Here we mainly focus on the estimated parameters in the $\boldsymbol{\delta}_{\mathbf{u}}$ and $\boldsymbol{\delta}_{\mathbf{w}}$ functions.

On the buyers' side, the negative signs on the Black dummy (`blkbuy`) and children dummy (`kidbuy`) in $\boldsymbol{\delta}_{\mathbf{w}}$ suggest that the costs of incomplete information for African American buyers and buyers with kids are lower. In contrast, buyers who have a higher income (`incbuy`), have a business (`busbuy`), are older (`agebuy`), and are educated (`edubuy`) pay higher information costs.

On the sellers' side, however, we find that income (`incsell`), having a business (`bussell`), age (`agesell`), being married (`marsell`), and having a college education (`edusell`) decrease information costs. Conversely, Black sellers (`blksell`) pay higher information costs.

6.3 Medical information asymmetry

6.3.1 Model estimation

Lu, Lian, and Lu (2011) estimated the effect of information asymmetry in the medical services market of China. Considering the ith doctor–patient pair, a doctor has a minimum price, $\underline{P}_i$, at which he or she would offer a particular medical service, and the patient has a maximum price, $\overline{P}_i$, that he or she could afford. The actual price can be written as

$$P_i = \underline{P}_i + \eta(\overline{P}_i - \underline{P}_i) = \mu(\mathbf{x}_i) + \eta\left\{\overline{P}_i - \mu(\mathbf{x}_i)\right\} - (1-\eta)\left\{\mu(\mathbf{x}_i) - \underline{P}_i\right\}$$

where $\mu(\mathbf{x}_i)$ is the optimal price in a doctor–patient pair and η represents the bargaining power of the doctor.

The 2TSF model fits data from the China Health and Nutrition Survey database. Here we add factor variables `i.province` and `i.year` into the syntax to absorb provincial and annual fixed effects.

```
. set seed 20220612
. use https://sftt.oss-cn-hangzhou.aliyuncs.com/lu11.dta, clear
. sftt lnprice lnage symp urban education job endurance insur i.province i.year
note: i_province_1 omitted because of collinearity.
note: i_year_1 omitted because of collinearity.
initial:       log likelihood = -3675.6139
rescale:       log likelihood = -3675.6139
rescale eq:    log likelihood =  -3645.644
Iteration 0:   log likelihood =  -3645.644
Iteration 1:   log likelihood = -3460.2012  (not concave)
  (output omitted)
Iteration 9:   log likelihood =  -3311.415
Two-tier stochastic frontier model with exponential specification
                                                Number of obs =  1,806
                                                Wald chi2(21) = 672.36
Log likelihood = -3311.415                      Prob > chi2   = 0.0000

------------------------------------------------------------------------------
     lnprice | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
frontier_ln_e|
       lnage |   .5982784   .1144348     5.23   0.000     .3739903    .8225664
    symptoms |   .7457815   .0548336    13.60   0.000     .6383096    .8532533
       urban |   .2049616   .0761177     2.69   0.007     .0557736    .3541495
   education |   .0630399    .034737     1.81   0.070    -.0050434    .1311233
         job |  -.2929144   .0777547    -3.77   0.000    -.4453109    -.140518
   endurance |  -1.015456   .0915982   -11.09   0.000    -1.194985   -.8359272
   insurance |   .0989838   .0857161     1.15   0.248    -.0690166    .2669843
  i_province_2 |   .2183171    .140904     1.55   0.121    -.0578497    .4944838
  i_province_3 |   .9542838   .1576787     6.05   0.000     .6452392    1.263328
  i_province_4 |   .4149037   .1523468     2.72   0.006     .1163095     .713498
  i_province_5 |   .5425718   .1485073     3.65   0.000     .2515029    .8336407
  i_province_6 |   1.158282   .1372435     8.44   0.000     .8892896    1.427274
  i_province_7 |   .6197577   .1360702     4.55   0.000      .353065    .8864504
  i_province_8 |   .6700105   .1298114     5.16   0.000     .4155849    .9244361
  i_province_9 |   .8400983   .1728504     4.86   0.000     .5013177    1.178879
    i_year_2 |  -.2171845   .2684136    -0.81   0.418    -.7432656    .3088965
    i_year_3 |   .4858937   .2390549     2.03   0.042     .0173548    .9544326
    i_year_4 |    .946091   .2338325     4.05   0.000     .4877878    1.404394
    i_year_5 |   1.093305   .2220374     4.92   0.000     .6581198     1.52849
    i_year_6 |   1.228319   .2168512     5.66   0.000     .8032985     1.65334
    i_year_7 |    1.20161   .2266252     5.30   0.000     .7574324    1.645787
       _cons |  -1.586893   .5319389    -2.98   0.003    -2.629474   -.5443124
-------------+----------------------------------------------------------------
ln_sig_v     |
       _cons |   .1269897   .1232088     1.03   0.303    -.1144951    .3684746
-------------+----------------------------------------------------------------
ln_sig_u     |
       _cons |  -1.177247   1.071734    -1.10   0.272    -3.277807    .9233118
-------------+----------------------------------------------------------------
ln_sig_w     |
       _cons |   .0036885   .0914932     0.04   0.968    -.1756349    .1830119
------------------------------------------------------------------------------
```

This command processes factor variables as a series of dummy variables, where `i_province_1` and `i_year_1` are dropped because of collinearity.

Next we decompose the composite error term into information asymmetry components.

```
. sftt eff, exp
The following variables have been generated:
_u_hat_exp
_w_hat_exp
_wu_diff_exp
_wu_net_effect
. summarize _u_hat_exp _w_hat_exp _wu_diff_exp
    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
  _u_hat_exp |      1,806    .2355296    .0406805   .1907761   .5762016
  _w_hat_exp |      1,806    .5016665    .1672711   .2073607   .9943287
_wu_diff_exp |      1,806     .266137    .2015247  -.3688409   .8035526
. summarize _wu_diff_exp, detail
                        _wu_diff_exp
-------------------------------------------------------------
      Percentiles      Smallest
 1%    -.1120609      -.3688409
 5%    -.0133144       -.332948
10%     .0352878      -.2467907       Obs               1,806
25%     .1196526      -.2240263       Sum of wgt.       1,806

50%     .2360287                      Mean            .266137
                        Largest       Std. dev.      .2015247
75%     .3937378       .7953292
90%     .5637176       .7954245       Variance       .0406122
95%     .6552005       .8010323       Skewness       .4913073
99%     .7682604       .8035526       Kurtosis       2.834541
```

From the first table in the results, the mean value of doctor surplus $E(e^{-w_i}|\varepsilon_i)$ (that is, the variable `_w_hat_exp`) is 0.5017, which means that, relative to the optimal price, doctor surplus makes the price 50.17% higher, while the patient surplus (`_u_hat_exp`) lowers the price by only 23.55%. The information asymmetry between doctors and patients eventually leads to medical service prices that are 26.62% (50.17% − 23.55%) higher than optimal prices.

The second table shows the quantile of net surplus $[E(e^{-u_i}|\varepsilon_i) - E(e^{-w_i}|\varepsilon_i)$; that is, the variable `_uw_diff_exp`]. The 10% quantile of the net surplus is $0.0353 > 0$, implying that at least 90% of patients must pay higher-than-optimal prices because of information asymmetry.

We then demonstrate the distribution of the surplus extracted by each side of patient–doctor pairs and the distributions of the net surplus.

```
. histogram _u_hat_exp, percent title(Percent, place(10) size(*0.7))
> ylabel(,angle(0)) ytitle("") xtitle("Surplus extracted by patients (%)")
> xscale(titlegap(3) outergap(-2)) scheme(sj)
(bin=32, start=.19077611, width=.01204455)
```

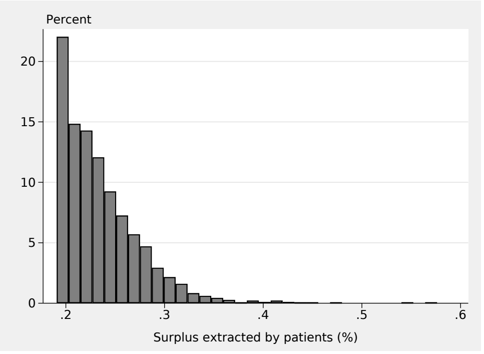

Figure 1. Surplus extracted by patients

```
. histogram _w_hat_exp, percent title(Percent, place(10) size(*0.7))
> ylabel(,angle(0)) ytitle("") xtitle("Surplus extracted by doctors (%)")
> xscale(titlegap(3) outergap(-2)) scheme(sj)
(bin=32, start=.20736069, width=.02459275)
```

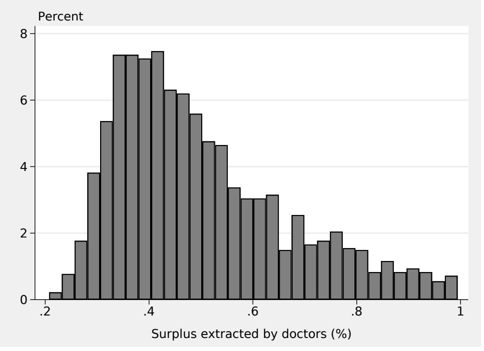

Figure 2. Surplus extracted by doctors

```
. histogram _wu_diff_exp, percent title(Percent, place(10) size(*0.7))
> ylabel(,angle(0)) ytitle("") xtitle("Net surplus (%)")
> xscale(titlegap(3) outergap(-2)) scheme(sj)
(bin=32, start=-.36884092, width=.0366373)
```

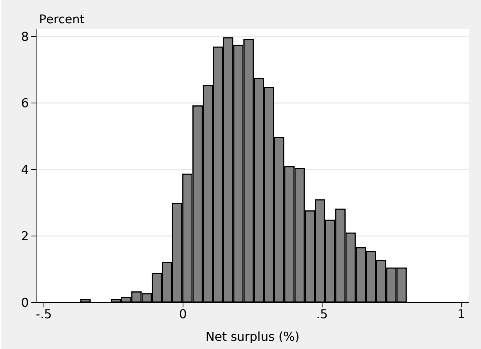

Figure 3. Distribution of net surplus

Figures 1 and 2 show that the patient surplus is much smaller than the doctor surplus. In contrast, both the patient surplus and the doctor surplus are right-skewed, indicating that only a few doctors have fully used their information superiority. Figure 3 illustrates the distribution of net surplus. From the histogram, we can see that most doctors have a positive net surplus, which means the prices of medical services are somewhat higher than optimal.

6.3.2 Contrasting the exponential and half-normal specifications

We then contrast the results from the exponential setting to the half-normal distributional setting. We run OLS estimation as well as the two specifications (exponential and half-normal) of the 2TSF model with the following syntax:

```
. use https://sftt.oss-cn-hangzhou.aliyuncs.com/lu11.dta, clear
. // OLS
. regress lnprice lnage symp urban education job endurance insur i.province
> i.year, vce(robust)
  (output omitted)
. // 2TSF - exponential specification
. sftt lnprice lnage symp urban education job endurance insur i.province i.year,
> findseed
  (output omitted)
. // 2TSF - half-normal specification
. sftt lnprice lnage symp urban education job endurance insur i.province i.year,
> hnormal findseed
  (output omitted)
```

The results are in table 1.

Table 1. Estimation results with different distributions

	(1) OLS	(2) Exponential	(3) Half-normal
lnage	0.666***	0.598***	0.557***
	(6.02)	(5.23)	(4.93)
symptoms	0.765***	0.746***	0.745***
	(13.23)	(13.60)	(13.65)
urban	0.210**	0.205**	0.206**
	(2.69)	(2.69)	(2.73)
education	0.0536	0.0630	0.0687*
	(1.47)	(1.82)	(2.05)
job	−0.274***	−0.293***	−0.286***
	(−3.44)	(−3.77)	(−3.71)
endurance	−0.969***	−1.015***	−0.990***
	(−10.43)	(−11.09)	(−11.00)
insurance	0.0947	0.0990	0.0998
	(1.07)	(1.15)	(1.17)
_cons	−1.255**	−1.587**	−1.512**
	(−2.61)	(−2.98)	(−3.05)
σ_u		0.3085	1.1468
σ_w		1.0037	2.0772
σ_v	1.5343	1.1354	0.5621
N	1806	1806	1806

NOTE: t statistics in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We can see that the coefficients are similar across the three columns, while the estimated standard errors vary between the exponential and the half-normal 2TSF specifications. The expected values of u_i and w_i also differ in table 2, as expected.

Table 2. Summary statistics of inefficiency

Specification	Variable	N	Mean	SD	Min	p50	Max
Exponential	`u_hat_e`	1806	0.309	0.0732	0.236	0.290	1.110
	`w_hat_e`	1806	1.004	0.695	0.261	0.775	5.841
Half-normal	`u_hat`	1806	0.915	0.376	0.442	0.780	3.753
	`w_hat`	1806	1.657	1.055	0.317	1.331	6.852

However, if we were interested in the rank of the one-sided terms, the choice of distributional assumptions may not significantly affect the results. Indeed, the rank correlation coefficients of u_i and w_i between the exponential and half-normal specifications are above 99%.

In figures 4 and 5, we present the scatterplots of rank(u_i) and rank(w_i). We see that the rankings of u_i and w_i are essentially linear, which also indicates that the ranking of inefficiency is not sensitive to the distributional assumptions.

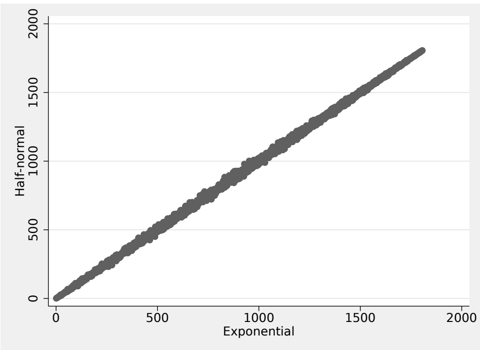

Figure 4. Ranking of u_i

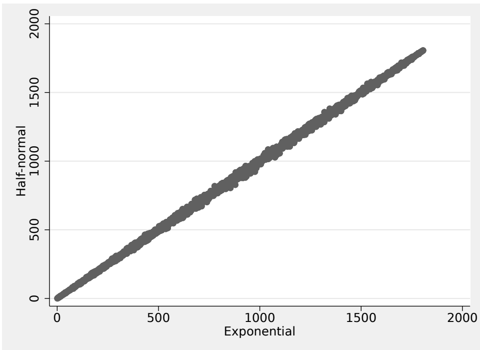

Figure 5. Ranking of w_i

7 Conclusions

In this article, we introduced a new command, **sftt**, to fit 2TSF models with cross-sectional data. When paired with distributional assumptions, this command can fit 2TSF models by using either an exponential (Polachek and Yoon 1987) or a half-normal (Papadopoulos 2015) specification. **sftt** also fits 2TSF models with the scaling property imposed (which are free of distribution assumptions), as in Parmeter (2018). We also provided two postestimation subcommands, **sftt sigs** and **sftt eff**, to help with variance identification and residual decomposition.

`sftt` provides the user with a simple way to fit 2TSF models that is intuitive and similar to `sfcross`. We illustrated the command's estimation capabilities through both simulated data and three distinct empirical datasets, using different frameworks to demonstrate the versatility of `sftt`. Our work here has covered the most popular cross-sectional approaches for the 2TSF model. There do exist several interesting extensions that could be added to the coding lexicon for the 2TSF. These include accounting for selection (Blanco 2017), the use of the fast Fourier transform to allow for different distributions for the separate one-sided shocks (Tsionas 2012), and extensions for the presence of panel data (Das and Polachek 2017). Any of these extensions into the Stata programming environment would help to further enhance the use of this flexible model.

8 Acknowledgments

We are grateful for the valuable comments and insightful ideas from Dr. Jenkins and an anonymous referee.

We also thank Alecos Papadopoulos for his helpful support.

9 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 23-1
. net install st0705      (to install program files, if available)
. net get st0705          (to install ancillary files, if available)
```

The corresponding code and results can be found on GitHub (https://github.com/arlionn/sftt).

10 References

Alvarez, A., C. Amsler, L. Orea, and P. Schmidt. 2006. Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics. *Journal of Productivity Analysis* 25: 201–212. https://doi.org/10.1007/s11123-006-7639-3.

Belotti, F., S. Daidone, G. Ilardi, and V. Atella. 2013. Stochastic frontier analysis using Stata. *Stata Journal* 13: 719–758. https://doi.org/10.1177/1536867X1301300404.

Blanco, G. 2017. Who benefits from job placement services? A two-sided analysis. *Journal of Productivity Analysis* 47: 33–47. https://doi.org/10.1007/s11123-016-0489-8.

Das, T., and S. W. Polachek. 2017. Estimating labor force joiners and leavers using a heterogeneity augmented two-tier stochastic frontier. *Journal of Econometrics* 199: 156–172. https://doi.org/10.1016/j.jeconom.2017.05.007.

Fé, E., and R. Hofler. 2020. sfcount: Command for count-data stochastic frontiers and underreported and overreported counts. *Stata Journal* 20: 532–547. https://doi.org/10.1177/1536867X20953566.

Fried, H. O., and L. W. Tauer. 2019. Efficient wine pricing using stochastic frontier models. *Journal of Wine Economics* 14: 164–181. https://doi.org/10.1017/jwe.2019.16.

Ge, T., J. Li, R. Sha, and X. Hao. 2020. Environmental regulations, financial constraints and export green-sophistication: Evidence from China's enterprises. *Journal of Cleaner Production* 251: 119671. https://doi.org/10.1016/j.jclepro.2019.119671.

Karakaplan, M. U. 2017. Fitting endogenous stochastic frontier models in Stata. *Stata Journal* 17: 39–55. https://doi.org/10.1177/1536867X1701700103.

———. 2022. Panel stochastic frontier models with endogeneity. *Stata Journal* 22: 643–663. https://doi.org/10.1177/1536867X221124539.

Kumbhakar, S. C., and C. F. Parmeter. 2009. The effects of match uncertainty and bargaining on labor market outcomes: Evidence from firm and worker specific estimates. *Journal of Productivity Analysis* 31: 1–14. https://doi.org/10.1007/s11123-008-0117-3.

———. 2010. Estimation of hedonic price functions with incomplete information. *Empirical Economics* 39: 1–25. https://doi.org/10.1007/s00181-009-0292-8.

Lin, Z. J., S. Liu, and F. Sun. 2017. The impact of financing constraints and agency costs on corporate R&D investment: Evidence from China. *International Review of Finance* 17: 3–42. https://doi.org/10.1111/irfi.12108.

Liu, Y., X. Yao, and T. Wei. 2019. Energy efficiency gap and target setting: A study of information asymmetry between governments and industries in China. *China Economic Review* 57: 101341. https://doi.org/10.1016/j.chieco.2019.101341.

Lu, H., Y. Lian, and S. Lu. 2011. Measurement of the information asymmetric in medical service market of China. *Economic Research Journal* 46(4): 95–107.

Lyu, X., C. Decker, and J. Ni. 2018. Compensation negotiation and corporate governance: The evidence from China. *Journal of Chinese Economic and Business Studies* 16: 193–213. https://doi.org/10.1080/14765284.2018.1445081.

Papadopoulos, A. A. 2015. The half-normal specification for the two-tier stochastic frontier model. *Journal of Productivity Analysis* 43: 225–230. https://doi.org/10.1007/s11123-014-0389-8.

———. 2018. The two-tier stochastic frontier framework (2TSF): Theory and applications, models and tools. PhD thesis, Department of Economics, Athens University of Economics and Business.

———. 2021. The two-tier stochastic frontier framework (2TSF): Measuring frontiers wherever they may exist. In *Advances in Efficiency and Productivity Analysis. NAPW 2018. Springer Proceedings in Business and Economics*, ed. C. F. Parmeter and R. C. Sickles, 163–194. Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-47106-4_8.

Parmeter, C. F. 2018. Estimation of the two-tiered stochastic frontier model with the scaling property. *Journal of Productivity Analysis* 49: 37–47. https://doi.org/10.1007/s11123-017-0520-8.

Polachek, S. W., and B. J. Yoon. 1987. A two-tiered earnings frontier estimation of employer and employee information in the labor market. *Review of Economics and Statistics* 69: 296–302. https://doi.org/10.2307/1927237.

Tsionas, E. G. 2012. Maximum likelihood estimation of stochastic frontier models by the Fourier transform. *Journal of Econometrics* 170: 234–248. https://doi.org/10.1016/j.jeconom.2012.04.001.

Wang, H.-J., and P. Schmidt. 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis* 18: 129–144. https://doi.org/10.1023/A:1016565719882.

About the authors

Yujun Lian is an associate professor of finance at the Sun Yat-sen University.

Chang Liu is a PhD candidate of finance at the Sun Yat-sen University.

Christopher Frank Parmeter is an associate professor of economics at the University of Miami.