Evaluating the accuracy of scope economies: comparisons among delta method, bootstrap, and Bayesian approach.

Liang-Cheng Zhang*[a] Andrew C. Worthington[a]

[a]*Department of Accounting, Finance and Economics, Griffith University, Australia*

Abstract

The estimate of scope economies is a nonlinear combination of estimated coefficients from an empirical model. This estimate usually involves out-of-sample predictions when calculating the separated costs (as a part of calculation of scope economies). These difficulties make it hard to give the precise prediction and to calculate the standard deviation of this estimate along with its confidence intervals. In this paper, we demonstrate methods for constructing confidence interval for scope economies to allow researchers to draw inferences from estimated economies of scope. We review the common approaches such as delta method or bootstrap adopted by previous studies. In contrast of the above approximation methods, this study also proposes an alternative method, Bayesian approach, to produce full predictive distribution for this measure with posterior distribution. To demonstrate these three approaches, we use a balanced panel data including 37 Australian public universities over the period 2003-12. All three approaches use a quadratic cost function with two outputs. Estimates of scope economies will be calculated with the sample data and estimated parameters from the model. Results shows that our Bayesian approach gives the most precise (the least standard deviations) among all approaches.

JEL classification: C11, C53, I23, I28

Keywords: Bayesian inference, delta method, bootstrap, scope economies, confidence interval

---

* Contact details of corresponding author. E-mail address: liang-cheng.zhang@griffithuni.edu.au.

Note

This paper has been accepted in 2015 Australian Conference of Economists PhD Colloquium.

1. Introduction

Since Panzar and Willig (1981) coined the term “economies of scope”, this measure of scope economies has been an important economic index to indicate whether there are the cost savings resulting from the scope of the firm’s operation. This means simultaneous production of multiple types of outputs in a single enterprise will be cheaper than their production in its own specialised firm. This index provides firms and their stakeholders with valuable information about the structure modifications including mergers and acquisitions. There has also been considerable interest in applying this estimate to different types of multi-output production industries such as higher education (Worthington & Higgs, 2011), bank (Berger, Hunter, & Timme, 1993; Cummins, Weiss, Xie, & Zi, 2010; Delgado, Parmeter, Hartarska, & Mersland, 2015), health care (Preyra & Pink, 2006), energy (Farsi, Fetz, & Filippini, 2008; Fetz & Filippini, 2010) and agriculture (Ohe & Kurihara, 2013; Schroeder, 1992). However, this important index is not build on a complete estimation process. Most previous studies focus on the functional form[1] of cost function or different approaches[2] to estimate but few studies concerns whether correct inference could be draw from this measure. Specifically, estimate of scope economies is usually estimated by simply plugging in sample data into the estimated cost function without considering its standard error. Therefore, intervals for this estimate could not be constructed, and we have little confidence in it. In other words, this naïve (Held & Bové, 2014) or plug-in (Aitchison & Dunsmore, 1975; Lancaster, 2004) procedure neglects uncertainty over estimates of scope economies as well as the uncertainty of regression coefficients.

In order to make inference about the estimate of scope economies, its standard deviation and confidence interval should be estimated as well. Two approximation methods are used by previous studies: delta method and bootstrap. Nonetheless, none of them produces the exact estimates. Furthermore, when constructing confidence intervals with bootstrap, different technique often produces different results (Eakin, McMillen, & Buono, 1990). Most studies also fail to specify which bootstrap method they used. Misleading conclusions could be drawn since different methods could lead to different conclusions. Most importantly, either delta method or bootstrap method still neglects the uncertainty of regression coefficients since it still treats estimated coefficients fixed during the process. The misleading results could be made due to the underestimated prediction variance.

---

[1] See (Pulley & Braunstein, 1992; Pulley & Humphrey, 1993) for the debates of functional form when estimating the scope economies.

[2] Approaches include semi-parametric (Delgado et al., 2015) and non-parametric methods (Cummins et al., 2010).

The preceding problems could be more appropriately treated with a Bayesian approach which is always able to produce probability distributions for prediction (Geisser, 1993). There is still no study using it to estimate scope economies and yet compare these estimates to those with delta method and bootstrap methods developed by Eakin et al.(1990). In this paper, we demonstrate methods for estimating standard deviations and constructing confidence interval for scope economies to allow researchers to draw inferences from estimated economies of scope. The remainder of this paper is organised into the following sections. We begin with the short introduction of scope economies and how they are usually gained. Common methods for constructing confidence intervals for scope economies will be introduced and then we propose a Bayesian approach as an alternative method. Section 3 uses a dataset composing of Australian universities and compares confidence intervals using different methods. Finally, we offer our suggestions in the final section.

2. Estimate of scope economies and its inference measures

The fundamental concept of scope economies is developed by Panzar and Willig (1981) and we follow their notations afterwards. Suppose that, firms produce n types of product whose quantities are contained in the vector $y = (y_1, \dots, y_n)$ with m types of inputs whose quantities are $x = (x_1, \dots, x_m)$ given a vector of input prices $w = (w_1, \dots, w_m)$, multiproduct minimum cost function could be specified as $C(y, w)$. Note that n is a subset of N, that is $N = (1, \dots, n)$. Let y_s denote the n vector whose elements are set equal to those of y for $i \in S \subset N$ and 0 for $i \neq S$. With these notations, the function $C(y_s, w)$ can be described as the minimum cost of producing only the products in the subset S, at the quantities of the vector y. The formal definition of scope economies is:

Let $T = (T_1, \dots, T_l)$ denote a nontrivial partition of $S \subset N$, which indicates $\cup_i T_i = S, T_i \cap T_j = \emptyset, for\ i \neq j, T_i \neq \emptyset, and\ l > 1$. Economies of scope exist at y_s and at input prices w with respective to the partition T if

$$\sum_{i=1}^{l} C(y_{T_i}, w) > C(y_s, w) \tag{1}$$

Intuitively, $C(y_{T_i}, w)$ is the cost of separate production (i.e., producing only y_{T_i}) since T_i a nontrivial partition of $S \subset N$, while $C(y_s, w)$ represents the cost of joint production due to $\cup_i T_i = S$. Therefore, (1) shows the total cost of separate production (i.e., the summation of all cost of separate production for $i = 1, \dots, l$) are larger than the cost of joint production. On the contrary, if the above equality is reversed, it indicates there are *dis*economies of scope.

In practice, to know whether there are scope economies, Baumol, Panzar, & Willig (1982) define the degree of scope economies (SE) by modifying (1) as

$$SE = \frac{\sum_{i=1}^{l} C\left(y_{T_i}, w\right) - C(y_s, w)}{C(y_s, w)} \quad (2)$$

In this way, we can gain the point estimate of scope economies. This equation is also more meaningful: it represents the proportion of cost savings from joint production.

Unfortunately, either costs of separate production or costs of joint production cannot be directly observed from the data and they are usually estimated from the cost function $C(y, w)$[3]. Therefore, we need to estimate cost function first to calculate the scope economies. In order to make our model more explicit for each firm, we further stack our observations (cost, outputs and prices) over time for each individual firm. Consider a panel data set, there are T obervations on each of the N individual firms

$$C_f = C(y, w) + \varepsilon_f = Y_f \beta + \varepsilon_f,$$
$$for\ f = 1 \dots N, \text{ with } \varepsilon_f \sim dnorm(0, \mathbf{I}_T \tau_C) \quad (3)$$

where

$$C_f = \begin{bmatrix} C_{f1} \\ \vdots \\ C_{fT} \end{bmatrix},\ y_f = \begin{bmatrix} 1 & Y_{2f1} & \cdots & Y_{Kf1} \\ \vdots & \vdots & & \vdots \\ 1 & Y_{2fT} & \cdots & Y_{KfT} \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}, \text{and } \mathbf{I}_T \tau_C = \begin{bmatrix} 1/\sigma_C^2 & 0 & \cdots & 0 \\ 0 & 1/\sigma_C^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1/\sigma_C^2 \end{bmatrix}$$

We can see that all coefficients, β (including intercepts), are included in a K vector of coefficients, C_f to be a $T \times 1$ vector containing T different points in time and Y_f which includes output and price vector to be a $T \times K$ matrix containing T different points in time on each of the K explanatory variables. In other words, our observations (cost, outputs and prices) over time have been stacked for each firm (*f*).

In practice, the output and price vector will be fixed at their means or the value of f-th firm. Suppose that the cost function $C(y, w)$ has been estimated with (3) and we can gain K estimated parameters $\widehat{\beta_1}, \widehat{\beta_2}, \dots, \widehat{\beta_K}$. Therefore SE in (2) actually the combinations of mean costs at the point $y_1 \dots y_n$ and w. In addition, since output and price vectors have been fixed at certain level, SE can be equal to a conditional expectation (or mean) function of K estimated parameters, that is,

[3] Numerous approaches are proposed to estimate scope economies without estimating cost function first including using distance function (Hajargasht, Coelli, & Rao, 2008).

$$\mathrm{SE} = \frac{\sum_{i=1}^{l} E\left(C|y_{T_i}, w, y_{T_j} = 0 \; for \; j \neq i\right) - E(C|y_s, \mathrm{w})}{E(C|y_s, \mathrm{w})} = \mathrm{SE}(\hat{\beta}) \tag{4}$$

where the hat symbol $\widehat{}$ indicates quantities calculated from the observed data and E denotes the expected value of cost[4].

From (4), we can observe that one difficulty of calculating scope economies is that we usually have to make out-of-sample extrapolation (or prediction) since zero output rarely happens in all output variables (Pulley & Humphrey, 1993). In other words, data of specialised firms are usually unavailable and the cost function is estimated with only joint producers who produce all types of products in study (Berger, Cummins, Weiss, & Zi, 2000)[5]. Out-of-sample extrapolation would result in inaccurate evaluations of costs of separate production (that is, $E\left(C|y_{T_i}, w, y_{T_j} = 0 \; for \; j \neq i\right)$ in (4)), which further leads to inaccurate estimation of scope economies.

Another difficulty which most often neglected by previous studies is how to make inference about the measure of scope economies. Several review articles(Byrnes & Dollery, 2002; Clark, 1988; Meyer, 2012; Tovar, Jara-Díaz, & Trujillo, 2007) have shown that there are at least two steps to gain the estimate: estimate the cost function with a proper functional form and then calculate the point estimate of scope economies based on the estimated cost function. However, their investigations show that most studies judge existence of scope economies *only* by whether the point estimate of (2) is larger than zero. In other words, the point estimate of scope economies is inferred purely by naïve (Held & Bové, 2014) or plug-in (Aitchison & Dunsmore, 1975; Lancaster, 2004) procedure and therefore it neglects uncertainty over estimates of scope economies as well as the uncertainty of the cost function coefficients. Therefore, intervals could not be constructed, and we have little confidence in this estimate.

In fact, intervals for the measure of scope economies (SE) could be constructed in a common probability structure(Casella & Berger, 2002):

$$Prob(L \leq \mathrm{SE} \leq U) = 1 - \alpha \tag{5}$$

---

[4] Notice that this hat symbol also explains the view of a frequentist: the model parameters are fixed but unknown. However, from a Bayesian view introduced in next section, we still use the same formula in (4) but replace $\hat{\beta}$ with β since they are considered to be random variables.

[5] To minimise the effect of the out-of-sample prediction on the estimate of scope economies, Berger et al.(2000) divided their samples based on whether produces all types of products and then allow cost function to differ with different dataset (joint producers and separate producers). Some studies prefer to use an appropriate model to control out-of-sample bias (see Farsi et al.(2008) and Fetz and Filippini (2010) for examples). The former approach is usually limited by the data. We consider the use of model approach to control the bias.

where L and U are the lower and upper interval limits, α is significance level, and $1-\alpha$ is confidence level.

Two interpretations could be applied to (6): Frequentist confidence interval and Bayesian confidence interval. A Frequentist confidence interval with a $100(1-\alpha)\%$ confidence interval is an interval to be expected to cover the true value of SE $100(1-\alpha)\%$ of the time in the repeated sampling of observations. On the other hand, a credible interval does not rely on repeated sampling but the posterior probability that the SE lies in the interval. This approach with non-informative priors is called Bayesian confidence interval (or credible interval) (Box & Tiao, 1992). So far, delta method and bootstrap are the most common methods to construct the intervals for scope economies but they are usually under the assumption of Frequentist statistical inference. These methods will be introduced and then we propose a Bayesian approach as an alternative method.

2.1 Frequentist statistical inference

To construct the confidence intervals for SE, variance for this estimate should be calculated first. Using delta method, SE can be approximated with a Taylor expansion around $\beta_1 \dots \beta_K$ but keeping only the first two terms of the Taylor series:

$$\mathrm{SE}(\hat{\beta}) \approx \mathrm{SE}(\beta) + \sum_{k=1}^{K} \left(\frac{\partial \mathrm{SE}(\widehat{\beta_k})}{\partial \widehat{\beta_k}} |_{\widehat{\beta_1}=\beta_1,\dots,\widehat{\beta_K}=\beta_K} \right) (\widehat{\beta_k} - \beta_k) \qquad (6)$$

which implies the variance of $\mathrm{SE}(\widehat{\beta_k})$ is approximately (see page 485 to 486 in Kmenta and Rafailzadeh (1997)or page 68 to 69 in Greene (2011) for further details)

$$\mathrm{Var}\left(\mathrm{SE}(\hat{\beta})\right) \approx \sum_k \left(\frac{\partial \mathrm{SE}(\widehat{\beta_k})}{\partial \widehat{\beta_k}} \right) Var(\widehat{\beta_k}) + 2 \sum_{j<k} \left(\frac{\partial \mathrm{SE}(\widehat{\beta_k})}{\partial \widehat{\beta_j}} \right) \left(\frac{\partial \mathrm{SE}(\widehat{\beta_k})}{\partial \widehat{\beta_k}} \right) Cov(\widehat{\beta_j}, \widehat{\beta_k}), \qquad (7)$$

$$for\ j,k = 1,2,\dots,K$$

This could be rewritten in a more well-known matrix form adopted by previous studies (Altunbas & Molyneux, 1996; Fields & Murphy, 1989; Mester, 1987; Suret, 1991): $\mathrm{D\Sigma D'}$ where D is the gradient of function f and Σ is a K by K variance-covariance matrix of the estimated parameters $\hat{\beta}$. However, it is difficult to calculate the derivatives of $\mathrm{SE}(\hat{\beta})$ since it is non-linear in the estimated parameters. Further, it could produce inaccurate estimates if higher-order terms of Taylor expansion are important (Eakin et al., 1990).

Another approximate approach is bootstrap, a data-based simulation method for statistical inference proposed by Efron (1979) and first applied to construct confidence interval s for scope economies by Eakin et al.(1990). This method starts from estimating parameters $\hat{\beta}$ and then randomly draw with replacement N by T times from the residuals ε_f (which are the differences between C_f and $C(y,w)$). These newly drew residuals could be used to construct new dependent variables denoted by ${C_f}^*$ and then we re-estimate model (2) R times. This yields R estimates of scope economies, denoted $\mathrm{SE}(\hat{\beta})^{(b)}$ for $b = 1, \dots, R$. The bootstrap estimate of variance of SE could be obtained:

$$\mathrm{Var}(\widehat{\mathrm{SE}}) = \sum_{b=1}^{R} \left(\mathrm{SE}(\hat{\beta})^{(b)} - \mathrm{SE}(\hat{\beta})^{(\cdot)}\right)^2 / (R-1) \tag{8}$$

where $\mathrm{SE}(\hat{\beta})^{(\cdot)} = \sum_{b=1}^{R} \mathrm{SE}(\hat{\beta})^{(b)} / R$ is the mean of $\mathrm{SE}(\hat{\beta})^{(b)}$

This is just a common formula for calculating variance and therefore there is no need to calculate the derivatives of SE since we can replace it with computer power. However, none of the both preceding methods can provide exact estimate for scope economies. The first method relies on a first-order variance approximation; the exact variance could not be calculated through this method (Mester, 1987). The estimate from the other method, bootstrap, has inherent error coming from two distinct sources: sampling variability and bootstrap resampling variability (Efron & Tibshirani, 1986).

From a Frequentist view, a null hypothesis $\mathrm{H_0}$ is usually needed to form certain sampling distribution. In our case, in order to construct the confidence interval for true scope economies (SE), we need to use the definition of scope economies in (2) to propose the following hypothesis:

$$\mathrm{H_0}: SE \leq 0 \ \text{ vs. } \ \mathrm{H_1}: SE > 0 \tag{9}$$

Under the above null hypothesis $\mathrm{H_0}$ and we know $\mathrm{SE}(\hat{\beta})$ is actually the mean estimate of SE from (4), $\mathrm{SE}(\hat{\beta})$ divided by its standard error $\sqrt{\mathrm{Var}\left(\mathrm{SE}(\hat{\beta})\right)}$ would be distributed as a student's t distribution (Mester, 1987) and $100(1-\alpha)\%$ confidence interval for SE at the point $y_1 \dots y_\mathrm{n}$ and w could therefore be approximated by

$$\mathrm{SE}(\hat{\beta}) - t\sqrt{\mathrm{Var}\left(\mathrm{SE}(\hat{\beta})\right)} \leq \mathrm{SE} \leq \mathrm{SE}(\hat{\beta}) + t\sqrt{\mathrm{Var}\left(\mathrm{SE}(\hat{\beta})\right)} \tag{10}$$

where t is the $\alpha/2$ upper percentile of the t distribution; for large N, the equivalent z

percentile could be used.

Notice that the left and right hand side of (10) represent the lower level and upper interval limits in **Error! Reference source not found.** and $\mathrm{Var}\left(\mathrm{SE}(\hat{\beta})\right)$ could be approximated by delta and bootstrap method introduced in early section. However, this process assumes that the sampling distribution of SE is normal or approximately normal. Once this assumption does not hold, misleading results could be produced.

To avoid the assumption of normality, we could fully take non-parametric advantage of bootstrap by taking direct use of the empirical sampling distribution. Therefore, the upper and lower confidence limits for the SE are $100(\alpha/2)$ and $100(1-\alpha/2)$ percentile values of the bootstrap sampling distribution estimate respectively. However, this method still makes a strong assumption that all the bootstrap replicates are unbiased (Hall, 1992). Efron (1987) therefore suggests a bias-corrected and accelerated (BC_a) method to adjust for the bias and skew of the sampling distribution.

2.2 Bayesian inference

From a frequentist view, SE is assumed fixed but unknown. Inferences about SE are based on the distribution of data statistics under repeated sampling as we introduced earlier. In contrast, from a Bayesian view, SE is considered to be a random variable and summarised in a probability distribution which will be derived in the following section[6].

Following the setting of (3), C_f has a normal distribution ($dnorm$), and then the likelihood function for N multiple linear regressions on K variables with balanced panel data could be derived as follows:

$$C_f|y_f, \beta, \tau_C \sim dnorm\left(y_f\beta, \mathbf{I}_T\tau_C\right), for\ \ f = 1 \dots N \tag{11}$$

To conduct a Bayesian approach, we need to assign prior information to each estimated parameters (β, τ_C). We start with the variance of cost τ_C. Without prior information on the variance, a gamma distribution with lower value of shape and scale parameters is a popular non-informative prior for τ_C.

$$\tau_C|a, b \sim dgamma(a, b) \tag{12}$$

We further assume that all coefficients are drawn from the normal distribution ($dnorm$):

6 The different views toward inferences between a frequentist and Bayesian are well documented in literatures (Gelman et al., 2013; Jaynes & Kempthorne, 1976; Little, 2006).

$\beta|\mathrm{c},\mathrm{d} \sim dnorm(c,d)$ (13)

With the prior and likelihood function, we can use Bayes' theorem to derive the posterior density function which is proportional to (∝) the likelihood function and prior distribution.

$$p(\theta|C_f) \propto p(C_f|\theta)p(\theta)$$
$$= \left[\prod_{f=1}^{N} p(C_f|y_f,\beta,\tau_C)\right] p(\beta|\mathrm{c},\mathrm{d})p(\tau_C|a,b) \quad (14)$$

where $\theta \equiv [\beta, \tau_C]$ includes all the parameters of our model.

This posterior density function is conditional on the observed costs, C_f and is also the joint posterior density for all the parameters of our model. The marginal posterior distribution for each parameter could be derived by integrating other parameters out of (14). Then this distribution would be further used for the inferences of each parameter.

With posterior density function in hand, we can apply our proceeding analysis to the formula (4) and predict the corresponding economies measures, $SE(\beta)$.

$$p(SE(\beta)|C_f) = \int p(SE(\beta)|C_f,\theta)p(\theta|C_f)d\theta \quad (15)$$

Here, $p(SE(\beta)|C_f,\theta)$ is the sampling distribution of future cost estimates. That is, we consider these economies measures as observations predicted by the estimated cost function. This predicted value is another dataset generated by the same model and therefore will still have a normal distribution with a mean ESS_{SE} and the same inverse variance (τ_C). In a Bayesian framework, these predictions are carried out by sampling from the posterior predictive distribution as follows.

$$SE(\beta)|C_f \sim dnorm(ESS_{SE},\tau_C) \quad (16)$$

where ESS_{SE}, is derived from using different output mean values.

Its Bayesian confidence interval is determined via

$$\int_L^U p(SE(\beta)|C_f)\, dSE(\beta) = 1-\alpha \quad (17)$$

where $p(SE(\beta)|C_f)$ is the posterior distribution of $SE(\beta)$ conditioned on observed cost C_f derived in (13)

To fairly compare with frequentist confidence intervals calculated from delta method and bootstrap, the credible interval limits U and L are determined using the equal-tailed method (Lu, Ye, & Hill, 2012):

$$Prob\left(SE(\beta) \geq U|C_f\right) = Prob\left(SE(\beta) \leq L|C_f\right) = \alpha/2 \quad (18)$$

One additional advantage of Bayesian approach is that we can further calculate the probability of scope economies:

$$Prob(SE(\beta) > 0|C_f) \quad (19)$$

This is evaluated with respect to the posterior predictive distribution in (16). It is worth noting that for those studies that gave the conclusion based on only point estimate of scope economies, they implicitly assume equation (19) is equal to 100%.

3. Empirical application to Australian university sector

To make our demonstration easy to understand, we limit the number of outputs and firms in the production process to demonstrate the model clearly in the preceding section. We also assume the price is constant across the firms to avoid the rules we need to obey (Chambers, 1988), so we can focus on how to calculate the economic estimates through a Bayesian approach and compare it with other two methods. These conditions lead us to use the dataset from higher education institutions whose main outputs could be reduced to two, that is, teaching and research (Cohn & Cooper, 2004). The assumption of constant price is also reasonable because institutions are usually in the highly regulated contexts (Agasisti & Johnes, 2010).

The Australian university sector provides us an appropriate data set for our model demonstration. This sector has experienced significant changes for the last 30 years. It used to contain many small and specialised higher education institutions (CAEs) but it is no longer the case. Universities are given more responsibilities for enrolling more students and raising their financial revenues (Bradley, 2008). We employ a balanced panel dataset composed of whole 37 Australian public universities over the years 2003-2012 which is the longest periods for public access[7]. These are totally 370 annual observations obtaining from the Australian Government Department of Education. The monetary variable, total cost (total expenses from continuing operations) has been converted to real values (year 2003 = 100) with consumer price index of Australian Bureau of Statistics. Two outputs are included in the cost function.

[7] The data set could be downloaded from http://docs.education.gov.au/ .

Outputs (y_i) consists of total completions (y_1), and number of publications (y_2). Table 1 summaries the average of year-end numbers from 2003 to 2012.

Table 1
Descriptive Statistics of the data over ten years (2003-2012)

Variable	Description	Min.	Max.	Mean	SD
C^a	Total cost	34,355	1,327,778	400,671	280,776
y_1	Number of student completions	546	17,085	6,602	3,551
y_2	Number of publications	60.76	5,118	1,222	1,087

Note:
[a] in 1,000 Australian dollars (year 2003=100);

To associate the production cost with outputs and other variables, a proper function form should be chose to approximate the true cost function $C(y, w)$. Baumol et al.(1982) have suggested three multi-product cost functional forms, Constant Elasticity of Substitution (CES), Hybrid Translog (TL) function, and Quadratic Cost Function (QCF). Each form makes the measurement of scope economies possible and allows returns to scale to vary. These functional forms also correspond to the ideal conditions of cost functional form: flexible enough to allow scale and scope economies to vary with the levels of output, allow data to decide the existence of scale and scope economics, and permit zero outputs of some services (Lloyd, Morgan, & Williams, 1993).

However, the difficulties of imposing linear homogeneity[8] and concavity in input prices in these functions have caused the concerns of violating the cost function assumptions (Baumol, 1982; Gao & Featherstone, 2008; Griffin, Montgomery, & Rister, 1987). More flexible functional forms are proposed (see Pulley and Braunstein (1992)) but the violation of cost function remains especially for imposing concavity in prices (Pulley & Humphrey, 1993).

We have no intention to join the debates of appropriate functional form for estimating cost economies but focus on the methods of constructing confidence interval. We therefore bypass these difficulties by assuming constant input prices across observations. Further, quadratic cost function is preferred in the estimation of scope economies since this cost function permits an output to have zero values without further transformation like other forms.

With our panel data set, we can modify (4) into

[8] Linear homogeneity for this function could be done by omitting one price variable but dividing total cost and other factor prices by the chosen price (Diewert & Wales, 1987; Martínez-Budría, Jara-Díaz, & Ramos-Real, 2003). However, the difficult part is which input is chosen for normalization since there is usually no solid theory to support the decision (Farsi et al., 2008) and estimated results are not invariant to the choices (J. D. Cummins & Weiss, 2013).

$$C_{ft} = \beta_0 + \beta_1 y_{1ft} + \beta_2 y_{2ft} + 0.5(\beta_{11}(y_{1ft})^2 + \beta_{22}(y_{2ft})^2) + \beta_{12}(y_{1ft})(y_{2ft}) + \varepsilon_{ft} \quad (20)$$
$$for\ f = 1 \ldots 37\ and\ t = 1 \ldots 10$$

All the notations follow the previous sections. We can notice that the number of variables, *K*, is 6 ($\beta_{k,k=\{0,1,2,11,22,12\}}$). Our model is fitted using the following non-informative priors. These choices could not only give posterior estimates similar to maximum likelihood point estimates but also offer additional uncertainty over estimates of scale and scope economies as well as the uncertainty of regression coefficients.

Table 2
Non-informative priors used in BUGS

Parameters	**Priors**
β_k	$dnorm$(0, 0.000001)
τ_c	$dgamma$(0.1, 0.01)

Notice that we assume prices are constant across institutions and therefore there is no need to include the price variables in the estimation process. We could get the corresponding posterior predictive distribution for the estimate of scope economies as follows.

$$SE(\beta)|C_{ft} \sim dnorm(ESS_{SE}, \tau_C) \quad (21)$$

where

$$ESS_{SE} = \frac{\beta_0 - \beta_{12} y_1 y_2}{\beta_0 + \beta_1 y_1 + \beta_2 y_2 + 0.5(\beta_{11}(y_1)^2 + \beta_{22}(y_2)^2) + \beta_{12} y_1 y_2}.$$

There is no analytical solution for the posterior density function above but a Gibbs sampler could be used to derive the estimates of the unknown parameters with the prior knowledge (Gelfand, Hills, Racine-Poon, & Smith, 1990). These estimates will be derived with statistical software, BUGS (Bayesian inference using Gibbs sampling) (Lunn, Spiegelhalter, Thomas, & Best, 2009) and it will choose a proper Markov Chain Monte Carlo (MCMC) sampling algorithm based on the model structure we specified above. All inferences are based on two parallel chains running for 50,000 with a burn-in of first 25,000 iterations[9]. The convergence diagnostic indicator, potential scale reduction ($\hat{R}$) (Brooks & Gelman, 1998), is used to monitor the convergence. This unbiased estimator (close to one as the number of draws approach infinity) also has the advantage of reliability compared with purely using plots.

As Table 3 shows, all the estimated parameters are obtained from our Bayesian approach. We

[9] We avoid the thinning of chains in our MCMC sampling to avoid the imprecision inference about the parameters (Link & Eaton, 2012).

obtained $\hat{R} < 1.04$ for all economic estimates indicating good convergence. To compare with this Bayesian model, we use (Nerlove, 1971)' random effect model to minimise the bias from out-of-sample prediction (Farsi et al., 2008; Fetz & Filippini, 2010).

Table 3

	Random Effect model		Bayesian approach	
	Coefficient	Std. Error	Mean	Std. Error
β_0	45770	2.2443	510.00	4.476
β_1	18.161	6.5645	23.42	0.0114
β_2	183.89	2.1633	214.3	0.0549
β_{11}	-0.0022	0.001	-0.0016	0.0000
β_{22}	-0.0432	0.01	-0.0261	0.0000
β_{12}	0.0119	0.0027	0.0056	0.0000

Notice that the positive figure of β_{12} implies sufficient condition of *dis*economies of scope (Hajargasht et al., 2008; Pulley & Humphrey, 1993). Negative values of scope economies are therefore expected from any methods. All estimates of scope economies could vary with output expansions and contractions and we extend the range to 50-150% mean output with random effect model and Bayesian approach in Table 3 Scope economies in delta method and bootstrap are based on the same model (random effect model in Table 3). However, these two methods produce quite different standard deviations of scope economies: the estimated standard deviations from bootstrap are consistently (all levels at means) lower than those from delta method[10]. However, these two methods give the indecisive results below or beyond 100% of output mean level since their intervals cover zero. Their estimates of scope economies also fail to comply the diseconomies of scope (positive figure of β_{12}) in Table 3 In contrast, our Bayesian approach gives the most precise (the least standard deviations) among all approaches. It also the only one to give the consistent results as shown in Table 3

[10] However, (Eakin et al., 1990)found the opposite result.

Table 4
Estimates of scope economies and their corresponding standard error and confidence intervals with delta method, bootstrap and Bayesian approach.

		Delta Method[a]	Bootstrap[a]		Bayesian
			Percentile	BC_a	
Levels	50%	0.0977 (0.1009)	0.5548 (0.0331)	0.5548 (0.0331)	-0.0521 (0.0198)
at means		[-0.101 0.296]	[0.5199 0.6525]	[0.4748 0.5903]	[-0.0895 -0.0147]
	100%	-0.124 (0.0785)	0.2564 (0.0471)	0.2564 (0.0471)	-0.1096 (0.0390)
		[-0.278 0.0301]	[0.2217 0.4086]	[0.1598 0.2820]	[-0.1818 -0.0346]
	150%	-0.285 (0.0905)	0.015 (0.0701)	0.015 (0.0701)	-0.1679 (0.0594)
		[-0.463 -0.107]	[-0.0413 0.2575]	[-0.0957 0.0635]	[-0.2704 -0.0527]

Note: point estimate of scope economies is calculated at output vector = its 50%, 100%, and 150% mean value; values in ( ) shows standard error of scope economies; values in [ ] shows the 95% confidence intervals (that is $\alpha = 5\%$); [a]denotes that estimates are calculated based on random effect model in Table 3 but the estimate from delta method follows the plug-in procedure.

4. Further studies

The main purpose of this study is to develop a new approach for estimating inference measures for scope economies. We demonstrate our model without incorporating the prices for the simplicity. Further studies especially applications in contexts of free market should consider the situation without constant prices across firms. In addition, cost function should obey several rules such as curvature. Fortunately, these restrictions are relatively easy imposed within a Bayesian framework (Griffiths, O'Donnell, & Cruz, 2000; O'Donnell & Coelli, 2005). Another extension of our model is to estimate economic estimates could also vary with L 100% output mean level by multiplying $y_{i,i=\{1,\ldots,n\}}$ and a constant, L. This could further give valuable information about scope modification to achieve cost economies. Final extension comes when researchers have more than one model to predict the economies estimates. Model averaging allows all models to be included in the posterior predictive distribution without deleting any of them (Lancaster, 2004).

5. Reference

Agasisti, T., & Johnes, G. (2010). Heterogeneity and the evaluation of efficiency: the case of Italian universities. *Applied Economics, 42*(11), 1365-1375. doi: 10.1080/00036840701721463

Aitchison, J., & Dunsmore, I. R. (1975). *Statistical prediction analysis*. New York, NY: Cambridge University Press.

Altunbas, Y., & Molyneux, P. (1996). Economies of scale and scope in European banking. *Applied Financial Economics, 6*(4), 367-375.

Baumol, W. J. (1982). Contestable markets: an uprising in the theory of industry structure. *American economic review, 72*(1), 1-15.

Baumol, W. J., Panzar, J. C., & Willig, R. D. (1982). *Contestable markets and the theory of industry structure*. New York, NY: Harcourt Brace Jovanovich.

Berger, A. N., Cummins, J. D., Weiss, M. A., & Zi, H. (2000). Conglomeration versus strategic focus: Evidence from the insurance industry. *Journal of Financial Intermediation, 9*(4), 323-362.

Berger, A. N., Hunter, W. C., & Timme, S. G. (1993). The efficiency of financial institutions: A review and preview of research past, present and future. *Journal of Banking & Finance, 17*(2–3), 221-249.

Box, G. E., & Tiao, G. C. (1992). *Bayesian inference in statistical analysis*. New York, NY: John Wiley & Sons.

Bradley, D. (2008). *Review of Australian higher education: Final report*. Canberra: Department of Education, Employment and Workplace Relations.

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics, 7*(4), 434-455.

Byrnes, J., & Dollery, B. (2002). Do economies of scale exist in Australian local government? a Review of the research evidence. *Urban Policy and Research, 20*(4), 391-414.

Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Pacific Grove, CA: Duxbury

Chambers, R. G. (1988). *Applied production analysis: a dual approach*: Cambridge University Press.

Clark, J. A. (1988). Economies of scale and scope at depository financial institutions: a review of the literature. *Economic Review - Federal Reserve Bank of Kansas City, 73*(8), 16.

Cohn, E., & Cooper, S. T. (2004). Multi-product cost functions for universities: economies of scale and scope. In G. Johnes & J. Johnes (Eds.), *International Handbook on the Economics of Education* (pp. 579-612). Cheltenham, UK: Edward Elgar Publishing Ltd.

Cummins, J. D., & Weiss, M. A. (2013). Analyzing firm performance in the insurance industry using frontier efficiency and productivity methods. In G. Dionne (Ed.), *Handbook of Insurance* (pp. 795-861). New York, NY: Springer

Cummins, J. D., Weiss, M. A., Xie, X., & Zi, H. (2010). Economies of scope in financial services: a DEA efficiency analysis of the US insurance industry. *Journal of Banking & Finance, 34*(7), 1525-1539.

Delgado, M. S., Parmeter, C. F., Hartarska, V., & Mersland, R. (2015). Should all microfinance institutions mobilize microsavings? Evidence from economies of scope. *Empirical Economics, 48*(1), 193-225.

Diewert, W. E., & Wales, T. J. (1987). Flexible Functional Forms and Global Curvature Conditions. *Econometrica, 55*(1), 43-68. doi: 10.2307/1911156

Eakin, B. K., McMillen, D. P., & Buono, M. J. (1990). Constructing Confidence Intervals Using the Bootstrap: An Application to a Multi-Product Cost Function. *The Review of Economics and Statistics, 72*(2), 339-344. doi: 10.2307/2109725

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The annals of Statistics*, 1-26.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American statistical Association, 82*(397), 171-185.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, 54-75.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall/CRC.

Farsi, M., Fetz, A., & Filippini, M. (2008). Economies of scale and scope in multi-utilities sector. *The Energy Journal, 29*(4), 123-143.

Feng, G., & Serletis, A. (2010). Efficiency, technical change, and returns to scale in large US banks: Panel data evidence from an output distance function satisfying theoretical regularity. *Journal of Banking & Finance, 34*(1), 127-138.

Fetz, A., & Filippini, M. (2010). Economies of vertical integration in the Swiss electricity sector. *Energy economics, 32*(6), 1325-1330.

Fields, J. A., & Murphy, N. B. (1989). An analysis of efficiency in the delivery of financial services: The case of life insurance agencies. *Journal of Financial Services Research, 2*(4), 343-356. doi: 10.1007/BF00114412

Gao, Z., & Featherstone, A. M. (2008). Estimating economies of scope using the profit function: A dual approach for the normalized quadratic profit function. *Economics Letters, 100*(3), 418-421.

Geisser, S. (1993). *Predictive Inference: an introduction*. London: CRC Press.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., & Smith, A. F. M. (1990). Illustration of bayesian inference in normal data models using Gibbs sampling. *Journal of the American statistical Association, 85*(412), 972-985.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Chapman and Hall - CRC.

Greene, W. H. (2011). *Econometric analysis* (7 ed.). New York: Prentice Hall.

Griffin, R. C., Montgomery, J. M., & Rister, M. E. (1987). Selecting functional form in production function analysis. *Western Journal of Agricultural Economics, 12*(2), 216-227.

Griffiths, W. E., O'Donnell, C. J., & Cruz, A. T. (2000). Imposing regularity conditions on a system of cost and factor share equations. *Australian Journal of Agricultural and Resource Economics, 44*(1), 107-127. doi: 10.1111/1467-8489.00101

Hajargasht, G., Coelli, T., & Rao, D. S. (2008). A dual measure of economies of scope. *Economics Letters, 100*(2), 185-188.

Hall, P. (1992). *The bootstrap and Edgeworth expansion*. Harrisonburg, VA: Springer.

Held, L., & Bové, D. S. (2014). *Applied statistical inference: likelihood and Bayes*. Berlin: Springer.

Jaynes, E. T., & Kempthorne, O. (1976). Confidence intervals vs Bayesian intervals. In W. Harper & C. Hooker (Eds.), *Foundations of probability theory, statistical inference, and statistical theories of science* (pp. 175-257). Netherlands: Springer.

Kmenta, J., & Rafailzadeh, B. (1997). *Elements of econometrics* (2nd ed.). Ann Arbor: The University of Michigan Press.

Lancaster, T. (2004). *An introduction to modern Bayesian econometrics*. Malden, MA: Blackwell.

Link, W. A., & Eaton, M. J. (2012). On thinning of chains in MCMC. *Methods in Ecology and Evolution, 3*(1), 112-115.

Little, R. J. (2006). Calibrated Bayes: A Bayes/Frequentist Roadmap. *The American Statistician, 60*(3), 213-223.

Lloyd, P., Morgan, M., & Williams, R. (1993). Amalgamations of universities: are there economics of scale or scope? *Applied Economics, 25*(8), 1081-1092.

Lu, D., Ye, M., & Hill, M. C. (2012). Analysis of regression confidence intervals and Bayesian credible intervals for uncertainty quantification. *Water Resources Research, 48*(9), 1-20.

Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: evolution, critique and future directions. *Statistics in Medicine, 28*(25), 3049-3067.

Martínez-Budría, E., Jara-Díaz, S., & Ramos-Real, F. J. (2003). Adapting productivity theory to the quadratic cost function: an application to the Spanish electric sector. *Journal of Productivity Analysis, 20*(2), 213-229.

Mester, L. J. (1987). A multiproduct cost study of savings and loans. *The Journal of Finance, 42*(2), 423-445.

Meyer, R. (2012). Vertical economies and the costs of separating electricity supply-a review of theoretical and empirical literature. *Energy Journal, 33*(4), 161-185.

Nerlove, M. (1971). A note on error components models. *Econometrica, 39*, 383-396.

O'Donnell, C. J., & Coelli, T. J. (2005). A Bayesian approach to imposing curvature on distance functions. *Journal of Econometrics, 126*(2), 493-523.

Ohe, Y., & Kurihara, S. (2013). Evaluating the complementary relationship between local brand farm products and rural tourism: Evidence from Japan. *Tourism Management, 35*, 278-283.

Panzar, J. C., & Willig, R. D. (1981). Economies of scope. *The American Economic Review, 71*(2), 268-272.

Preyra, C., & Pink, G. (2006). Scale and scope efficiencies through hospital consolidations. *Journal of Health Economics, 25*(6), 1049-1068.

Pulley, L. B., & Braunstein, Y. M. (1992). A composite cost function for multiproduct firms with an application to economies of scope in banking. *The Review of Economics and Statistics, 74*(2), 221-230.

Pulley, L. B., & Humphrey, D. B. (1993). The role of fixed costs and cost complementarities in determining scope economies and the cost of narrow banking proposals. *The Journal of Business, 66*(3), 437-462.

Schroeder, T. C. (1992). Economies of Scale and Scope for Agricultural Supply and Marketing Cooperatives. *Review of Agricultural Economics, 14*(1), 93-103.

Suret, J. M. (1991). Scale and scope economies in the Canadian property and casualty insurance industry. *Geneva Papers on Risk and Insurance, 16*(59), 236-256.

Tovar, B., Jara-Díaz, S., & Trujillo, L. (2007). Econometric estimation of scale and scope economies within the port sector: a review. *Maritime Policy & Management, 34*(3), 203-223.

Worthington, A. C., & Higgs, H. (2011). Economies of scale and scope in Australian higher education. *Higher Education, 61*(4), 387-414.