

Derivation of marginal effects of determinants of technical inefficiency

Subal C. Kumbhakar¹ and Kai Sun*²

¹Department of Economics, State University of New York at Binghamton, NY 13902, USA

²Economics and Strategy Group, Aston Business School, Aston University, Birmingham, B4 7ET, UK

Abstract

In efficiency studies using the stochastic frontier approach the main focus is to explain inefficiency in terms of some exogenous variables and computation of marginal effects of each of these determinants. Although inefficiency is estimated by its mean conditional on the composed error term (the Jondrow et al. (1982) estimator), the marginal effects are computed from the unconditional mean of inefficiency (Wang (2002)). In this paper we derive the marginal effects based on the Jondrow et al. estimator and use the bootstrap method to compute confidence intervals of the marginal effects.

Keywords: Stochastic frontier model; Inefficiency; Environmental variable

JEL codes: C16, C24

*Corresponding author. Economics and Strategy Group, Aston Business School, Aston University, Birmingham, B4 7ET, UK. Email: k.sun@aston.ac.uk. Tel: +44 121 204 3162.

The authors would like to thank an anonymous referee for useful comments.

1 Introduction

In stochastic frontier models the two main objectives are to estimate underlying production technology and observation-specific technical inefficiency. While estimating inefficiency, the empirical studies in this literature examine whether differences in inefficiency among producers can be explained in terms of some exogenous (environmental) variables. A natural question in this context is how to compute the marginal effects of these environmental variables on inefficiency. For this, first we need a model that includes these environmental variables in the specification of inefficiency, and then a point estimator of inefficiency. Models in which the environmental variables enter into the mean and/or the variance of inefficiency have been proposed in some earlier studies, e.g., Kumbhakar, Ghosh and McGuckin (1991), Reifschneider and Stevenson (1991), Huang and Liu (1994), and Battese and Coelli (1995).¹ Wang (2002) examined this model thoroughly by allowing the environmental variables to enter into the mean and the variance of inefficiency and derived the formula for calculating marginal effects of the environmental variables on inefficiency. However, his derivation of the marginal effects was based on the unconditional mean of inefficiency, although the estimator of inefficiency was based on the conditional mean – the Jondrow et al. (1982) (henceforth JLMS) estimator.

In this paper we derive the marginal effects of environmental variables on inefficiency where inefficiency is estimated using the Jondrow et al. (1982) formula. We consider a model in which the environmental variables appear in both the mean and the variance of inefficiency as well as in the variance of the noise term. Based on this model we show that there are three channels through which the environmental variables can effect the estimated inefficiency. We show that even if these variables do not enter into either the mean or the variance of inefficiency, they can affect inefficiency via the variance of the noise component. This is a new result which comes from the fact that the JLMS estimator (shown later) depends on the variance of the noise component. Since the JLMS estimator is universally used for estimating inefficiency, it should also be used to compute the marginal effects. That is, both inefficiency and its marginal effects should be based on the same formula.

The variance of the noise component in our model is made a function of the same environmental variables that affect the mean and/or the variance of inefficiency. Since we use the JLMS estimator which (as shown later) is a function of the mean and the variance of inefficiency as well as the variance of the noise component, the marginal effects of environmental variables will have a component coming from the variance of the noise term. This extra component was absent from the Wang (2002) formula and has not been discussed in any studies before because the JLMS estimator was not used in the literature to compute marginal effects.²

¹Although chronologically last, in the efficiency literature these models are known as the Battese-Coelli (1995) model.

²Wang (2002) used the unconditional mean of inefficiency (which is independent of the variance of the noise term) to compute

We apply European banking data and estimate inefficiency as well as the marginal effects. We compare our marginal effects with those based on Wang (2002). We also compute confidence intervals of the marginal effects using bootstrap procedure. Our results show that the marginal effects based on the Wang (2002) formula tend to overestimate marginal effects in our application.

2 A Stochastic Frontier Model with Environmental Variables

Consider a stochastic production frontier model in a cross-sectional setting, viz.,

$$y_i = \beta' x_i + v_i - u_i, \quad (1)$$

$$u_i \sim N^+(\mu_i, \sigma_{ui}^2), \quad (2)$$

$$v_i \sim N(0, \sigma_{vi}^2), \quad (3)$$

$$\mu_i = c_0 + \delta' z_i, \quad (4)$$

$$\sigma_{ui} = \exp(c_1 + \gamma' z_i), \text{ and} \quad (5)$$

$$\sigma_{vi} = \exp(c_2 + \rho' z_i), \quad (6)$$

where u_i is the non-negative technical inefficiency component, which follows a truncated normal distribution. The vector of environmental variables z_i are allowed to affect the pre-truncation mean and variance of u_i , μ_i and σ_{ui}^2 , respectively. The noise component is v_i distributed normally with zero mean and variance σ_{vi}^2 which is assumed to be a function of z_i as well.³ Following Jondrow et al. (1982), it can be shown that the distribution of u_i given the composed error $\varepsilon_i = v_i - u_i$ is truncated normal with mean $\tilde{\mu}_i = (\mu_i \sigma_{vi}^2 - \varepsilon_i \sigma_{ui}^2) / \sigma_i^2$ and standard deviation $\sigma_{*i} = \sigma_{ui} \sigma_{vi} / \sigma_i$, where $\sigma_i^2 = \sigma_{ui}^2 + \sigma_{vi}^2$. Thus the point estimator of u_i is given by the conditional mean, i.e.,

$$E(u_i | \varepsilon_i) = \tilde{\mu}_i + \sigma_{*i} \frac{\phi(\tilde{\mu}_i / \sigma_{*i})}{\Phi(\tilde{\mu}_i / \sigma_{*i})}, \quad (7)$$

where ϕ and Φ denote the standard normal density and distribution functions, respectively. The estimator in (7) is known as the JLMS estimator in the efficiency literature.

Wang (2002) used the formula in (7) to calculate inefficiency but he used the post-truncation mean of u_i , i.e., $E(u_i | u_i > 0) = \sigma_{ui} [\Lambda_i + \frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)}]$, where $\Lambda_i = \mu_i / \sigma_{ui}$, to compute the marginal effects. In other words,

the marginal effects and this is why the extra term was missing in his derivation.

³Note that $\sigma_{ui}^2 = \exp(2c_1 + 2\gamma' z_i)$ and $\sigma_{vi}^2 = \exp(2c_2 + 2\rho' z_i)$.

his marginal effects are computed from $\frac{\partial E(u_i|u_i>0)}{\partial z_{li}}$, where z_{li} is the l -th element of z_i . More specifically, the formula for the marginal effects in Wang (2002) is:

$$\begin{aligned} \frac{\partial E(u_i)}{\partial z_{li}} = & \delta_l \left\{ 1 - \Lambda_i \frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)} - \left[\frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)} \right]^2 \right\} \\ & + \gamma_l \sigma_{ui} \left\{ (1 + \Lambda_i^2) \frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)} + \Lambda_i \left[\frac{\phi(\Lambda_i)}{\Phi(\Lambda_i)} \right]^2 \right\}, \end{aligned} \quad (8)$$

where δ_l and γ_l are the coefficients associated with z_{li} in (4) and (5), respectively.

To derive the formula for the marginal effects based on the JLMS estimator in (7), i.e., $\frac{\partial E(u_i|\varepsilon_i)}{\partial z_{li}}$, we define $m_i = \tilde{\mu}_i/\sigma_{*i}$ and $g_i = \phi(m_i)/\Phi(m_i)$. After a lengthy and tedious algebra (which are skipped here but available from the authors upon request) the marginal effects of the l -th environmental variable on $E(u_i|\varepsilon_i)$ is found to be:

$$\begin{aligned} \frac{\partial E(u_i|\varepsilon_i)}{\partial z_{li}} = & \delta_l \left[\frac{\sigma_{vi}^2}{\sigma_i^2} (1 - m_i g_i - g_i^2) \right] \\ & + \gamma_l \frac{1}{\sigma_i^2} \left\{ \sigma_{vi}^2 \sigma_{*i} [g_i(1 + m_i^2) + m_i g_i^2] - 2\sigma_{*i}^2 (\varepsilon_i + \mu_i)(1 - g_i^2 - m_i g_i) \right\} \\ & + \rho_l \frac{1}{\sigma_i^2} \left\{ \sigma_{ui}^2 \sigma_{*i} [g_i(1 + m_i^2) + m_i g_i^2] + 2\sigma_{*i}^2 (\varepsilon_i + \mu_i)(1 - g_i^2 - m_i g_i) \right\}, \end{aligned} \quad (9)$$

where δ_l , γ_l and ρ_l are the coefficients associated with z_{li} in (4), (5) and (6), respectively. This result shows that the marginal effects of z_l have three components which identify three separate channels through which z_l affects the estimated inefficiency. These components/channels are related to μ_i , σ_{ui} and σ_{vi} functions. That is, if μ_i , σ_{ui} and σ_{vi} are functions of z_i , then each element of z_i affects inefficiency via the three channels given by the three terms on the right-hand-side of (9). On the other hand, if σ_{ui} and σ_{vi} are constants, then the marginal effects come only from the mean and it is δ_l multiplied by an adjustment function which is positive.⁴ Similarly, if μ_i is constant (i.e., u_i follows a truncated-normal distribution with heteroskedasticity) the first term in (9) drops out and the channels by which z_l affects inefficiency are through the variances of u_i and v_i . The same holds true when $\mu_i = 0$ (i.e., u_i follows a half-normal distribution with heteroskedasticity). If $\mu_i = 0$ and σ_{vi}^2 is a constant, then z_l affects inefficiency through σ_{ui}^2 and this is captured by the second term in (9). Finally, the new result in (9) comes from the last term that captures the effect via σ_{vi}^2 . This term is new and is not explored in the literature.⁵ This component is interesting because it shows that if σ_{vi}^2 is a function of z_i , these variables can also affect inefficiency even if $\mu_i = 0$ and σ_{ui}^2 is a constant.

⁴Note that this function is similar to the one in (8) (equation (9) in Wang (2002, p.244)), except for the extra term of σ_{vi}^2/σ_i^2 .

⁵It can be seen from (8) that the Wang (2002) formula did not have this component although he allowed the variance of the noise component, σ_{vi}^2 , to depend on z variables (in the application part). This is because $E(u_i|u_i > 0)$ does not depend on σ_{vi}^2 .

3 Empirical Application

As an empirical illustration, we use an unbalanced panel data of European banking industry obtained from Bankscope.⁶ The data set covers 15 European countries⁷ for 17 years (during the period 1993 and 2009) with a total of 6,733 bank-year observations. Although we considered a cross-sectional model and our derivation of marginal effects is based on that cross-sectional model, the panel extension of it is trivial. For this we add an extra subscript t to $z_i, \mu_i, \sigma_{ui}^2$ and σ_{vi}^2 , and specify the model as in Wang (2002):⁸

$$y_{it} = \beta' x_{it} + v_{it} - u_{it}, \quad (10)$$

$$u_{it} \sim N^+(\mu_{it}, \sigma_{uit}^2), \quad (11)$$

$$v_{it} \sim N(0, \sigma_{vit}^2), \quad (12)$$

$$\mu_{it} = c_0 + \delta' z_{it}, \quad (13)$$

$$\sigma_{uit} = \exp(c_1 + \gamma' z_{it}), \text{ and} \quad (14)$$

$$\sigma_{vit} = \exp(c_2 + \rho' z_{it}). \quad (15)$$

In this application we use a single output (loans) and three inputs (personnel expenses, physical capital expenses, and interest expenses). We use three different measures of bank risk as environmental variables, viz., credit risk (loan loss provision/total assets), solvency risk (equity/total assets) and liquidity risk (liquidity/total assets). We also control for year and country effects. Thus the data matrices for our model are:

y_{it} : ln(loans),

x_{it} : 1, ln(personnel expenses), ln(physical capital expenses), ln(interest expenses), country dummies, year dummies,⁹

z_{it} : credit risk, solvency risk, liquidity risk.

The parameters in (10) and (13)-(15) are estimated by maximizing the log-likelihood function, which for

⁶See *bankscope2.bvdep.com* for details.

⁷These are Austria, Belgium, Denmark, Finland, France, Germany, Great Britain, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, and Sweden.

⁸There are, of course, many other ways of specifying inefficiency in panel models (see for example, Kumbhakar and Lovell (2000), Kumbhakar et al. (2012), among others.)

⁹One country dummy and one year dummy are dropped to avoid multicollinearity.

the i -th bank at time t is (Kumbhakar and Lovell (2000)):

$$L_{it} = -\frac{1}{2} \ln \sigma_{it}^2 + \ln \left[\phi \left(\frac{y_{it} - \beta' x_{it} + \mu_{it}}{\sigma_{it}} \right) \right] - \ln \left[\Phi \left(\frac{\mu_{it}}{\sigma_{uit}} \right) \right] + \ln \left[\Phi \left(\frac{\tilde{\mu}_{it}}{\sigma_{*it}} \right) \right]. \quad (16)$$

The MLE of the parameters are used to calculate the point estimates for technical inefficiency using (7) and the marginal effects of an environmental variable on the estimated technical inefficiency using (9). To save space and concentrate on the marginal effects, we report and compare the marginal effects on $E(u_{it}|\varepsilon_{it})$ and on $E(u_{it}|u_{it} > 0)$.¹⁰

Since our focus is on the marginal effects we are not reporting parameter estimates and estimated inefficiency. The marginal effects of each risk variable against itself, holding all other variables at their means, are reported in Figure 1. For comparison we report the marginal effects calculated from both (8) and (9). It can be seen that the marginal effects based on (8) are much higher for the last two risk measures, while for credit risk it is opposite only in the range -0.034 to -0.004. Thus policy implications will be different depending on which formula is used to calculate the marginal effects. Since the JLMS estimator is used to estimate inefficiency, it is natural to use it for calculating marginal effects.

A natural question after estimating the marginal effects of a particular environmental variable is to check whether these are statistically significant. The JLMS estimator does not have a known distribution. The same holds for the marginal effects. However, it is possible to use bootstrap procedure to construct confidence intervals for the marginal effects. Figure 2 reports the marginal effects along with their confidence intervals obtained via the following bootstrap procedure:

Step 1: Calculate the marginal effects from (9), using the estimated parameter vector via MLE, i.e., $[\hat{\beta}, \hat{c}_0, \hat{\delta}, \hat{c}_1, \hat{\gamma}, \hat{c}_2, \hat{\rho}]$, in (10).

Step 2: Generate a pseudo-sample using the parameter estimates in Step 1. Draw v_{it}^* from $N(0, \hat{\sigma}_{vit}^2)$, and u_{it}^* from $N^+(\hat{\mu}_{it}, \hat{\sigma}_{uit}^2)$, $\forall i = 1, \dots, N$ and $t = 1, \dots, T$, where $\hat{\mu}_{it} = \hat{c}_0 + \hat{\delta}' z_{it}$, $\hat{\sigma}_{uit}^2 = \exp(2\hat{c}_1 + 2\hat{\gamma}' z_{it})$ and $\hat{\sigma}_{vit}^2 = \exp(2\hat{c}_2 + 2\hat{\rho}' z_{it})$. Then, generate $y_{it}^* = \hat{\beta}' x_{it} + v_{it}^* - u_{it}^*$.

Step 3: Obtain MLE of the parameters using the pseudo-sample, $\{y_{it}^*, x_{it}, z_{it}\}_{i=1, t=1}^{N, T}$, generated in Step 2. Then, use the estimated parameters to calculate the marginal effects from (9).

Step 4: Repeat the previous two steps B times.¹¹ This gives us a sample of $(1 + B)$ marginal effects for each observation. The 95% confidence interval for a particular observation is then constructed using the 2.5th and 97.5th percentiles from the sample of $(1 + B)$ marginal effects.

¹⁰The simplest case is to compute marginal effects from the pre-truncation mean of u_{it} , i.e., $E(u_{it}) = \mu_{it}$ which are simply δ_l . This will make the marginal effects monotonic, while in our case as well as in Wang (2002) the marginal effects are non-monotonic.

¹¹ B in this example is set to be 99.

Using (9) and the bootstrap procedure for each risk measure, it can be seen that the sign, magnitude, and significance of the marginal effects do not always stay the same along the distribution of each risk variable. Note that the marginal effects and their 95% confidence intervals are observation-specific. If the zero line is inside the 95% confidence interval, then the marginal effect is insignificant.

4 Conclusion

In this paper we derived marginal effects of environmental variables (z) on estimated inefficiency using the Jondrow et al. (1982) (JLMS) estimator which is widely used in the efficiency literature. Our results show that the z variables can affect inefficiency through three channels. We show that even if these variables do not enter into either the mean or the variance of inefficiency, they can affect inefficiency via the variance of the noise component if it depends on z . This is a new result. We also show how to compute confidence intervals of these marginal effects using bootstrap. Finally, we illustrate these using a panel of European banks. In our application, we find that the Wang (2002) formula overestimates the marginal effects compared to our formula.

References

- [1] Battese, G. E. and T. J. Coelli (1995), "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data," *Empirical Economics* 20, pp. 325-332.
- [2] Huang, C. J. and J. T. Liu (1994), "Estimation of a Non-neutral Stochastic Frontier Production Function," *Journal of Productivity Analysis* 5, pp. 171-180.
- [3] Jondrow, J., C. A. K Lovell, I. S. Materov and P. Schmidt (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics* 19, pp. 233-238.
- [4] Kumbhakar, S. C. and C. A. K. Lovell (2000), *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge, New York.
- [5] Kumbhakar, S. C., S. Ghosh and J. T. McGuckin (1991), "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms," *Journal of Business and Economic Statistics* 9, pp. 279-286.
- [6] Kumbhakar, S.C., G. Lien and J. B. Hardaker (2012), "Technical Efficiency in Competing Panel Data Models: A Study of Norwegian Grain Farming," *Journal of Productivity Analysis* (forthcoming), DOI 10.1007/s11123-012-0303-1.
- [7] Reifschneider, D. and R. Stevensen (1991), "Systematic Departure from the Frontier: A Framework for the Analysis of Firm Efficiency," *International Economic Review* 32, pp. 713-723.
- [8] Wang, H-J. (2002), "Heteroscedasticity and Non-monotonic Efficiency Effects of a Stochastic Frontier Model," *Journal of Productivity Analysis* 18, pp. 241-253.

Figure 1: Marginal effects of risk variables

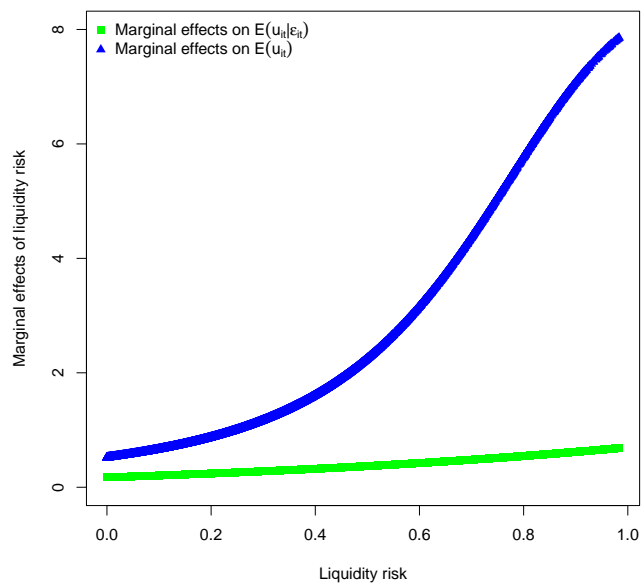
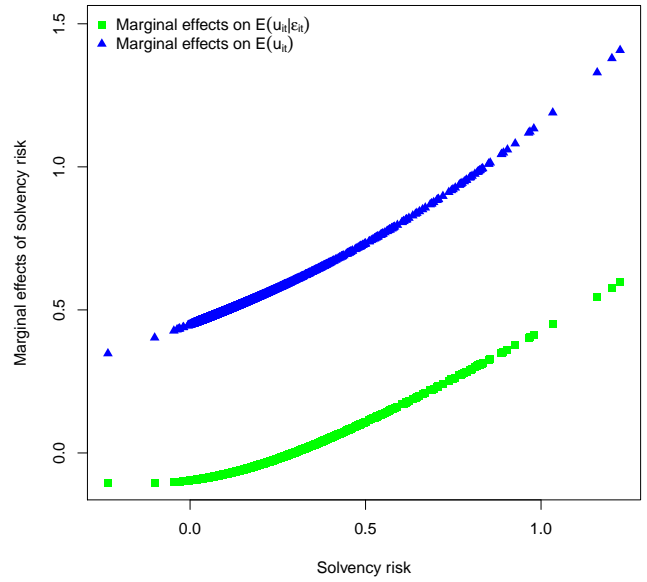
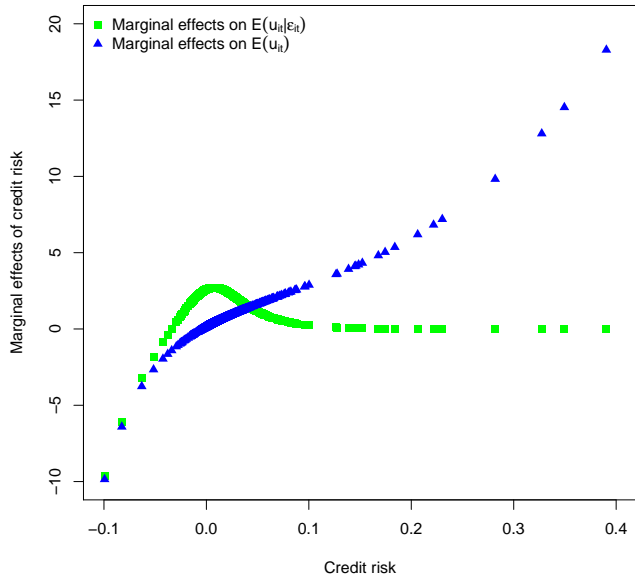


Figure 2: Confidence intervals of marginal effects of risk variables

