Author/s:
Griffiths, W;Zhang, X;Zhao, X

Title:
Estimation and efficiency measurement in stochastic production frontiers with ordinal outcomes

Date:
2014-08-01

Citation:
Griffiths, W., Zhang, X. & Zhao, X. (2014). Estimation and efficiency measurement in stochastic production frontiers with ordinal outcomes. Journal of Productivity Analysis, 42 (1), pp.67-84. https://doi.org/10.1007/s11123-013-0365-8.

Persistent Link:
https://hdl.handle.net/11343/283077

Estimation and Efficiency Measurement in Stochastic Production Frontiers with Ordinal Outcomes

William Griffiths

Department of Economics, University of Melbourne

Xiaohui Zhang[1]

School of Management and Governance, Murdoch University

Xueyan Zhao

Department of Econometrics and Business Statistics, Monash University

25 July 2013

Corresponding Author:

William Griffiths
Department of Economics
University of Melbourne, Vic 3010
Australia
Email: wegrif@unimelb.edu.au
Phone: +613 8344 3622,
Fax: +613 8344 6899

[1] This work was completed while Zhang was a postdoctoral fellow at Monash University, funded by ARC Discovery Grants DP0880086 and DP0878765.

Abstract

We consider Bayesian estimation of a stochastic production frontier with ordered categorical output, where the inefficiency error is assumed to follow an exponential distribution, and where output, conditional on the inefficiency error, is modelled as an ordered probit model. Gibbs sampling algorithms are provided for estimation with both cross-sectional and panel data, with panel data being our main focus. A Monte Carlo study and a comparison of results from an example where data are used in both continuous and categorical form supports the usefulness of the approach. New efficiency measures are suggested to overcome a lack-of-invariance problem suffered by traditional efficiency measures. Potential applications include health and happiness production, university research output, financial credit ratings, and agricultural output recorded in broad bands. In our application to individual health production we use data from an Australian panel survey to compute posterior densities for marginal effects, outcome probabilities, and a number of within-sample and out-of-sample efficiency measures.

Keywords Bayesian estimation · Gibbs sampling · ordered probit · production efficiency

JEL Classification C11 · C23

1 Introduction

Since its introduction by Aigner *et al*. (1977) and Meeusen and van den Broeck (1977), the stochastic frontier model has been widely used in the analysis of productivity and firm efficiency, and has been extended in numerous directions. Extensions include alternative assumptions for the distribution of the inefficiency error, the use of panel as well as cross-section data, the specification of time-varying inefficiencies, or inefficiencies that are related to firm characteristics, the introduction of heteroskedasticity, the use of dual cost and profit frontiers as well as production frontiers, and multiple output models. Applications have used both firm level and country level datasets, and have evaluated the performance of production units for both traditional and service industries. Estimation has been carried out from both the sampling theory and Bayesian standpoints. Surveys of different aspects of the literature can be found in Bauer (1990), Kim and Schmidt (2000), and Greene (2005). Of particular relevance to this study is Bayesian estimation of the stochastic frontier model, introduced by van den Broeck *et al*. (1994) and surveyed by Koop and Steel (2001). Books with substantial reviews of the literature are Kumbhakar and Lovell (2000) and Coelli *et al*. (2005).

An assumption common to all past studies is that the dependent variable (logarithm of output or cost) is a continuous random variable that is fully observed. In this paper we extend modelling and estimation of the stochastic frontier model to the case where the dependent variable is latent and is observed only as an ordered categorical variable. Requiring special attention in this case is the modelling and estimation of efficiency. In the traditional model where the continuous output variable is observed, the efficiency of a firm is defined unambiguously as the expectation of output conditional on that firm's inefficiency error, relative to the conditional expectation of output given an inefficiency error of zero. When output is a latent variable partially observed through discrete categorical outcomes, this efficiency measure is no longer invariant. It will depend on the normalization setting of an unidentifiable parameter in the ordered probit model. To overcome this problem, we explore several alternative measures of efficiency that do not suffer from a lack of invariance.

The inefficiency error in our model is assumed to follow an exponential distribution, and, conditional on the inefficiency error, categorical output is modelled using an ordered probit model. For estimation, we use a Bayesian approach with a Markov chain Monte Carlo (MCMC) algorithm that is a modification of previous Bayesian algorithms for (1) stochastic frontier models with continuous outputs (van den Broeck, *et al*. 1994), and (2) ordered probit models (Albert and Chib, 1993; Nandram and Chen, 1996; Li and Tobias, 2006). The proposed model has wide applications to broad definitions of 'firm' production (or cost) functions where output (or cost) is observed as ordered categories. The application in this paper, used to illustrate the model, estimation technique, and efficiency measures, is a production frontier for individual health production estimated using panel data from an Australian longitudinal study. The outcome variable is self-assessed health measured as ordered multiple discrete choices over the range: excellent, very good, good, fair and poor. Other potential applications include (1) the production of health, wellbeing and happiness at a country level using panel data on grades from 1 to 10 published by the United Nations' World Happiness Report[2], (2) university research output, graded by the Australian government on a scale between 1 and 5 for both narrow and broad discipline groupings in Australian universities, (3) production of financial credit ratings of AAA, AA, A, BBB, etc. at

[2] http://www.earth.columbia.edu/articles/view/2960

country or bank level, (4) output or cost functions at farm level when only bands of output or cost are provided for privacy reasons, as is the case for data provided by the Australian Bureau of Agricultural and Resource Economics, (5) quality of patient care recorded as a discrete grade in hospitals, and (6) modelling of output that is recorded as counts when there is a desire to avoid the restrictive nature of the Poisson and negative binomial models.

In Section 2 we describe the categorical-output model for panel data and present conditional posterior densities useful for Bayesian estimation of the parameters[3]. Section 3 describes how to obtain posterior densities for probabilities of output categories and their marginal effects. Efficiency measurement is considered in Section 4. We summarize results on efficiency measurement in the continuous output model, explain why these measures are not invariant with respect to reparameterizations of the categorical-output model, and then present some alternative, invariant efficiency measures, for both within-sample and out-of-sample firms. In Section 5, we apply the models and efficiency measures to a production frontier for individual health. The paper concludes in Section 6 with a summary of the contribution and avenues for future research.

2 Modelling and Estimation

2.1 Model specification

We consider a panel data scenario with *N* firms and T_i time period observations on the *i*-th firm, $i = 1, 2, ..., N$. Output for the *i*-th firm in the *t*-th time period y_{it} is a discrete observable random variable that takes one of $J+1$ ordered values from 0 to *J*. As in the

[3] The conditional posterior densities for estimation of the model from cross-sectional data are provided in an earlier version of the paper which is available from the authors upon request.

standard ordered probit model, an unobserved continuous latent variable y_{it}^* can be mapped to the observed discrete outcome y_{it} via some boundary parameters. The latent production output variable y_{it}^* is assumed positive and is related to a $1\times(K+1)$ vector of input variables $\boldsymbol{x}_{it}$, whose first element is unity. Following the typical stochastic frontier model setup, we write the natural logarithm of latent variable y_{it}^* as

$$\ln\ y_{it}^* = f(\boldsymbol{x}_{it},\boldsymbol{\beta}) + v_{it} - u_i, \qquad t = 1,2,\ldots,T_i, \quad i = 1,2,\ldots,N. \tag{1}$$

The $(K+1)\times 1$ vector $\boldsymbol{\beta}$ contains unknown parameters. The v_{it}'s reflect measurement and specification errors. They are assumed to be independent identically distributed symmetric errors following a normal distribution, with mean zero and variance σ_v^2. Specifying a variance σ_v^2 that is not necessarily equal to one differs from the typical textbook treatment of the ordered probit model, but it is in line with specifications that use one less boundary parameter (Nandram and Chen, 1996; Li and Tobias, 2006), and it is a convenient one for the MCMC algorithm that we describe. The u_i's are assumed to be time-invariant, independent identically distributed non-negative error terms, with a given u_i measuring the inefficiency level of firm *i*. The value $u_i = 0$ indicates technical efficiency, while $u_i > 0$ is an indication of technical inefficiency where production of the *i*-th firm lies below the production frontier. While we treat the u_i as random, in classical analysis they have been treated as fixed or random effects. In Bayesian analysis, the difference between fixed and random effect models can be defined through the prior distribution for u_i (Koop *et al*. 1997). A number of authors has extended the basic model by allowing for u_i to be time-varying (see Kumbhakar and Lovell, 2000 for a review). Greene (2004, 2005) has considered a model with random effects for both individual heterogeneity and inefficiency.

In the traditional frontier model where y_{it}^* is observed, and u_i is random and time-invariant, efficiency of the i-th firm is defined as

$$r_i = \frac{E(y_{it}^* \mid u_i)}{E(y_{it}^* \mid u_i = 0)} = \exp(-u_i). \quad (2)$$

As we see later, when y_{it}^* is unobservable and we observe instead an ordered category denoted by y_{it}, the measure in (2) is not invariant with respect to reparameterizations of the model. We develop alternative efficiency measures designed to overcome this problem.

Several one-sided distributions of u_i have been considered in the literature. Early work on the stochastic frontier model presented by Meeusen and van den Broeck (1977) adopts an exponential distribution. Aigner *et al*. (1977) assume u_i follows a half normal distribution. Other suggestions include truncated normal (Stevenson, 1980) and gamma (Greene, 1990) distributions. In this paper, we assume u_i follows an exponential distribution with mean λ and variance λ^2. That is, $p(u_i \mid \lambda) = \lambda^{-1}\exp(-\lambda^{-1}u_i)$. Also we assume that v_{it} and u_i are independent.

Defining $g_{it}^* = \ln(y_{it}^*)$, we write the mapping between the latent variable g_{it}^* and the observed categorical variable y_{it} as

$$
y_{it} = \begin{cases} 0 & g_{it}^* \le 0 \\ 1 & 0 < g_{it}^* \le \gamma_1 \\ \vdots & \\ j & \gamma_{j-1} < g_{it}^* \le \gamma_j \quad (j = 2,\dots, J-2) \\ \vdots & \\ J-1 & \gamma_{J-2} < g_{it}^* \le 1 \\ J & g_{it}^* > 1. \end{cases} \tag{3}
$$

Note that there are $J-2$ unknown boundary or threshold parameters $\gamma_1, \gamma_2, \dots, \gamma_{J-2}$, three less than the number of categories. This parameterization, where $\gamma_{J-1} = 1$ and $\sigma_v^2 \ne 1$, is in line with Nadram and Chen (1996) and Li and Tobias (2006); the traditional ordered-probit parameterization leaves γ_{J-1} unrestricted and sets $\sigma_v^2 = 1$.

In the special case where $f(\boldsymbol{x}_i, \boldsymbol{\beta})$ is linear, for example, a Cobb-Douglas production technology, where we redefine $\boldsymbol{x}_i$ to be the logarithms of the inputs, (1) can be written as

$$
g_{it}^* = \ln(y_{it}^*) = \boldsymbol{x}_{it}\boldsymbol{\beta} + v_{it} - u_i. \tag{4}
$$

It is this linear specification for which we describe a MCMC algorithm. It is convenient at this point to introduce some matrix notation. Let $\boldsymbol{y}_i$, $\boldsymbol{g}_i^*$ and $\boldsymbol{v}_i$ be $T_i \times 1$ vectors and $\mathbf{X}_i$ a $T_i \times (K+1)$ matrix, containing T_i observations for individual *i*. Further, we define $\boldsymbol{y} = (\boldsymbol{y}_1', \boldsymbol{y}_2', \dots, \boldsymbol{y}_N')'$, $\boldsymbol{g}^* = (\boldsymbol{g}_1^{*\prime}, \boldsymbol{g}_2^{*\prime}, \dots, \boldsymbol{g}_N^{*\prime})'$ and $\boldsymbol{v} = (\boldsymbol{v}_1', \boldsymbol{v}_2', \dots, \boldsymbol{v}_N')'$ as $S \times 1$ vectors, and $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2', \dots, \mathbf{X}_N')'$ as an $S \times (K+1)$ matrix, where $S = \sum_{i=1}^{N} T_i$. We also let $\boldsymbol{u} = (u_1, u_2, \dots, u_N)'$ and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_{J-2})'$.

Our immediate objective is to describe how Bayesian estimation of this model can be carried out. The first step in this direction is to specify conditional posterior densities which can be used in a Gibbs sampling algorithm to draw values from the joint posterior density of the unknown parameters $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and λ, and the latent variables $\boldsymbol{g}^*$ and $\boldsymbol{u}$. Conditional posterior densities are also specified for an extension where λ is allowed to vary over firms' characteristics. Later, we describe how the parameter draws can be used to get posterior densities on other quantities of interest such as probabilities for each level of production conditional on $\boldsymbol{x}_{it}$, and various measures of efficiency.

2.2 Conditional posterior densities

In this section, we present a Gibbs sampling algorithm that combines the Nadram-Chen algorithm for the ordered probit model (Nadram and Chen, 1996) and an algorithm for the stochastic frontier model with continuous output (Koop, *et al*., 1997; Koop and Steel, 2001). Both g_{it}^* and the inefficiency error u_i are treated as unknown parameters with values being drawn from their conditional posterior densities. Then, conditional on $\boldsymbol{g}^*$ and $\boldsymbol{u}$, the stochastic frontier model for ordinal outcomes reduces to the standard linear regression model, facilitating draws from the conditional posterior densities for the parameters.

For a prior density we assume all parameters are *a priori* independent with $p\left(\boldsymbol{\beta},\boldsymbol{\gamma}\right) \propto 1$, $p\left(\sigma_v^2\right) \propto 1/\sigma_v^2$, and $\lambda \sim IG(a_\lambda, b_\lambda)$, where $IG(a,b)$ denotes the inverted-gamma density with parameters a and b. Specifically, if $z \sim IG(a,b)$, then $p(z\,|\,a,b) = \left[b^a/\Gamma(a)\right] z^{-a-1}\exp(-b/z)$. If only cross-sectional data are available, a proper inverted-gamma density needs to be specified for σ_v^2 to obtain proper posteriors (Fernández *et al*. 1997).

Under these prior distributions and the assumed model specification, the joint posterior density for all unknown parameters and the latent variables $\boldsymbol{g}^*$ and $\boldsymbol{u}$ can be written as

$$p\left(\boldsymbol{\beta},\sigma_v^2,\boldsymbol{\gamma},\boldsymbol{u},\lambda,\boldsymbol{g}^* \mid \mathbf{X},\boldsymbol{y}\right) \propto \left[\prod_{i=1}^{N}\prod_{t=1}^{T_i} I\left(\gamma_{y_{it}-1} < g_{it}^* \le \gamma_{y_{it}}\right)\right]\left[\prod_{i=1}^{N} I\left(u_i \ge 0\right)\right]$$
$$\times \frac{1}{\left(\sigma_v^2\right)^{S/2+1}\lambda^{N+a_\lambda+1}} \exp\left\{-\frac{1}{2\sigma_v^2}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(g_{it}^* - \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i\right)^2 - \frac{1}{\lambda}\left(\sum_{i=1}^{N} u_i + b_\lambda\right)\right\}. \tag{5}$$

In (5) $I(\cdot)$ denotes an indicator function, and we adopt the notation $\gamma_{-1} = -\infty$, $\gamma_0 = 0$, $\gamma_{J-1} = 1$ and $\gamma_J = +\infty$. The remaining thresholds that make up the vector $\boldsymbol{\gamma}$ are unknown parameters.

From the joint posterior density we can derive the following conditional posterior densities that can be used for Gibbs sampling. The conditional posterior densities for the latent variables g_{it}^* are independent truncated normal distributions

$$g_{it}^* \mid \boldsymbol{\beta},\sigma_v^2,\boldsymbol{\gamma},\boldsymbol{u},\lambda,\mathbf{X},\boldsymbol{y} \sim N\left(\boldsymbol{x}_{it}\boldsymbol{\beta} - u_i, \sigma_v^2\right) I\left(\gamma_{y_{it}-1} < g_{it}^* \le \gamma_{y_{it}}\right), \tag{6}$$

The conditional posterior density for $\boldsymbol{\beta}$ is the normal distribution

$$\boldsymbol{\beta} \mid \sigma_v^2,\boldsymbol{\gamma},\boldsymbol{u},\lambda,\boldsymbol{g}^*,\mathbf{X},\boldsymbol{y} \sim N\left(\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\left(\boldsymbol{g}^* + \tilde{\boldsymbol{u}}\right), \sigma_v^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right), \tag{7}$$

where $\tilde{\boldsymbol{u}}'$ is the $1\times S$ vector, $\tilde{\boldsymbol{u}}' = \left(u_1\boldsymbol{\iota}_{T_1}', u_2\boldsymbol{\iota}_{T_2}', \ldots, u_N\boldsymbol{\iota}_{T_N}'\right)$ and $\boldsymbol{\iota}_{T_i}$ is a $T_i \times 1$ vector of ones. The conditional posterior density for σ_v^2 is the inverted gamma distribution

$$\sigma_v^2 \mid \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{u},\lambda,\boldsymbol{g}^*,\mathbf{X},\boldsymbol{y} \sim IG\left(\frac{S}{2},\frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(g_{it}^* - \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i\right)^2\right). \tag{8}$$

To obtain the conditional posterior for $\boldsymbol{\gamma}$, we first note that for $J \le 2$, there are no unknown threshold parameters. For $J \ge 3$, rather than using (5) to derive the conditional posterior density for $\boldsymbol{\gamma}$, it is more convenient to consider the joint posterior density obtained without the introduction of the latent vector $\boldsymbol{g}^*$ as unknown parameters. It is given by

$$p\left(\boldsymbol{\beta},\boldsymbol{\gamma},\lambda,\boldsymbol{u},\sigma_v^2 \mid \boldsymbol{y},\mathbf{X}\right) \propto \prod_{i=1}^{N}\prod_{t=1}^{T_i}\left[\Phi\left(\frac{\gamma_{y_{it}} - \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i}{\sigma_v}\right) - \Phi\left(\frac{\gamma_{y_{it}-1} - \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i}{\sigma_v}\right)\right]$$
$$\times\left[\prod_{i=1}^{N} I\left(u_i \ge 0\right)\right]\frac{1}{\sigma_v^2\lambda^{N+a_\lambda+1}}\exp\left\{-\frac{1}{\lambda}\left[b_\lambda + \sum_{i=1}^{N} u_i\right]\right\}. \quad (9)$$

The conditional posterior density for $\boldsymbol{\gamma}$ derived from (9) is

$$p\left(\gamma \mid \boldsymbol{\beta},\sigma_v^2,\boldsymbol{u},\lambda,\mathbf{X},\boldsymbol{y}\right) \propto \prod_{i=1}^{N}\prod_{t=1}^{T_i}\left\{\Phi\left(\frac{\gamma_{y_{it}} - \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i}{\sigma_v}\right) - \Phi\left(\frac{\gamma_{y_{it}-1} - \boldsymbol{x}_{it}\boldsymbol{\beta} + u_i}{\sigma_v}\right)\right\}. \quad (10)$$

Because (10) is not a density function whose form is recognizable, a Metropolis-Hastings step is used to draw from it. A Dirichlet proposal density is constructed for $p\left(\boldsymbol{\gamma} \mid \boldsymbol{\beta},\sigma_v^2,\boldsymbol{u},\lambda,\mathbf{X},\boldsymbol{y}\right)$ in the following way. Let the differences between adjacent thresholds be defined as $q_j = \gamma_j - \gamma_{j-1}$ for $j = 1,2,\ldots,J-1$, and let $\boldsymbol{q} = (q_1,q_2,\ldots,q_{J-1})'$. Then, $q_j \ge 0$, and $\sum_{j=1}^{J-1} q_j = 1$, making the Dirichlet distribution a possible proposal density for $\boldsymbol{q}$. Its density is given by

$$p\left(\boldsymbol{q} \mid \boldsymbol{\alpha},\boldsymbol{y}\right) \propto \prod_{j=1}^{J-1} q_j^{\alpha_j n_j - 1}, \quad (11)$$

where $0 \le \alpha_j \le 1,\ j = 1,2,\ldots,J-1$, are tuning parameters, and $n_j = \sum_{i=1}^{N}\sum_{t=1}^{T_i} I(y_{it} = j)$ is the number of sample observations in category *j*. The advantages of the proposal density (11) are that the entire vector $\boldsymbol{q}$ can be drawn at once, and it does not depend on $\boldsymbol{\beta}$ and

σ_v^2. The tuning parameters, α_j $(j = 1, 2, \ldots, J-1)$ are chosen to make the dispersion of the distribution of $\boldsymbol{q}$ comparable to or at least as large as that of the posterior distribution of $\boldsymbol{\gamma}$. To perform the Metropolis-Hastings step, a set of candidate values $\boldsymbol{q}^{can}$ is drawn from $p(\boldsymbol{q} | \boldsymbol{\alpha}, \boldsymbol{y})$, and transformed to a set of candidate values $\boldsymbol{\gamma}^{can}$. Given values $\boldsymbol{\gamma}$ from the previous iteration, the vector $\boldsymbol{\gamma}^{can}$ is accepted with probability $a = \min\{R, 1\}$ where

$$R = \left\{ \prod_{i=1}^{N} \prod_{t=1}^{T_i} \frac{\Phi\left[\left(\gamma_{y_{it}}^{can} - \boldsymbol{x}_i \boldsymbol{\beta}\right)/\sigma_v\right] - \Phi\left[\left(\gamma_{y_{it}-1}^{can} - \boldsymbol{x}_i \boldsymbol{\beta}\right)/\sigma_v\right]}{\Phi\left[\left(\gamma_{y_{it}} - \boldsymbol{x}_i \boldsymbol{\beta}\right)/\sigma_v\right] - \Phi\left[\left(\gamma_{y_{it}-1} - \boldsymbol{x}_i \boldsymbol{\beta}\right)/\sigma_v\right]} \right\} \left\{ \prod_{j=1}^{J-1} \left(\frac{q_j}{q_j^{can}} \right)^{\alpha_j n_j - 1} \right\}. \quad (12)$$

Values $\boldsymbol{\gamma}$ from the previous iteration are accepted with probability $1 - a$.

The remaining conditional posterior densities are the truncated normal distribution for the inefficiency error u_i

$$u_i \,|\, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_v^2, \lambda, \boldsymbol{g}^*, \mathbf{X}, \boldsymbol{y} \sim N\left(\bar{\boldsymbol{x}}_i \boldsymbol{\beta} - \bar{g}_i^* - (T_i \lambda)^{-1} \sigma_v^2,\; T_i^{-1} \sigma_v^2\right) I(u_i \geq 0), \quad (13)$$

where $\bar{\boldsymbol{x}}_i$ and $\bar{g}_i^*$ are the respective means of $\boldsymbol{x}_{it}$ and g_{it}^* over the T_i observations for firm *i*, and the inverted-gamma density for λ

$$\lambda \,|\, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{u}, \sigma_v^2, \boldsymbol{g}^*, \mathbf{X}, \boldsymbol{y} \sim IG\left(N + a_\lambda,\; \sum_{i=1}^{N} u_i + b_\lambda \right). \quad (14)$$

2.3 Generalising the inefficiency term

In line with the literature for the stochastic frontier model with a continuous output variable, in this section we extend our model to allow the inefficiency term u_i to be related to explanatory variables (Kumbhakar and Lovell, 2000, Ch.7). Of interest is whether firms with some special characteristics are more likely to be more efficient than others. Our model specification and algorithm for Bayesian estimation follow that in Koop *et al*. (1997).

Suppose there are m time-invariant variables, w_{ik} $(i=1, 2,\ldots,N;\ k=1, 2,\ldots,m)$, that impact on the efficiency of firms, where $w_{i1}\equiv 1$ is a constant and w_{ik} $(k=2,\ldots,m)$ are dummy variables representing firm characteristics. Continuous w variables can be introduced, but only at a cost of computational complexity. Define $\mathbf{W}=(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_N)'$ as an $N\times m$ matrix. Assume that u_i follows the exponential distribution:

$$p\left(u_i\,|\,\boldsymbol{\phi},\mathbf{W}\right)=\lambda_i^{-1}\exp(-\lambda_i^{-1}u_i),\qquad \lambda_i^{-1}=\prod_{k=1}^{m}\phi_k^{w_{ik}}, \tag{15}$$

where $\boldsymbol{\phi}'=\left(\phi_1,\phi_2,\ldots,\phi_m\right)$ are unknown parameters. Since the mean of the inefficiency distribution λ_i is always positive, the ϕ_k should all be positive. If $\phi_k\equiv 1$ for $k=2,\ldots,m$, $\lambda_i^{-1}\equiv\phi_1$ is a constant and the model collapses to the standard model in equation (4). Otherwise, the ϕ_k $(k=2,\ldots,m)$ are to be estimated and the magnitude of a ϕ_k (in particular whether $\phi_k<1$ or $\phi_k>1$) determines whether the attribute w_k is a "bad" or "good" attribute in terms of its contribution to mean inefficiency. Since the kth term enters the product in (15) as $\phi_k{}^{w_k}=1$ if $w_k=0$, and $\phi_k{}^{w_k}=\phi_k$ if $w_k=1$, firms with attribute w_k will have a higher mean inefficiency λ_i if $\phi_k<1$ (i.e., w_k is a "bad" attribute), and a lower mean inefficiency if $\phi_k>1$ (w_k is a "good" attribute). Note that $\phi_k<1$ does not mean that a firm with attribute w_k is definitely more inefficient than those without this characteristic, but rather that the former has an inefficiency u_i drawn from a distribution with a higher mean, assuming all other characteristics are the same. It is also relevant to note that the good-bad interpretation of the $\phi_h \ \ h \geq 2$ is invariant with respect to scale transformations of u_i. Such transformations will change the constant coefficient ϕ_1, but not the remaining coefficients.

For a prior distribution for the new parameters $\boldsymbol{\phi}$, we use independent gamma priors $\phi_k \sim G(a_k, b_k)$ and $p(\boldsymbol{\phi}) = p(\phi_1) \cdots p(\phi_m)$; $G(a,b)$ denotes the gamma density with parameters a and b. With this prior, the conditional posterior densities for $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, σ_v^2, and $\boldsymbol{g}^*$ remain the same as in the model without $\mathbf{W}$. The conditional posterior density for u_i in (13) is replaced by

$$u_i \mid \boldsymbol{\phi}, \boldsymbol{g}^*, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_v^2, \boldsymbol{y}, \mathbf{X}, \mathbf{W} \ \sim N \ \ \bar{\mathbf{x}}_i \boldsymbol{\beta} - \bar{g}_i^* - \ T_i \lambda_i \ ^{-1} \sigma_v^2, \ T_i^{-1} \sigma_v^2 \ \ I \ \ u_i \geq 0 \qquad (16)$$

The new parameters, ϕ_k ($k = 1, 2, \ldots, m$), can be drawn from

$$\phi_k \mid \boldsymbol{u}, \boldsymbol{\phi}^{(-k)}, \boldsymbol{g}^*, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_v^2, \boldsymbol{y}, \mathbf{X}, \mathbf{W} \ \sim G\left(a_k + \sum_{i=1}^{N} w_{ik}, \ b_k + \sum_{i=1}^{N} \left(w_{ik} u_i \prod_{s \neq k} \phi_s^{w_{is}} \right) \right). \qquad (17)$$

where $\boldsymbol{\phi}^{(-k)}$ denotes the vector $\boldsymbol{\phi}$ with the k-th component omitted.

2.4 Evaluating the Methodology

Two questions about the above methodology naturally arise. Does it produce satisfactory estimates of true underlying parameters, and how do its estimates compare with those obtained when continuous output is observed? To answer these questions we performed a small Monte Carlo experiment and also examined how continuous-output estimates compared with categorical-output estimates in a single sample of data.

For the Monte Carlo study we generated repeated samples from the data generating process in (4). An unbalanced panel was used with $N = 5{,}000$ and a maximum $T_i = 4$, leading to a total of 15,042 observations for each sample. We set

$X = (1, x_1, x_2, x_3)$, with two dummy variables $x_1 = I(U[0,1] > 0.45)$ and $x_3 = I(U[0,1] > 0.25)$, and one continuous variable $x_2 = \ln(U[0,100])$, where $U[a,b]$ denotes the uniform distribution on the interval $[a,b]$. The parameter values were set as $\beta = (0.2,\ 0.55,\ -0.05,\ 0.4)'$, $\sigma_v = 0.5$ and $\lambda = 0.5$, with thresholds $\gamma = (0, 0.3,\ 0.6, 1)'$, defining 5 ordered categories. The simulation was carried out for 200 replications. The burn-in for the MCMC was taken as 2,000 iterations and the number of total recorded iterations after the burn-in was 5,000. The results presented in Table 1 are reassuring. The Monte Carlo averages of the posterior means are almost identical to the true parameter values, suggesting the posterior means are unbiased estimators. Moreover, the Monte Carlo standard deviations of the posterior means are relatively small, implying low repeated-sample variability of the posterior mean. They are also similar to the Monte Carlo averages of the posterior standard deviations, suggesting that the posterior standard deviation is a good reflection of the repeated sampling variability. However, the standard deviations of the thresholds are an exception; in this case the posterior standard deviation understates the repeated-sample variability of the posterior mean.

The precision of the estimates in the Monte Carlo experiment suggest that estimation with ordinal categorical data can perform favourably relative to estimation when continuous output data are available. However, it is of interest to examine how different the two sets of estimates might be in a single sample of data. To make this comparison we estimated both models using some Philippine data on rice production (see Hill et al 2011, Exercise 5.24, p.209). In this example $N = 44$ and $T = 8$, and output is a function of area, labour and fertiliser, with all variables measured in logs. For estimating with categorical data, 5 categories were established with thresholds for

$\ln(y^*)$ given by 0, $\ln(2)=0.69$, $\ln(4)=1.39$, and $\ln(10)=2.30$. The results from the 2 estimations are presented in Table 2. Those reported for the categorical output model have been adjusted to accommodate the fact that the parameters are only identified relative to a fixed scale. Since the normalisation used by our estimation procedure is to set the upper threshold to 1, and the true setting is 2.3, the adjustment we made was to multiply all posterior means and standard deviations by 2.3. In practice, when only ordinal data are available, and the thresholds are not known, this adjustment cannot be made. The normalisation $\gamma_{J-1}=1$ can be retained, or it might be convenient to rescale in another way. With the exception of λ, the two sets of posterior means in Table 2 are of similar magnitudes. The posterior standard deviations from the categorical output are larger than their continuous output counterparts, reflecting differences in the amount of information available and the number of parameters being estimated.

The parameter draws obtained using the algorithms described in the previous subsections and evaluated using the Monte Carlo experiment in this section can be used to estimate various characteristics of the posterior distributions of the parameters. However, we are typically more interested in various functions of the parameters $\boldsymbol{\beta},\gamma,\sigma^2,\lambda$ and $\boldsymbol{\phi}$, than in the parameters themselves. Two such functions are the probabilities of each ordered outcome for given values of ***x*** and ***w***, and the marginal effects of changes in an *x* or a *w* on those probabilities. Also of interest are the efficiencies of various firms – those within the sample, and those outside the sample with specific characteristics. In Section 3 we give expression for probabilities and their marginal effects. In Section 4 we suggest a number of possible efficiency measures particularly suited to the stochastic frontier model with ordinal output.

3 Probabilities and marginal effects

There are two output probabilities for out-of-sample firms that might be of interest. The first is the probability that a randomly selected firm with characteristics $(\boldsymbol{x}_s, \boldsymbol{w}_s)$ has output in category *j*. Letting $\theta = (\beta', \gamma', \sigma_v, \phi')'$, we can write this probability as

$$\Pr\left(y_s = j \mid \boldsymbol{x}_s, \boldsymbol{w}_s, u_s, \boldsymbol{\theta}\right) = \Phi\left(\frac{\gamma_j - \mathbf{x}_s\boldsymbol{\beta} + u_s}{\sigma_v}\right) - \Phi\left(\frac{\gamma_{j-1} - \mathbf{x}_s\boldsymbol{\beta} + u_s}{\sigma_v}\right). \qquad (18)$$

Draws from the posterior-predictive density for this probability can be obtained by drawing a value u_s from the exponential density $p(u_s \mid \boldsymbol{\phi}, \boldsymbol{w}_s)$ for each MCMC draw for $\boldsymbol{\phi}$, and then computing (18) for all draws of (θ, u_s).

The second output probability of interest is the average probability of achieving output in category *j* for all out-of-sample firms with characteristics $(\boldsymbol{x}_s, \boldsymbol{w}_s)$. It is given by the expectation of (18) with respect to the distribution of u_s. That is,

$$\Pr\left(y_s = j \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}\right) = \int_0^{+\infty} \left(\Phi\left(\frac{\gamma_j - \mathbf{x}_s\boldsymbol{\beta} + u_s}{\sigma_v}\right) - \Phi\left(\frac{\gamma_{j-1} - \mathbf{x}_s\boldsymbol{\beta} + u_s}{\sigma_v}\right) \right) p\left(u_s \mid \boldsymbol{\theta}\right) du_s. \quad (19)$$

Draws from the posterior density for this probability can be obtained by numerically evaluating the integral in (19) for each MCMC draw of $\boldsymbol{\theta}$.

Posterior densities for the marginal effects corresponding to (18) and (19) can also be obtained. For example, the marginal effect on $\Pr\left(y_s = j \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}\right)$ of a change in a continuous covariate, say x_{sk}, evaluated at settings $(\boldsymbol{x}_s, \boldsymbol{w}_s)$, is given by

$$\frac{\partial \Pr\left(y_s = j \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}\right)}{\partial x_{sk}} = \int_0^{+\infty} \left(\phi_{SN}\left(\frac{\gamma_{j-1} - \mathbf{x}_s\boldsymbol{\beta} + u_s}{\sigma_v}\right) - \phi_{SN}\left(\frac{\gamma_j - \mathbf{x}_s\boldsymbol{\beta} + u_s}{\sigma_v}\right) \right) \frac{\beta_k}{\sigma_v} p\left(u_s \mid \boldsymbol{\theta}\right) du_s$$

(20)

Like before, draws from the posterior density for this quantity can be obtained by numerically evaluating the integral in (20) for each MCMC draw of $\boldsymbol{\theta}$.

To obtain the posterior density for the marginal effect of any binary variable d in $\boldsymbol{x}$, where $\boldsymbol{x}_s^*, \boldsymbol{w}_s$ denotes the settings of all other variables at which the effect of d is evaluated, we compute draws from

$$\Pr\left(y = j \mid d = 1, \boldsymbol{x}_s^*, \boldsymbol{w}_s, \boldsymbol{\theta} \right) - \Pr\left(y = j \mid d = 0, \boldsymbol{x}_s^*, \boldsymbol{w}_s, \boldsymbol{\theta} \right), \tag{21}$$

When no variables influence the distribution of the inefficiency error, $\boldsymbol{w}_s$ is omitted from (18)-(21), $p(u_s \mid \boldsymbol{\phi}, \boldsymbol{w}_s)$ becomes $p(u_s \mid \lambda)$, and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \sigma_v, \lambda)'$.

4. Efficiency measures

Two main aims of traditional production frontier analysis are (1) to evaluate and rank the efficiencies of all firms in the sample, given observed input and output levels of these firms, and (2) to assess the efficiency of out-of sample firms with particular characteristics. In this section we begin by briefly summarizing relevant posterior density functions for efficiency measurement in the stochastic production frontier model with continuous observable output. We then explain why these posterior densities are not invariant with respect to reparameterizations when output is categorical. Finally, in Sections 4.3 to 4.5, we propose some alternative efficiency measures which do not suffer from a lack of invariance when output is categorical.

4.1 Efficiency measurement in the traditional model

In traditional stochastic production frontier analysis where y_{it}^* is observable, a convenient measure of efficiency is that given in equation (2), namely, $r_i = \exp(-u_i)$. Related to r_i, are three posterior densities likely to be of interest. They are (1) the efficiency of a within-sample firm, (2) the efficiency of an out-of-sample firm with particular characteristics, and (3) the average efficiency of all out-of-sample firms with particular characteristics.

Noting that when y_{it}^* is observable, $g_{it}^* = \ln(y_{it}^*)$ is also observable, the posterior density for the efficiency of a particular within-sample firm can be written as

$$p\left(r_i \mid \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) = \int p\left(r_i \mid \boldsymbol{\theta}, \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) p\left(\boldsymbol{\theta} \mid \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) d\boldsymbol{\theta}, \tag{22}$$

where

$$p\left(r_i \mid \boldsymbol{\theta}, \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) = \frac{I(0 < r_i \le 1)}{\sqrt{2\pi}\,\sigma_{u_i} r_i \Phi\left(\mu_{u_i} / \sigma_{u_i}\right)} \exp\left\{-\frac{1}{2}\left(\frac{-\ln r_i - \mu_{u_i}}{\sigma_{u_i}}\right)^2\right\}, \tag{23}$$

with $\mu_{u_i} = \bar{x}_i\beta - \bar{g}_i^* - \left(T_i\lambda_i\right)^{-1}\sigma_v^2$, and $\sigma_{u_i}^2 = \sigma_v^2 / T_i$. The conditional posterior density in (23) is obtained by transforming the truncated normal density for u_i in (16) to that for $r_i = \exp(-u_i)$. Its first and second moments can be shown to be

$$E\left(r_i \mid \boldsymbol{\theta}, \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) = \frac{\Phi\left(\mu_{u_i} / \sigma_{u_i} - \sigma_{u_i}\right)}{\Phi\left(\mu_{u_i} / \sigma_{u_i}\right)} \exp\left(\frac{\sigma_{u_i}^2}{2} - \mu_{u_i}\right), \tag{24}$$

and

$$E\left(r_i^2 \mid \boldsymbol{\theta}, \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) = \frac{\Phi\left(\mu_{u_i} / \sigma_{u_i} - 2\sigma_{u_i}\right)}{\Phi\left(\mu_{u_i} / \sigma_{u_i}\right)} \exp\left(2\sigma_{u_i}^2 - 2\mu_{u_i}\right). \tag{25}$$

The unconditional posterior density $p\left(r_i \mid \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right)$ and its moments can be obtained by averaging (23), (24) and (25) over MCMC draws for $\boldsymbol{\theta}$.

For an out-of-sample firm with characteristics $\boldsymbol{w}_s$ and corresponding inefficiency error u_s, that is a drawing from an exponential distribution with inverse mean $\lambda_s^{-1} = \prod_{k=1}^{m} \phi_k^{w_{sk}}$, the density function for efficiency $r_s = \exp(-u_s)$ is

$$p\left(r_s \mid \boldsymbol{w}_s,\boldsymbol{\phi}\right) = \lambda_s^{-1} r_s^{\lambda_s^{-1}-1}\, I\left(0 < r_s \le 1\right). \tag{26}$$

Its first and second moments are $E\left(r_s \mid \boldsymbol{w}_s,\boldsymbol{\phi}\right) = \left(\lambda_s + 1\right)^{-1}$ and $E\left(r_s^2 \mid \boldsymbol{w}_s,\boldsymbol{\phi}\right) = \left(2\lambda_s + 1\right)^{-1}$, respectively. These results are different from those for a within-sample firm because we no longer condition on the firm's $\boldsymbol{g}^*$ and $\boldsymbol{x}$ values which are not observed. However, the sample values $\boldsymbol{g}^*$, $\mathbf{X}$ and $\mathbf{W}$ provide information on $\boldsymbol{\phi}$ through its posterior density which is used to obtain the Bayesian predictive density

$$p\left(r_s \mid \boldsymbol{w}_s, \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) = \int p\left(r_s \mid \boldsymbol{w}_s,\boldsymbol{\phi}\right) p\left(\boldsymbol{\phi} \mid \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right) d\boldsymbol{\phi}. \tag{27}$$

This density and its moments can be obtained by averaging $p(r_s|\boldsymbol{w}_s,\boldsymbol{\phi})$, $E(r_s|\boldsymbol{w}_s,\boldsymbol{\phi})$ and $E(r_s^2|\boldsymbol{w}_s,\boldsymbol{\phi})$ over the MCMC draws for $\boldsymbol{\phi}$.

The third efficiency measure likely to be of interest is the average efficiency of out-of-sample firms with particular attributes $\boldsymbol{w}_s$. The relevant posterior density is that for $E\left(r_s \mid \boldsymbol{w}_s,\boldsymbol{\phi}\right) = \left(\lambda_s + 1\right)^{-1}$ which can be estimated from the MCMC draws for $\boldsymbol{\phi}$. In this case variation comes only from the uncertainty in $\boldsymbol{\phi}$. The posterior mean of $\left(\lambda_s + 1\right)^{-1}$ is the same as the mean of the predictive density $p\left(r_s \mid \boldsymbol{w}_s, \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right)$, but the posterior variance of $\left(\lambda_s + 1\right)^{-1}$ will be much smaller than that of $p\left(r_s \mid \boldsymbol{w}_s, \mathbf{X}, \mathbf{W}, \boldsymbol{g}^*\right)$ because it does not include the randomness of selecting a particular firm.

4.2 Lack of invariance in efficiency measures

When we move from a model where the log of output g_{it}^* is fully observed to one where it is a latent variable that maps output into a number of discrete categories according to some threshold parameters, the efficiency measures described in the previous section will be different for different parameterizations of the same model. To appreciate this fact, consider again the panel data frontier model

$$g_{it}^* = \ln(y_{it}^*) = \boldsymbol{x}_{it}\boldsymbol{\beta} + v_{it} - u_i, \quad v_{it} \sim N(0, \sigma_v^2),$$
$$p(u_i|\lambda_i) = \lambda_i^{-1}\exp(-\lambda_i^{-1}u_i), \quad \lambda_i^{-1} = \phi_1 \prod_{k=2}^{m} \phi_k^{w_{ik}}, \tag{28}$$

where the category $y_{it} = j$ is observed when

$$\gamma_{j-1} < g_{it}^* \leq \gamma_j \quad j = 0, \ldots, J, \text{ with } \gamma_{-1} = -\infty,\ \gamma_0 = 0 \text{ and } \gamma_J = +\infty.$$

If g_{it}^* is observed, then no restriction on σ_v^2 or on one of the threshold parameters $\gamma_1, \gamma_2, \ldots, \gamma_{J-1}$ is needed for identification. As described in the previous section, efficiency is measured as $\exp(-u_i)$. When y_{it} but not g_{it}^* is observed, a restriction is needed for identification. Suppose we use the restriction $\sigma_v^2 = 1$. Imposing this restriction is equivalent to reparameterizing the model as

$$\tilde{g}_{it}^* = \boldsymbol{x}_{it}\tilde{\boldsymbol{\beta}} + \tilde{v}_{it} - \tilde{u}_i, \quad \tilde{v}_{it} \sim N(0,1),$$
$$p(\tilde{u}_i|\tilde{\lambda}_i) = \tilde{\lambda}_i^{-1}\exp(-\tilde{\lambda}_i^{-1}\tilde{u}_i), \quad \tilde{\lambda}_i^{-1} = \tilde{\phi}_1 \prod_{k=2}^{m} \phi_k^{w_{ik}}, \tag{29}$$

where the category $y_{it} = j$ is observed when $\tilde{\gamma}_{j-1} < \tilde{g}_{it}^* \leq \tilde{\gamma}_j$, with $\tilde{g}_{it}^* = g_{it}^*/\sigma_v$, $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}/\sigma_v$, $\tilde{v}_{it} = v_{it}/\sigma_v$, $\tilde{u}_i = u_i/\sigma_v$, $\tilde{\lambda}_i = \lambda_i/\sigma_v$, $\tilde{\phi}_1 = \phi_1/\sigma_v$, and $\tilde{\gamma}_j = \gamma_j/\sigma_v$. The efficiency measure becomes $\exp(-\tilde{u}_i) = \exp(-u_i/\sigma_v)$.

Similarly, if we achieve identification by setting $\gamma_{J-1} = 1$, the restriction used in our estimation procedure, the reparameterized inefficiency error becomes $\tilde{u}_i = u_i/\gamma_{J-1}$, and the efficiency measure becomes $\exp(-\tilde{u}_i) = \exp(-u_i/\gamma_{J-1})$. Using (u_i/γ_{J-1}) will give different efficiency measurements to those obtained using (u_i/σ_v) in the alternative parameterization. Also, any other threshold value could be set to unity to achieve identification, leading again to a different efficiency measure.

To avoid these problems, in what follows we consider a number of scale-invariant measures of efficiency. In Section 4.3 we suggest two possible relative efficiency measures for a within-sample firm, where efficiency is measured relative to a reference firm. In Section 4.4 we consider relative efficiency between two out-of-sample firms. A measure based on the probability of being in the highest output category is considered in Section 4.5.

4.3 Relative efficiency of within-sample firm

When assessing the efficiency of a within-sample firm relative to the remainder of the firms in the sample, two obvious comparisons come to mind: How does a firm compare to (1) the most efficient firm in the sample, and (2) a firm with "average" efficiency? For the first comparison, a natural efficiency measure is

$$RE_{i;best} = \frac{\min_i\{u_i\}}{u_i}. \tag{30}$$

Because $u_i \geq 0$, we have $0 < RE_{i;best} \leq 1$, with $RE_{i;best} = 1$ for the most efficient firm in the sample, and $RE_{i;best} \to 0$ when $u_i \to +\infty$.

For a comparison with an average firm in the sample, we suggest the measure

$$RE_{i;med} = \frac{u_{median}}{u_i + u_{median}}, \tag{31}$$

where u_{median} is the median u_i in the sample. This measure also satisfies $0 < RE_{i;med} \leq 1$. When $u_i = 0$, $RE_{i;med} = 1$, indicating the firm is most efficient; when $u_i \to +\infty$,

$RE_{i;med} \to 0$, indicating a least efficient firm; and when $u_i = u_{median}$, $RE_{i;med} = 0.5$, implying a firm of median efficiency.

The posterior densities for $RE_{i;best}$ and $RE_{i;med}$ and their moments can be estimated from the MCMC draws for u_i. Note that $\min_i\{u_i\}$, u_{median}, and u_i will all change in each iteration of the MCMC algorithm.

4.4 Relative efficiency for two out-of-sample firms

When considering the efficiency of out-of-sample firms, we are more likely to be interested in how the efficiency of a firm with characteristics $\boldsymbol{w}_s$ (say) compares with another firm with characteristics $\boldsymbol{w}_c$. Their inefficiency errors u_s and u_c follow exponential distributions with densities

$$p\left(u_j \mid w_j, \boldsymbol{\phi}\right) = \lambda_j^{-1} \exp\left(-u_j \lambda_j^{-1}\right), \quad \text{where } \lambda_j^{-1} = \phi_1 \prod_{k=2}^{m} \phi_k^{w_{jk}}, \quad j = s,\ c. \tag{32}$$

We define the relative efficiency of firm s to c as

$$RE_{s,c} = \frac{u_c}{u_s}. \tag{33}$$

A value $RE_{s,c} > 1$ indicates firm s is more efficient than firm c; $RE_{s,c} < 1$ implies s is less efficient than c; and $RE_{s,c} = 1$ means firms s and c are equally efficient. From (32) and (33), the density function for $RE_{s,c}$ can be derived as (see Appendix A)

$$p\left(RE_{s,c} \mid \boldsymbol{w}_s, \boldsymbol{w}_c, \boldsymbol{\phi}\right) = \frac{\lambda_c^{-1} \lambda_s^{-1}}{\left(\lambda_c^{-1} RE_{s,c} + \lambda_s^{-1}\right)^2}. \tag{34}$$

Because this density does not depend on ϕ_1, which is the only parameter in (32) which changes when the scale changes, it is invariant with respect to scale transformations. The moments of the density in (34) do not exist, but the median is equal to

$$\text{Median}\left(RE_{s,c} \mid \boldsymbol{w}_s, \boldsymbol{w}_c, \boldsymbol{\phi}\right) = \frac{\lambda_c}{\lambda_s} = \prod_{k=2}^{m} \phi_k^{(w_{sk} - w_{ck})}. \tag{35}$$

To estimate the Bayesian predictive density for $RE_{s,c}$, we average $p\left(RE_{s,c} \mid \boldsymbol{w}_s, \boldsymbol{w}_c, \boldsymbol{\phi}\right)$ over the MCMC draws for $\boldsymbol{\phi}$, for a grid of $RE_{s,c}$ values. The posterior density for the median in (35), and its moments, can also be found from the MCMC draws for $\boldsymbol{\phi}$.

Equation (34) can be used to highlight the effect of a single attribute on efficiency. Consider two firms with exactly the same $\boldsymbol{w}$ except for one attribute w_{k*}; that is, $w_{sk} = w_{ck}$ for $k \neq k*$, $w_{sk*} = 1$ and $w_{ck*} = 0$. Then, denoting the resulting relative efficiency by $RE_{s,c}^{(k*)}$, equations (34) and (35) become

$$p\left(RE_{s,c}^{(k*)} \mid \boldsymbol{w}_s, \boldsymbol{w}_c, \boldsymbol{\phi}\right) = \frac{\phi_{k*}}{\left(RE_{s,c}^{(k*)} + \phi_{k*}\right)^2}, \qquad k* = 2, 3, \ldots, m, \tag{36}$$

$$\text{Median}\left(RE_{s,c}^{(k*)} \mid \boldsymbol{w}_s, \boldsymbol{w}_c, \boldsymbol{\phi}\right) = \phi_{k*}. \tag{37}$$

Equation (37) gives us an interpretation for the coefficients $\boldsymbol{\phi}$. For example, suppose $\phi_{k*} = 2$. Then we can say that a firm without characteristic w_{k*} will, on average, be twice the distance from the frontier as a firm with characteristic w_{k*}. Here we use the term "on average" loosely; half the population without characteristic w_{k*} will be more than double the distance from the frontier and half will be less than double the distance.

4.5 Relative probability as an efficiency measure

Another way to measure efficiency for a particular firm, and one which has parallels with the traditional definition of efficiency in equation (2) is to examine

$$r_{sj} = \frac{\Pr\left(y_s = j \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}, u_s\right)}{\Pr\left(y_s = j \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}, u_s = 0\right)}. \tag{38}$$

This ratio gives the probability of being in category j for a firm with characteristics $(\boldsymbol{x}_s, \boldsymbol{w}_s)$ and inefficiency error u_s, relative to the probability for a firm that is on the frontier, and has the same characteristics. It could be evaluated for all $j = 0,1,\ldots,J$. We expect it to be less than one for $j = J$ and greater than one for $j = 0$. Whether it is greater or less than one for the intermediate categories is less clear. Since we are primarily interested in efficiency relative to a best-case scenario, we consider r_{sJ}, the ratio of probabilities for the highest output category, as our measure of efficiency.

Note that probability statements about output are invariant to alternative parameterizations. Thus, although r_{sJ} depends on only one inefficiency error u_s, it is free from the scaling problem that arises when using simple functions of u_s, like $\exp(-u_s)$. Another characteristic of r_{sJ} as an efficiency measure is that it depends not only on the variables $\boldsymbol{w}_s$ that influence the mean of the inefficiency error u_s, but also on the variables $\boldsymbol{x}_s$ that appear in the mean function for output, $E\left(y_s^* \mid u_s\right)$.

We can use r_{sJ} to assess the efficiency of a within-sample firm or an out-of-sample firm. For a within sample firm *i*, we set $y_s = y_{it}$, $\boldsymbol{x}_s = \boldsymbol{x}_{it}$, $\boldsymbol{w}_s = \boldsymbol{w}_i$, $u_s = u_i$, and $r_{sJ} = r_{itJ}$; the efficiency measure will be different in each time period because $\boldsymbol{x}_{it}$ changes over time. In an empirical study it may be useful to define the efficiency of the *i*-th firm as the average $r_{iJ} = T_i^{-1}\sum_{t=1}^{T_i} r_{itJ}$. Using (21) and recalling that $\gamma_J = \infty$ and $\gamma_{J-1} = 1$, we can write

$$r_{itJ} = \frac{1-\Phi\left(\dfrac{1-\boldsymbol{x}_{it}\boldsymbol{\beta}+u_i}{\sigma_v}\right)}{1-\Phi\left(\dfrac{1-\boldsymbol{x}_{it}\boldsymbol{\beta}}{\sigma_v}\right)}. \tag{39}$$

We can estimate the posterior density for r_{itJ} from the MCMC draws for $\boldsymbol{\beta}$, σ_v, and u_i.

For an out-of-sample firm with characteristics ($\boldsymbol{x}_s, \boldsymbol{w}_s$), the efficiency measure r_{sJ} is given by (39) with $\boldsymbol{x}_{it}$ replaced by $\boldsymbol{x}_s$, and u_i replaced by u_s. In this case u_s has not been realized and we are interested in the density function for r_{sJ}. It can be derived from that for u_s as (see Appendix A)

$$p\left(r_{sJ} \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}\right) = \frac{\sigma_v p_0 \lambda_s^{-1} \exp\left(-\lambda_s^{-1}\left[\boldsymbol{x}_s\boldsymbol{\beta} - 1 + \sigma_v \Phi^{-1}\left(1 - r_{sJ} p_0\right)\right]\right)}{\phi_{SN}\left(\Phi^{-1}\left(1 - r_{sJ} p_0\right)\right)}, \tag{40}$$

where $p_0 = 1 - \Phi\left(\left(1 - \boldsymbol{x}_{it}\boldsymbol{\beta}\right)/\sigma_v\right)$. The predictive density $p\left(r_{sJ} \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \mathbf{X}, \mathbf{W}, \boldsymbol{y}\right)$ can be estimated by averaging (40) over the MCMC draws for $\boldsymbol{\theta}$ for a grid of r_{sJ} values, or by drawing u_s from $p\left(u_s \mid \boldsymbol{\theta}\right)$ for each MCMC draw of $\boldsymbol{\theta}$, computing the corresponding values of r_{sJ} from (39), and constructing a histogram. The average of these r_{sJ} values is an estimate of the mean of the predictive density and a point estimate of efficiency. The absence of closed form expressions for the moments of $p\left(r_{sJ} \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}\right)$ prevents us from getting this estimate by averaging $E\left(r_{sJ} \mid \boldsymbol{x}_s, \boldsymbol{w}_s, \boldsymbol{\theta}\right)$ over the draws $\boldsymbol{\theta}$.

5 An application to health production of individuals

The proposed model is applicable to any production or cost function model where the dependent variable is observed as ordered categories. These include the analyses of discrete classifications of broader types of ‘firms’ where performance evaluation and ranking may be of interest. Examples are universities producing education and research, hospitals producing healthcare, banks producing financial services, and countries producing income or financial or political stability. In this section, we use an example of individual health production to illustrate our model. Although we are less interested in ranking individuals in this example, we still illustrate and compare efficiency measures for individual "producers" as these are generally useful in applications where performance ranking of firms may be of interest.

5.1 Data and model specification

In this case what is typically a "firm" becomes an individual. The outcome variable is self-reported health, classified into five categories, and collected through the question: "*In general, would you say your health is: Excellent (4), Very good (3), Good (2), Fair (1), or Poor (0)"*. Individual health production functions of this type, using self-reported health as an overall indicator of health status, and relating it to a variety of inputs, have been estimated in the health economics literature (see for example Desai, 1987; Contoyannis and Jones, 2004), but such studies have not been concerned with efficiency measurement and have not been estimated within a stochastic frontier framework. Efficiency measurement within the health economics literature has so far been confined to macro functions with continuous outcomes, examining either the production of health care and the benchmarking of hospital performance (e.g., Gerdtham *et al.*, 1999; Rosko, 2001; Brown, 2003; Puig-Junoy and Orton, 2004), or the health output of countries, measured as continuous variables such as mortality rate or life expectancy (Puig-Junoy, 1998; Evan *et al.*, 2000; Thornton, 2002; Fayissa and Gutema, 2005).

The data used in this application are from the first fives waves of the Household, Income and Labour Dynamics in Australia (HILDA) surveys conducted from 2001 to

2005. See http://melbourneinstitute.com/hilda/. These nationally representative longitudinal surveys contain detailed information on demographic, socioeconomic, geographic and lifestyle characteristics of individuals, as well as individual health status and health related behaviour. For this study, our sample is restricted to respondents 18 years or older, which involves 65,449 records. After removal of missing values, a sample of 53,164 records for 15,450 individuals was used. It is an unbalanced panel data set. The proportions of the sample reporting each category of health status were poor 3.41%, fair 13.97%, good 34.44%, very good 35.60%, and excellent 12.58%.

It is not straightforward to decide what variables should be considered production inputs $\boldsymbol{x}$ and what factors $\boldsymbol{w}$ should influence inefficiency. In the production frontier literature, variables influencing the structure of the technology by which conventional inputs are converted to outputs are usually included in $\boldsymbol{x}$, while firm characteristics judged to influence the efficiency of conversion are related to the inefficiency term (Pitt and Lee 1981; Kumbhakar, et al, 1991; Battese and Coelli, 1995; Kumbhakar and Lovell, 2000, Ch.7). As first pointed out by Deprins and Simar (1989), it is not always obvious whether a particular variable affects the technology directly or is a determinant of efficiency, and this is frequently a judgement call for particular applications (Kumbhakar and Lovell, 2000). To illustrate, we consider two alternative model specifications, with and without variables $\boldsymbol{w}$ that effect the mean of u.

In Model 1, we specify production input variables $\boldsymbol{x}$ to include lifestyle factors, living location, social connectedness, employment and education. Personal demographic characteristics, including gender, marital status, country of birth and age bands, are assumed to influence the efficiency of health production via the variables $\boldsymbol{w}$. Definitions for all variables specified in the vectors $\boldsymbol{x}$ and $\boldsymbol{w}$ are given in Appendix B.

As these specification for $\boldsymbol{x}$ and $\boldsymbol{w}$ may seem arbitrary, we also consider, as Model 2, an alternative specification where all exogenous variables in $\boldsymbol{x}$ and $\boldsymbol{w}$ are included in the frontier function and there are no regressors influencing u. Our results suggest that the two different specifications do not significantly affect the distributions for efficiency measures or individuals' rankings. We recognize that possible endogeneity of some variables may mean the example is not an ideal one. For example, amount of exercise and degree of loneliness may depend on health. For some evidence on the effect of endogeneity on efficiency estimates in the traditional model, see Mutter et al. (2012).

5.2 Results for estimated parameters and marginal effects

For the gamma prior densities on the elements of $\boldsymbol{\phi}$ for Model 1, we set $a_1 = 1,\ b_1 = 0.36$ and $a_k = 2$, $b_k = 2$, $k = 2, 3, \ldots, m$, to yield relatively noninformative priors. For priors on other parameters for both models, we used $p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_v^2) \propto 1/\sigma_v^2$. The burn-in period was taken as 2,000 iterations and the number of recorded iterations after the burn-in was 10,000. For assessing mixing performance, we plotted the MCMC sample paths for selected parameters. These graphs suggested the sample paths were well mixed. There was no evidence of nonconvergence. The posterior means and standard deviations for all parameters are summarised in Table 3 for both models. The parameters $\boldsymbol{\beta}$ represent the effects of the $\boldsymbol{x}$ variables on latent health. Results for the marginal effects of all $\boldsymbol{x}$ variables on the probabilities of health levels are presented in Table 4 for both models. They are evaluated at the sample means of all variables.

The first point to note is that the marginal effects from the two models are very similar. In other words, regardless whether individual characteristics ($\boldsymbol{w}$) are assumed to affect the frontier function directly or to affect inefficiency, the effects of $\boldsymbol{x}$ on health status do not changed significantly. Note next that the signs of the marginal effects are generally as expected, and the posterior standard deviations are relatively small, with most 95% credibility intervals (not presented to save space) not containing zero. For example, examining some of the posterior means for Model 1, we find that people who have never smoked regularly are 1.59% less likely to produce poor health and 4.67% more likely to produce excellent health. Education level, as an important proxy of socioeconomic status and health behaviour, is a positive input in health production. Compared with those who have less than year 12 education, people with a higher degree are 2.17% less likely to report poor health and 6.93% more likely to have excellent health. These two numbers are 0.5% and 1.18% for persons with a diploma and 1.51% and 4.25% for those with year 12 education.

5.3 Efficiency measures for within-sample individuals

As examples of the invariant within-sample efficiency measures, we consider efficiency relative to the median, $RE_{i;med} = u_{median} \big/ \; u_i + u_{median}$, and the relative probability measure r_{iJ} obtained by averaging (39) over time. In Figures 1 and 2, we present histograms for the posterior means of $RE_{i;med}$ and r_{iJ}, respectively, for all individuals in the sample for both Model 1 and Model 2. We note that the distributions do not change much with the two model specifications.

Assuming we take the posterior mean as an estimate of relative efficiency, Figure 1 shows a wide range of relative efficiency estimates in the sample, varying from about 0.16 to 0.94. The distribution of the estimates will depend on the sample distribution of the variables in $\boldsymbol{w}$. Given these variables are dummy variables, and include variables for age groupings, the multi-modal nature of the distribution is likely to be attributable to the varying sample numbers in each of the categories defined by the dummy variables. In Figure 2, because the frequency of means close to zero is very large relative to frequencies in bins away from zero, we graphed the histograms in two segments, one from 0 to 0.05 and the other from 0.05 to 1, with a different scale for each segment, but the same bin width. The concentration of means towards zero is surprising at first. However, the proportion of individuals reporting excellent health in the raw sample is only 0.1258, and so one would expect the probability of reporting excellent health to be very low for many individuals.

Next, to illustrate how the posterior densities for efficiency measures for individual 'firms' can be used to assess and rank firm performance, we chose from Model 1 three individuals: a "best" or most efficient individual who had the highest posterior mean for $RE_{i;med}$ (0.9345), a "worst" or least efficient individual who had the lowest posterior mean (0.1591), and a median individual who had the median posterior mean (0.5280). Estimates of the posterior densities for each of these individuals are presented in Figure 3. The best person has a relative efficiency ranging from 0.6 to 1, and only a small probability of having a relative efficiency less than 0.8. Possible efficiency values for the worst person are concentrated around 0.2, and the range of possible values is relatively small compared with those for the median person, who has possible relative efficiency values from 0.3 to 0.9.

In Figure 4 we graph the posterior densities for r_{iJ} from Model 1 for the same three individuals that were considered in Figure 3. In this case efficiency is defined as the probability of an individual in the sample having excellent health relative to the probability of that same individual having excellent health if they were on the frontier

$(u = 0)$. This ratio depends not only on an individual's values for the $\boldsymbol{w}$ variables but also on their values for the $\boldsymbol{x}$ variables. The posterior density for the least efficient person is concentrated heavily at zero, with a mean of 1.7×10^{-7}. That for the most efficient person has a mean of 0.8446 and is such that efficiency values from 0.3 to 1 are possible. The efficiency of the intermediate person ranges from 0 to 0.8, with most of the probability less than 0.25, and a mean of 0.1532. These results give a much grimmer picture of efficiency than those from $RE_{i;med}$. They suggest that, using this definition, the efficiency of most of the sample could be very poor. This conjecture is supported by examining Figure 2 that contains the histogram of the posterior means of r_{iJ} for all individuals in the sample.

A final point to note regarding in-sample individual efficiency measures is that whether regressors are included in the inefficiency term, or as part of the frontier function, has little effect on the efficiency ranking of individual 'firms'. For example, for the same three individuals we chose in Model 1, the best person is also the most efficient in Model 2, the worst person is the third worst in Model 2, and the median person in Model 1 is not far from the median in Model 2 (ranked 8116th rather than the median 7725th). The rank correlation coefficient for the $RE_{i;med}$ measures from the two models is 0.976. This result is reassuring for the robustness of the efficiency ranking to alternative specifications of the location of the regressors.

5.4 Efficiency measures for out-of-sample individuals

When using Model 1, we may also be interested in the efficiency of out-of-sample individuals with specific characteristics $\boldsymbol{w}$. In Section 4.4, we showed that the coefficients ϕ_k represent the medians of distributions of relative efficiency. Specifically,

ϕ_k is the median of the distribution for the relative efficiency of someone with characteristic *k*, relative to someone without characteristic *k*, holding all other $\boldsymbol{w}$ variables constant. The posterior means, standard deviations and 95% credibility intervals for each of the ϕ_k are given in Table 3.

Taking the effect of the firm characteristic gender on efficiency as an example, in Figure 5 we graph the posterior density for median relative efficiency of female to male. We see that, "on average", females are slightly more efficient than males. The posterior mean for the median is 1.065, and the spread of the distribution suggests the median lies somewhere between 1.02 and 1.12. If we consider not just the median of relative efficiency, but the complete predictive density of relative efficiency for a female and a male selected at random, we introduce a great deal more uncertainty. Figure 6 contains the graph of the density for the relative efficiency of females to males, with all other characteristics being held constant. The spread is much greater than that for the distribution of the median in Figures 5. For values less than 1, females are less efficient than males; the probability for this occurrence is $\Pr\left(RE_{females;males} < 1\right) = 0.48$.

Our final out-of-sample invariant efficiency measure is r_{sJ}, the probability of an individual with specified characteristics $\left(\boldsymbol{x}_s, \boldsymbol{w}_s\right)$ having excellent health relative to the same probability for the same person, but with an inefficiency error of zero. The predictive density for r_{sJ} for a person selected at random from the population of all people with a given $\left(\boldsymbol{x}_s, \boldsymbol{w}_s\right)$ is given by equation (40) averaged over MCMC draws for $\boldsymbol{\theta}$. Two of these predictive densities are plotted in Figure 7. One is for an individual with the worst $\boldsymbol{x}$ and the worst $\boldsymbol{w}$, and the other is for someone with the best $\boldsymbol{x}$ and the best $\boldsymbol{w}$, where best and worst were chosen by examining the posterior mean estimates for

β and ϕ, and choosing the variable categories accordingly. Both densities cover the complete range from 0 to 1. That for the worst x, w is highest at zero, and then decreases monotonically, with most of its probability less than 0.2. That for best x, w has an asymmetric "U" shape, declining for values of r_{sJ} from zero to approximately 0.18, and then increasing over the remaining range up to $r_{sJ} = 1$. Most of the probability is at the upper end of the distribution, but the "U" shape means that efficiency defined in this way can be very low, even for someone with favourable characteristics.

6 Conclusion

We present a stochastic frontier model for discrete ordinal outcomes for both cross-sectional and panel data. The model is a meaningful extension of the stochastic frontier model with a continuous output variable. It has potential applications in many fields where there is a need for performance evaluation of organisations and firms, or when survey data that comes in the form of discrete categories is to be used. Gibbs sampling with data augmentation is adopted as the posterior simulator, adapting and marrying algorithms suggested previously in the literature for the ordered probit and traditional stochastic frontier models. Posterior distributions for quantities of interest, including probabilities of outcome status, the marginal effects of inputs on output status, efficiency measures for within-sample and outside-sample individuals are presented. We show that the efficiency measure used in conventional frontier models for continuous dependent variables is no longer identifiable in the categorical-dependent variable case, and we propose several efficiency measures, in particular relative efficiencies and efficiency in terms of the probability for the 'best' category, that are uniquely identified and invariant to the unidentifiable scaling parameter. We also show that the inefficiency measures and rankings for individual firms are robust to alternative ways of incorporating exogenous regressors into the model. The model is illustrated with a health production analysis using panel data from the HILDA survey. The marginal effects of inputs on the probabilities of health status are used to present the impact of inputs on health production output. A range of efficiency measures for within and outside individuals are presented as posterior density functions.

Our extension of the stochastic frontier model to discrete ordered dependent variables is based on traditional stochastic frontier models, in the spirit of Battese and Coelli (1988, 1995), Kumbhakar *et al*. (1991), and Koop *et al*. (1997). Greene (2004, 2005) discusses the issue of distinguishing between individual heterogeneity and inefficiency in stochastic frontier analysis. He examined several extensions to the commonly used stochastic frontier model specifications for panel data to allow for more flexibility in accommodating firm heterogeneity while preserving the inefficiency measurement feature of the frontier models. These include the 'true' fixed and random effect models that have both the traditional fixed/random individual-specific term, as typically used in panel data linear regression models, as well as the one-sided inefficiency error term. He also presented random coefficient and latent class versions of the stochastic frontier model for isolating individual heterogeneity. In Greene's (2004, 2005) context, our model has allowed for individual heterogeneity to affect both the production function and the inefficiency term via observable time-invariant characteristics; it does not separately identify individual heterogeneity and inefficiency due to unobservable factors. Allowing for these possibilities is a potential avenue for future research.

Appendix A: Density functions for relative efficiency

To derive the density function for $RE_{s,c} = u_c / u_s$ given in (34) we use the short hand notation $r = RE_{s,c}$ and transform the variables as $r = u_c / u_s$ and $z = u_s$, where

$$p(u_j) = \lambda_j^{-1} \exp\left(-u_j \lambda_j^{-1}\right) \qquad j = s, c.$$

Thus, $u_c = rz$ and

$$\begin{aligned} p(r,z) &= p(u_c, u_s) \begin{vmatrix} \partial u_c / \partial r & \partial u_c / \partial z \\ \partial u_s / \partial r & \partial u_s / \partial z \end{vmatrix} \\ &= \lambda_c^{-1} \lambda_s^{-1} \exp\left(-\lambda_c^{-1} rz - \lambda_s^{-1} z\right) \begin{vmatrix} z & r \\ 0 & 1 \end{vmatrix} \\ &= z \lambda_c^{-1} \lambda_s^{-1} \exp\left(-z\left(\lambda_c^{-1} r + \lambda_s^{-1}\right)\right). \end{aligned}$$

The density for r is given by

$$p(r) = \lambda_c^{-1} \lambda_s^{-1} \int_0^{+\infty} z \exp\left(-z\left(\lambda_c^{-1} r + \lambda_s^{-1}\right)\right) dz = \frac{\lambda_c^{-1} \lambda_s^{-1}}{\left(\lambda_c^{-1} r + \lambda_s^{-1}\right)^2}.$$

Now we consider the density function for

$$r_{sJ} = \frac{1 - \Phi\left(\dfrac{1 - \boldsymbol{x}_s \boldsymbol{\beta} + u_s}{\sigma_v}\right)}{p_0}, \tag{A1}$$

where $p_0 = 1 - \Phi\left((1 - \boldsymbol{x}_s \boldsymbol{\beta}) / \sigma_v\right)$. We wish to derive the density for r_{sJ} from that for u_s which is given by $p(u_s) = \lambda_s^{-1} \exp\left(-u_s \lambda_s^{-1}\right)$. Solving (A1) for u_s yields $u_s = \boldsymbol{x}_s \boldsymbol{\beta} - 1 + \sigma_v \Phi^{-1}\left(1 - r_{sJ} p_0\right)$. Then, noting that, if $x = \Phi^{-1}(p)$, then $dx/dp = 1 / \phi_{SN}\left(\Phi^{-1}(p)\right)$, we have $du_s / dr_{sJ} = -\sigma_v p_0 / \phi_{SN}\left(\Phi^{-1}\left(1 - r_{sJ} p_0\right)\right)$. Thus,

$$p(r_{sJ}) = p(u_s) \left| \frac{du_s}{dr_{sJ}} \right| = \frac{\sigma_v p_0 \lambda_s^{-1} \exp\left(-\lambda_s^{-1}\left[\boldsymbol{x}_s \boldsymbol{\beta} - 1 + \sigma_v \Phi^{-1}\left(1 - r_{sJ} p_0\right)\right]\right)}{\phi_{SN}\left(\Phi^{-1}\left(1 - r_{sJ} p_0\right)\right)}.$$

Appendix B: Definition of variables

Variables	Definition
y	
SRH	self-reported health, 0 for poor, 1 for fair, 2 for good, 3 for very good and 4 for excellent
x	
LT3EX	1 if doing exercise for less than 3 times but at least 1 time per week and 0 otherwise
MT3	1 if doing exercise for more than 3 times per week, including doing exercise every day and 0 otherwise
NOEX	1 if doing no exercise at all and 0 otherwise. This variable is used as the base for exercise level and is dropped off in the estimation
NOSM	1 if never smoke and 0 otherwise
LRA	1 if having low alcohol risk [a] or alcohol risky and 0 otherwise
HIGHRA	1 if having high alcohol risk and 0 otherwise
NORA	1 if having no alcohol risk and 0 otherwise. This variable is used as the base for alcohol risk and is dropped off in the estimation
LONELY1	1 if sometime feel lonely [b] and 0 otherwise
LONELY2	1 if always feel lonely and 0 otherwise.
LONELY0	1 if never feel lonely and 0 otherwise. This variable is used as the base for social net work and is dropped off in the estimation
INNER	1 if living in inner region of Australia and 0 otherwise
OUTER	1 if living in outer region of Australia and 0 otherwise
REMOTE	1 if living in remote region of Australia and 0 otherwise
MAJOR	1 if living in major cities of Australia and 0 otherwise. This variable is used as the base for living region and is dropped off in the estimation
STUDENT	1 if full time study and 0 otherwise
PARTTIME	1 if part-time employed and 0 otherwise
UNEMP	1 if unemployed and 0 otherwise
RETD	1 if completely retired from labour market and 0 otherwise
NOTINLAB	1 if not in labour force and 0 otherwise
FULLTIME	1 if full-time employed and 0 otherwise. This variable is used as the base for major activity and is dropped off in the estimation
DEGREE	1 if the highest qualification is a tertiary degree and 0 otherwise
DIPLOMA	1 if the highest qualification is diploma or trade certificate and 0 otherwise
YEAR12	1 if the highest qualification is Year 12 and 0 otherwise
LOWER12	1 if still in school or cannot finish Year 12 and 0 otherwise. This variable is used as the base for education level and is dropped off in the estimation
w	
GENDER	1 for male and 0 for female
AUSABO	1 if born in Australia and aboriginal and 0 otherwise
MAINENG	1 if born in other main English speaking countries and 0 otherwise
OTHERC	1 if born in other countries rather than Australia and main English speaking countries and 0 otherwise
AUSNABO	1 if born in Australia and not aboriginal and 0 otherwise. This variable is used as the base for country born status and is dropped off in the estimation
MARRIAGE	1 if living with somebody in a relationship for most of the time periods and 0 otherwise
AGEG1	1 if aged from 18 to 24 and 0 otherwise
AGEG2	1 if aged from 25 to 34 and 0 otherwise
AGEG3	1 if aged from 35 to 44 and 0 otherwise
AGEG4	1 if aged from 45 to 54 and 0 otherwise
AGEG5	1 if aged from 55 to 64 and 0 otherwise
AGEG6	1 if aged 65 or over and 0 otherwise. This variable is used as the base for age band and is dropped off in the estimation

NOTE: (*a*) Generally, no alcohol risk means 0 standard drinks per week; low alcohol risk means, for males, 1-6 standard drinks per week, or, for females, 1-4 standard drinks per week; high alcohol risk means, for males, at least 7 standard drinks per week, for females, at least 5 standard drinks per week.

(*b*) The information on loneliness is collected through the statement, 'I often feel very lonely'. Respondents are assigned a number from 1 to 7 representing from strongly disagree to strongly agree; that is, the higher the number the individual chooses, the more she or he agrees with the statement. We re-classify the respondents into three groups as never feel lonely (1 or 2), sometimes feel lonely (from 3 to 5) and always feel lonely (6 or 7).

References

Aigner, D.J., Lovell, C.A.K., Schmidt P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6: 21-37.

Albert, J.H., Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* 88: 669-679.

Battese, G.E., Coelli, T.J. (1988). Prediction of farm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* 38: 387-399.

Battese, G.E., Coelli, T.J. (1995). A model for technical efficiency effects in a stochastic frontier production function with panel data. *Empirical Economics* 20: 325-332.

Bauer, P.W. (1990). Recent developments in the econometric estimation of frontiers. *Journal of Econometrics* 46: 39-56.

Brown, H.S. (2003). Managed care and technical efficiency. *Health Economics* 12: 149-158.

Coelli, T., O'Donnell, C.J., Battese, G.E., Prasada Rao, D.S. (2005). *An Introduction to Efficiency and Productivity Analysis*, second edition. New York: Springer.

Contoyannis, P., Jones, A.M. (2004). Socio-econornic status, health and lifestyle. *Journal of Health Economics* 23: 965-995.

Deprins, D., Simar, L. (1989). Estimating technical inefficiencies with correction for environmental conditions. *Annals of Public and Cooperative Economics* 60: 81-102.

Desai, S. (1987). The estimation of the health production function for low-income working men. *Medical Care* 25: 604-615.

Evan, D., Tandon, A., Murray, C., Lauer, J. (2000). The comparative efficiency of national health systems in producing health: an analysis of 191 countries. World Health Organization, GPE discussion paper, No. 29, EIP/GPE/EQC.

Fayissa, B., Gutema, P. (2005). Estimating a health production function for Sub-Saharan Africa (SSA). *Applied Economics* 37: 155-164.

Fernández, C., Osiewalski, J., Steel, M.F.J. (1997). On the use of panel data in stochastic frontier models with improper priors. *Journal of Econometrics* 79: 169-173.

Gerdtham, U.G., Lothgren, M., Tambour, M., Rehnberg, C. (1999). Internal markets and health care efficiency: a multiple output stochastic frontier analysis. *Health Economics* 8: 151-164.

Greene, W. (1990). A gamma-distributed stochastic frontier model. *Journal of Econometrics* 46: 141-163.

Greene, W. (2004). Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics* 13: 959-980.

Greene, W. (2005). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis* 23: 7-32.

Hill, R.C., Griffiths, W.E., Lim, G.C. (2011), *Principles of Econometrics, Fourth Edition*, New York: John Wiley and Sons.

Kim, Y., Schmidt, P. (2000). A review and empirical comparison of Bayesian and classical approaches to inference on efficiency levels in stochastic frontier models with panel data. *Journal of Productivity Analysis* 14: 91-118.

Koop, G., Osiewalski, J., Steel, M.F.J. (1997). Bayesian efficiency analysis through individual effects: hospital cost frontiers. *Journal of Econometrics* 76: 77-105.

Koop, G., Steel, M.F.J. (2001). Bayesian analysis of stochastic frontier models. In *A Companion to Theoretical Econometrics*, Baltagi, B.H. (ed.). Chichester: Wiley.

Kumbhakar, S.C., Lovell, C.A.K. (2000). *Stochastic Frontier Analysis*. New York: Cambridge University Press.

Kumbhakar, S.C., Ghosh, S., McGuckin J.T. (1991). A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business & Economic Statistics* 9: 279-86.

Li, M., Tobias, J.L. (2006). Bayesian analysis of structural effects in an ordered equation system. *Studies in Nonlinear Dynamics and Econometrics* 10: Article 7.

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

Mutter, R.L., Greene, W.H., Spector, W., Rosko, M.D., Mukamel, D.B. (2012). Investigating the impact of endogeneity on inefficiency estimates in the application of stochastic frontier analysis to nursing homes. Journal of Productivity Analysis, published online 20 March 2012.

Nandram, B., Chen, M.-H. (1996). Reparameterizing the generalized linear model to accelerate Gibbs sampler convergence. *Journal of Statistical Computation and Simulation* 54: 129-144.

Pitt, M., Lee, L.F. (1981). The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics* 9: 43-64.

Puig-Junoy, J. (1998). Measuring health production performance in the OECD. *Applied Economics Letters* 5: 255-259.

Puig-Junoy, J., Ortun, V. (2004). Cost efficiency in primary care contracting: a stochastic frontier cost function approach. *Health Economics* 13: 1149-1165.

Rosko, M.D. (2001). Cost efficiency of US hospitals: a stochastic frontier approach. *Health Economics* 10: 539-551.

Stevenson, R.E. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics* 13: 57-66.

Thornton, J. (2002). Estimating a health production function for the US: some new evidence. *Applied Economics* 34: 59-62.

van den Broeck, J., Koop, G., Osiewalski, J., Steel, M.F.J. (1994). Stochastic frontier models – a Bayesian perspective. *Journal of Econometrics* 61: 273-303.

Table 1. Monte Carlo Results

	True Parameter	Average of Posterior Means	St.devn. of Posterior Means	Average of Posterior St.devns.
β	0.2	0.203	0.026	0.026
	0.55	0.549	0.014	0.014
	-0.05	-0.050	0.006	0.006
	0.4	0.400	0.012	0.013
γ	0.3	0.303	0.007	0.003
	0.6	0.606	0.007	0.003
σ_v	0.5	0.501	0.005	0.005
λ	0.5	0.500	0.013	0.013

Table 2. Rice Data Estimates

	Continuous Output		Categorical Output	
	Posterior mean	Posterior St. Devn.	Posterior mean	Posterior St. Devn.
constant	-1.63	0.26	-1.50	0.37
ln(area)	0.36	0.06	0.38	0.09
ln(labor)	0.43	0.07	0.41	0.10
ln(fert)	0.22	0.04	0.24	0.05
σ_v	0.34	0.01	0.38	0.03
λ	0.07	0.03	0.15	0.04
	True thresholds			
γ_1	0.69		0.64	0.06
γ_2	1.39		1.31	0.05
γ_3	2.30		2.30	0.00

Table 3. Posterior Statistics for Estimated Coefficients

Model 1 (x in frontier, w in u)

	Mean	**St.D**	**95% CI**
β			
ONE	0.782	0.008	(0.767, 0.798)
LT3EX	0.059	0.005	(0.050, 0.068)
MT3	0.133	0.005	(0.124, 0.143)
NOSM	0.069	0.004	(0.061, 0.076)
LRA	0.042	0.005	(0.032, 0.051)
HIGHRA	0.026	0.006	(0.013, 0.038)
LONELY1	-0.063	0.003	(-0.068, -0.057)
LONELY2	-0.073	0.004	(-0.081, -0.065)
INNER	-0.003	0.004	(-0.012, 0.005)
OUTER	-0.015	0.006	(-0.027, -0.004)
REMOTE	-0.009	0.012	(-0.032, 0.014)
STUDENT	0.041	0.007	(0.027, 0.055)
PARTTIME	-0.013	0.004	(-0.021, -0.005)
UNEMP	-0.016	0.008	(-0.031, 0.000)
RETD	-0.122	0.006	(-0.134, -0.110)
NOTINLAB	-0.075	0.005	(-0.085, -0.066)
DEGREE	0.097	0.006	(0.085, 0.108)
DIPLOMA	0.020	0.005	(0.010, 0.030)
YEAR12	0.064	0.006	(0.052, 0.076)
ϕ			
ONE	2.352	0.070	(2.218, 2.492)
GENDER	0.939	0.017	(0.906, 0.974)
AUSABO	0.830	0.054	(0.728, 0.939)
MAINENG	1.079	0.032	(1.018, 1.142)
OTHERC	0.943	0.026	(0.893, 0.993)
MARRIAGE	1.094	0.022	(1.051, 1.137)
AGEG1	1.634	0.059	(1.522, 1.751)
AGEG2	1.538	0.050	(1.443, 1.640)
AGEG3	1.411	0.044	(1.328, 1.500)
AGEG4	1.173	0.037	(1.103, 1.247)
AGEG5	1.047	0.035	(0.981, 1.117)
γ			
γ_1	0.189	0.005	(0.184, 0.193)
γ_2	0.595	0.002	(0.593, 0.597)
γ_3	1.000		
σ_v			
	0.205	0.001	(0.203, 0.207)

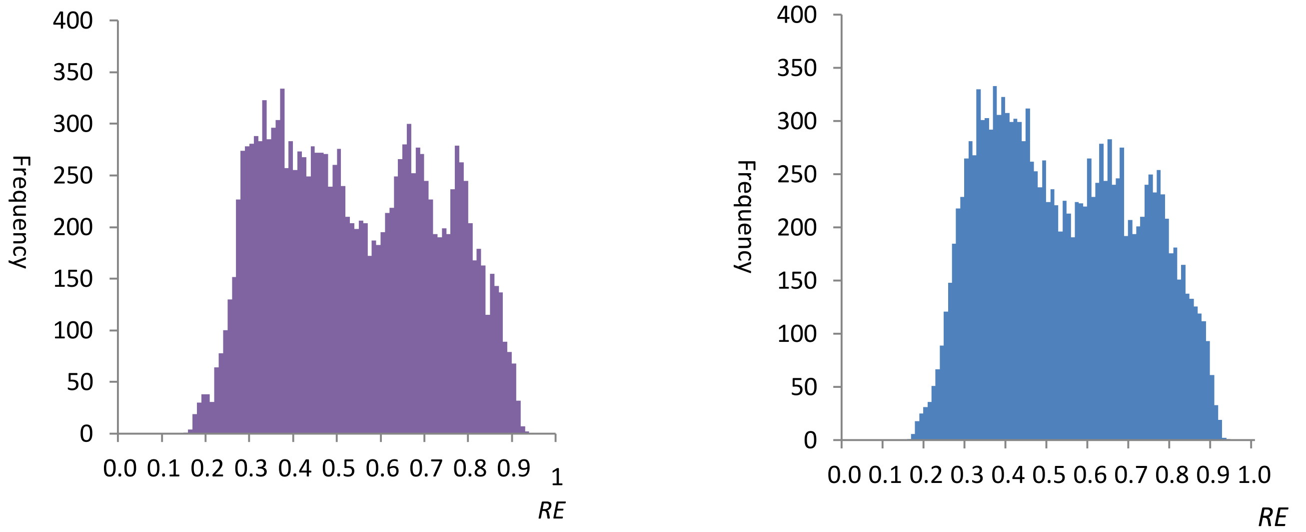
Model 2 (x and w in frontier, no regressors in u)

	Mean	**St.D**	**95% CI**
β			
ONE	0.672	0.011	(0.651, 0.693)
LT3EX	0.061	0.004	(-0.030, -0.012)
MT3	0.138	0.005	(-0.059, 0.006)
NOSM	0.066	0.004	(0.024, 0.051)
LRA	0.044	0.005	(-0.013, 0.013)
HIGHRA	0.019	0.007	(-0.009, 0.007)
LONELY1	-0.064	0.003	(0.162, 0.199)
LONELY2	-0.073	0.004	(0.151, 0.182)
INNER	-0.002	0.004	(0.117, 0.147)
OUTER	-0.013	0.006	(0.064, 0.094)
REMOTE	-0.004	0.012	(0.033, 0.061)
STUDENT	0.030	0.008	(0.052, 0.069)
PARTTIME	-0.013	0.004	(0.128, 0.147)
UNEMP	-0.020	0.008	(0.058, 0.073)
RETD	-0.085	0.007	(0.035, 0.053)
NOTINLAB	-0.072	0.005	(0.007, 0.032)
DEGREE	0.094	0.006	(-0.070, -0.058)
DIPLOMA	0.021	0.005	(-0.082, -0.065)
YEAR12	0.053	0.006	(-0.010, 0.007)
GENDER	-0.021	0.004	(-0.024, -0.002)
AUSABO	-0.027	0.017	(-0.027, 0.019)
MAINENG	0.037	0.007	(0.015, 0.045)
OTHERC	0.000	0.007	(-0.021, -0.004)
MARRIAGE	-0.001	0.004	(-0.035, -0.004)
AGEG1	0.181	0.009	(-0.098, -0.072)
AGEG2	0.166	0.008	(-0.082, -0.062)
AGEG3	0.132	0.008	(0.083, 0.106)
AGEG4	0.080	0.008	(0.011, 0.031)
AGEG5	0.047	0.007	(0.040, 0.065)
γ			
γ_1	0.188	0.001	(0.187, 0.189)
γ_2	0.590	0.003	(0.587, 0.593)
γ_3	1.000		
σ_v			
	0.205	0.001	(0.203, 0.207)
λ			
	0.324	0.004	(0.317, 0.332)

Table 4. Posterior Mean (Standard Deviation) for Marginal Effects of X on Health Status

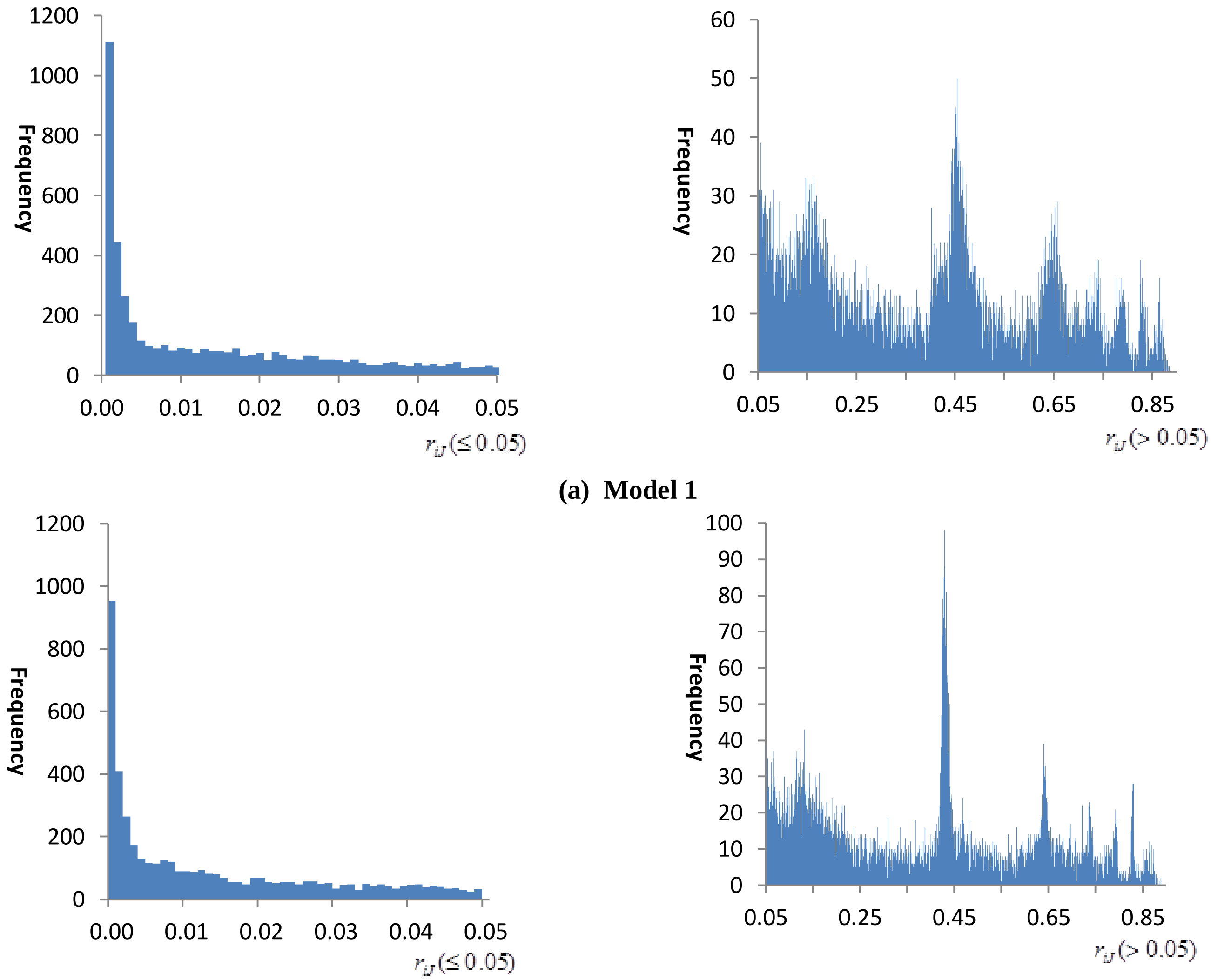
	Pr($y=0$)		**Pr($y=1$)**		**Pr($y=2$)**		**Pr($y=3$)**		**Pr($y=4$)**	
	Model 1	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**	**Model 1**	**Model 2**
LT3EX	-1.59(0.13)*	-1.71(0.26)*	-1.26(0.11)*	-1.32(0.20)*	-4.50(0.35)*	-4.58(0.57)*	4.26(0.38)*	4.59(0.49)*	3.08(0.22)*	3.01(0.29)*
MT3	-3.23(0.15)*	-3.44(0.50)*	-2.56(0.14)*	-2.68(0.39)*	-10.13(0.41)*	-10.32(1.02)*	7.34(0.41)*	7.98(0.87)*	8.58(0.29)*	8.47(0.61)*
NOSM	-1.59(0.09)*	-1.58(0.25)*	-1.27(0.09)*	-1.23(0.19)*	-5.25(0.31)*	-4.97(0.53)*	3.45(0.22)*	3.49(0.44)*	4.67(0.27)*	4.28(0.37)*
LRA	-1.01(0.12)*	-1.11(0.20)*	-0.80(0.10)*	-0.86(0.16)*	-3.20(0.38)*	-3.36(0.48)*	2.38(0.31)*	2.66(0.41)*	2.63(0.29)*	2.68(0.32)*
HIGHRA	-0.64(0.16)*	-0.50(0.19)*	-0.51(0.13)*	-0.39(0.14)*	-2.00(0.49)*	-1.47(0.52)*	1.58(0.39)*	1.27(0.44)*	1.57(0.39)*	1.10(0.38)*
LONELY1	1.48(0.08)*	1.58(0.24)*	1.18(0.07)*	1.22(0.18)*	4.79(0.25)*	4.87(0.50)*	-3.30(0.20)*	-3.58(0.43)*	-4.14(0.20)*	-4.08(0.32)*
LONELY2	1.74(0.11)*	1.81(0.28)*	1.39(0.10)*	1.42(0.22)*	5.57(0.34)*	5.55(0.61)*	-4.01(0.30)*	-4.23(0.52)*	-4.69(0.25)*	-4.55(0.38)*
INNER	0.07(0.10)	0.04(0.10)	0.06(0.08)	0.03(0.08)	0.24(0.33)	0.12(0.33)	-0.16(0.22)	-0.09(0.23)	-0.21(0.29)	-0.10(0.28)
OUTER	0.36(0.14)*	0.32(0.15)*	0.29(0.11)*	0.25(0.12)*	1.18(0.46)*	1.00(0.45)*	-0.83(0.33)*	-0.74(0.34)*	-1.01(0.38)*	-0.83(0.36)*
REMOTE	0.22(0.27)	0.11(0.29)	0.17(0.22)	0.08(0.22)	0.70(0.89)	0.32(0.89)	-0.50(0.62)	-0.25(0.65)	-0.59(0.76)	-0.25(0.75)
STUDENT	-0.81(0.14)*	-0.64(0.19)*	-0.65(0.11)*	-0.50(0.15)*	-2.99(0.52)*	-2.22(0.58)*	0.98(0.13)*	1.00(0.30)*	3.47(0.65)*	2.36(0.64)*
PARTTIME	0.28(0.09)*	0.28(0.11)*	0.23(0.07)*	0.22(0.08)*	1.00(0.31)*	0.94(0.34)*	-0.51(0.16)*	-0.56(0.21)*	-1.00(0.31)*	-0.89(0.31)*
UNEMP	0.35(0.18)*	0.46(0.20)*	0.28(0.14)*	0.36(0.16)*	1.21(0.61)*	1.49(0.62)*	-0.64(0.34)*	-0.93(0.43)*	-1.20(0.58)*	-1.37(0.55)*
RETD	3.13(0.19)*	2.17(0.35)*	2.48(0.17)*	1.69(0.28)*	9.30(0.47)*	6.45(0.78)*	-7.79(0.50)*	-5.25(0.68)*	-7.12(0.32)*	-5.05(0.48)*
NOTINLAB	1.79(0.13)*	1.79(0.29)*	1.42(0.11)*	1.40(0.22)*	5.76(0.39)*	5.48(0.64)*	-4.04(0.32)*	-4.24(0.54)*	-4.93(0.31)*	-4.43(0.41)*
DEGREE	-2.17(0.14)*	-2.18(0.35)*	-1.73(0.13)*	-1.71(0.27)*	-7.31(0.46)*	-7.07(0.78)*	4.27(0.28)*	4.43(0.65)*	6.93(0.46)*	6.52(0.60)*
DIPLOMA	-0.50(0.13)*	-0.55(0.16)*	-0.40(0.11)*	-0.43(0.12)*	-1.53(0.40)*	-1.62(0.42)*	1.25(0.33)*	1.38(0.36)*	1.18(0.31)*	1.22(0.31)*
YEAR12	-1.51(0.15)*	-1.30(0.24)*	-1.20(0.13)*	-1.01(0.19)*	-4.90(0.48)*	-3.98(0.60)*	3.36(0.32)*	3.01(0.46)*	4.25(0.44)*	3.28(0.46)*

NOTE: The marginal effect of a dummy variable is estimated as the difference between the probabilities when the dummy is turned on and turned off, keeping all the other variables at their sample mean values. All numbers in this table are presented as percentages. For example, from Model 1 the posterior mean for the marginal effect of doing exercise for less than 3 times but at least 1 time on the probability of poor health status is -1.59%, the posterior standard deviation is 0.13%. “*” denotes the 95% confidential interval of the corresponding marginal effect does not include 0.



(a) Model 1 **(b) Model 2**

Figure 1.Histogram for Posterior Means of $RE_{i;med}$ for All Within-Sample Individuals.



(a) Model 1

(b) Model 2

Figure 2. Histogram of Posterior Means for r_{iJ} for All Within-Sample Individuals

NOTE: The histogram is drawn in two segments, with a different scale, but the same bin widths, for each segment.

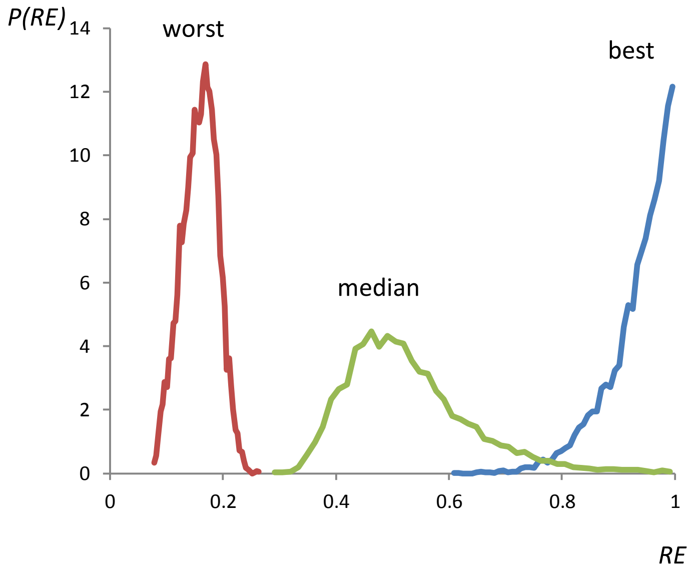


Figure 3. Posterior Densities for Relative Efficiency $RE_{i;med}$ for Three Within-Sample Individuals from Model 1.

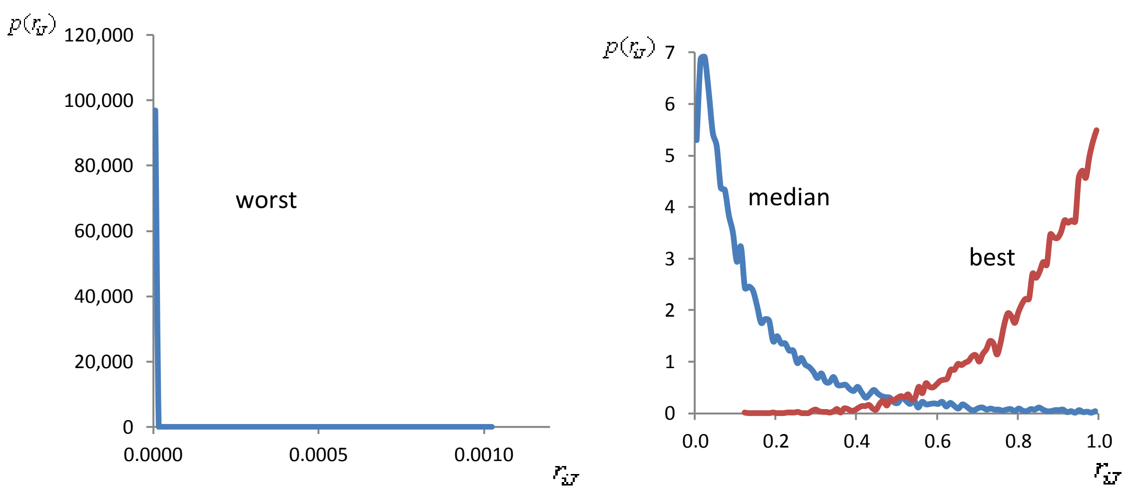


Figure 4. Posterior Densities for Relative Probability r_{iJ} for Three Within-Sample Individuals from Model 1.

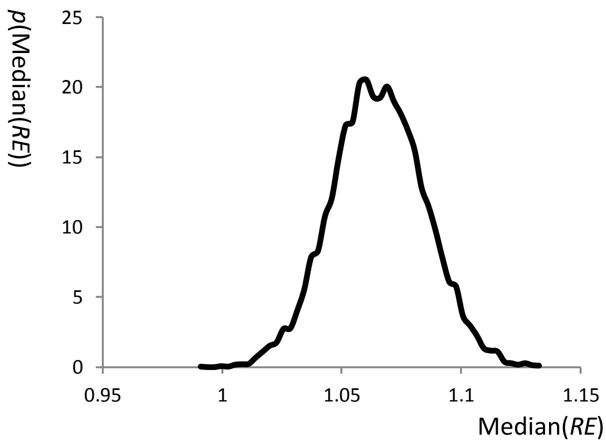


Figure 5. Posterior Density for Median Relative Efficiency of Female to Male

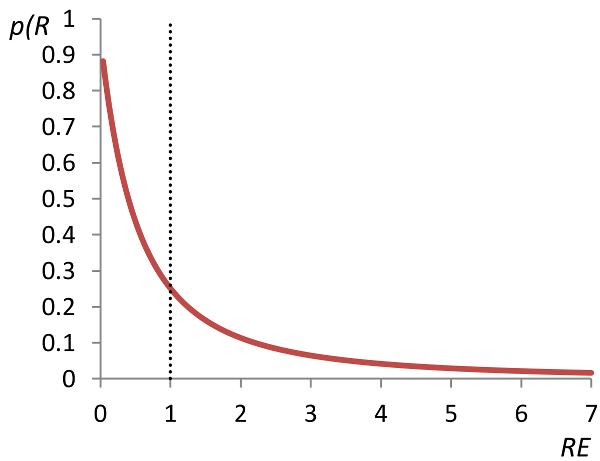


Figure 6. Predictive Densities for Relative Efficiency for Out-of-Sample Female vs. Male

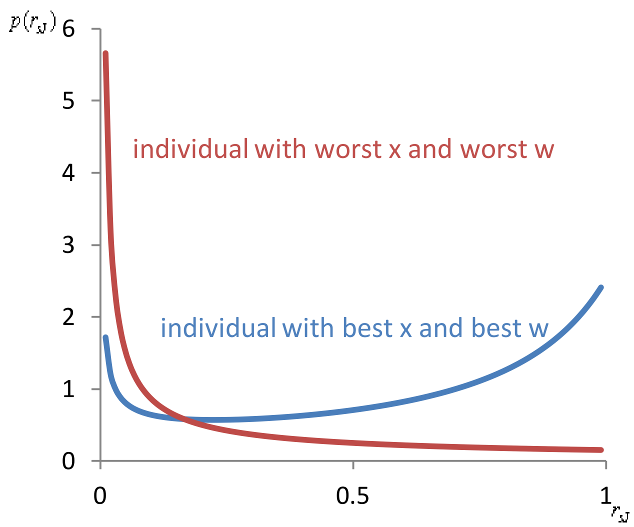


Figure 7. Predictive Densities for Relative Probability for Out-of-Sample Individuals