RISE | RICE INITIATIVE *for the* STUDY *of* ECONOMICS

RISE Working Paper 15-019

"Productivity Measurement, Model Averaging, and World Trends in Growth and Inequality"

by Robin C. Sickles, Jiaqi Hao and Chenjun Shang

RICE

Department of Economics
Baker Hall, MS22
6100 Main Street, Houston, Texas 77005
https://economics.rice.edu

Productivity Measurement, Model Averaging, and World Trends in Growth and Inequality*

Robin C. Sickles, Jiaqi Hao and Chenjun Shang

Abstract Our paper provides new methods to robustify productivity growth measurement by utilizing various economic theories explaining economic growth and productivity and the econometric model generated by that particular theory. We utilize the World Productivity Database from the UNIDO to analyze productivity during the period 1960-2010 for OECD countries. We focus on three competing models from the stochastic frontier literature, Cornwell, Schmidt, and Sickles (1990), Kumbhakar (1990) and Battese and Coelli (1992) to estimate productivity growth and its decomposition into technical change and efficiency change and utilize methods due to Hansen (2010) to construct optimal weights in order to model average the results from these three approaches.

Keywords: productivity, panel data, stochastic frontiers, time varying heterogeneity, model averaging, United Nations Industrial Development Organization

Robin C. Sickles
Rice University, Houston, Texas, e-mail: rsickles@rice.edu

Jiaqi Hao
Credit Scorecard and Portfolio Management, ATB Financial, Edmonton Alberta, e-mail: jhao@atb.com

Chenjun Shang
Rice University, Houston, Texas, e-mail: Chenjun.Shang@rice.edu

* This Pleanary Session presentation was prepared for the North American Productivity Workshop VIII, hosted by Carleton University, Ottawa, Canada, June 4 - June 7, 2014 and is based in part on Sickles, Hao, and Shang (Oxford Handbook on Panel Data, 2015), Shang (2015), and Sickles, Hao, and Shang (2014). The authors would especially like to thank Marcel Viola for his extensive and careful editing of our contribution. The usual caveat applies.

1 Introduction

Proper measurement of nations' productivity growth is essential to understand current and future trends in world income levels, growth in per/capita income, political stability, and international trade flows. In measuring such important economic statistics is it also essential that a method that is robust to misspecification error is used. This talk addresses the robustification of productivity growth measurement by utilizing various economic theories explaining economic growth and productivity and the econometric model generated by that particular theory. We start from a realistic assumption that all models are misspecified in one way or another. Just as the famous quote by Box, "essentially, all models are wrong, but some are useful", carefully designed procedures to approximate the underlying DGP based on all collected information are needed. We address the heterogeneity problem by grouping countries according to their geographical, cultural and development characteristics as well as by the use of various panel data techniques. We utilize the World Productivity Database from the UNIDO to analyze productivity during the period 1960-2010. We consolidate the empirical findings from a number of statistical treatments consistent with the various economic models of economic growth and productivity. We discuss methodologies for averaging these various empirical findings. We also construct consensus estimates of world productivity TFP growth and find that, compared to efficiency catch-up, innovation plays a much more important factor in generating TFP growth.

2 Traditional Explanations for Sources of Economic Growth

The primary sources of economic growth and development are centered on 2 basic explanations: factor-accumulation and productivity-growth.

Rapid economic growth in East Asia in the 1970's and 1980's were thought by Kim and Lau (1994), Young (1992, 1995) and Krugman (1994) to be largely explained by the mobilization of resources. An alternative explanation to the neoclassical hypothesis explains economic growth in terms of intensive and extensive utilization of input factors as well as governmental industrial policies and liberalization policies. The sources of economic growth can be derived by explicitly introducing the role of catch-up due to an increase in productivity efficiency (Hultberg, Nadiri, and Sickles, 1999; 2004).

Introducing the role of efficiency in production means introducing some form of frontier production process, i.e., stochastic frontier production. Total Factor Productivity (TFP) growth is often decomposed into technological (technical innovation) change and technical efficiency change. Modifications of the neoclassical model can be found in the new growth theory. Endogenous growth models were developed to weaken the strong neoclassical assumption that long-run productivity growth could only be explained by an exogenously driven change in technology. Sources of productivity differences in post WWII industrialized countries can be explained by neoclassical growth models that incorporate knowledge spillovers, technological diffusion, and convergence to a best practice production process (Smolny, 2000).

2.0.1 Classical Residual based Partial and Total Factor Productivity Measurement

Productivity historically has been specified as the ratio of some function of outputs (Y_i) to some function of inputs (X_i), which may be further adjusted by accounting for changing output and input mix. For a single such total factor productivity (TFP) output is often written as:

$$TFP = \frac{Y}{\sum a_i X_i}. \tag{1}$$

The a_i weights can be assigned as *arithmetic weighted averages (*Kendrick, 1961) wherein the weights are typically based on expenditure shares, or as *geometric weighted averages* (Solow, 1957). As growth in TFP is usually of primary concern, geometric averages have usually been used and this leads to the Solow measure, which is adopted by most central governments and statistical agencies and is based on the Cobb-Douglas production function with constant returns to scale, $Y = AX_L^{\alpha} X_K^{1-\alpha}$:

$$TFP = \frac{Y}{X_L^{\alpha} X_K^{1-\alpha}}. \tag{2}$$

Assuming cost mnimization, the parameter α is the expenditure share of labor and a measure of TFP growth is the simple time derivative of TFP:

$$T\dot{F}P = \frac{dY}{Y} - \left[\alpha \frac{dX_L}{X_L} + (1-\alpha)\frac{dX_K}{X_K}\right] \tag{3}$$

and thus a total factor productivity index is simply the difference between the log of the output index and the log of the input index. Growth in the index is thus the first difference over a time period in the differences of the log output aggregator and the log input aggregator (Jorgenson and Griliches, 1972). Of course index numbers themselves have a long standing literature that has been surveyed by a number of scholars, based on part on the pioneering work of Fisher (1927) who formulated a number of desirable properties for index numbers. One such survey can be found in Good, Nadiri, and Sickles (1997).

2.0.2 Modifications of the Neoclassical Model: The New Growth Theory

Endogenous growth models were developed to weaken the strong neoclassical assumption that long-run productivity growth could only be explained by an exogenously driven change in technology. An alternative interpretation to the endogenous growth literature is that it was a response to the simplistic view that the benefits of technical change (aka 'manna from heaven', Scherer, 1971) were determined 'outside the system.' However, technological change as result of economic factors was discussed in Griliches' 1957 Ph.D. dissertation and his concurrent article Griliches (1957), wherein he pointed out that hybrid corn seed penetration followed a logistic distribution. The diffusion of innovations and the technological change it engenders has much in common with the penetration of seeds varieties in agricultural production. It is thus no surprise that numerous instances of such patterns were found by many researchers, including a productivity pioneer in his own right, Edwin Mansfield (1961). Mansfield's treatment of technological change and the rate of imitation was in its own right equally prescient. The classic model put forth by Romer (1986), which began the "new growth theory," allowed for non-diminishing returns to capital due to external effects. For example, research and development by a firm could spill over and affect the stock of knowledge available to all firms. In the simple Romer model firms face constant returns to scale to all private inputs. The level of technology A can vary depending on the stock of some privately provided input R (such as knowledge) and the production function is formulated as

$$Y = A(R)f(K,L,R). \tag{4}$$

In the "new" growth theory the production frontier is shifted by factors that are endogenous, such as "learning-by-doing" Arrow (1962) , the "stock of research and development" (Romer, 1986), "human capital (Lucas, 1988), "trade spillovers" (Coe and Helpman, 1995; Coe, Helpman, and Hoffmaister, 1997), and "trade openness" (Diao, Rattsø, and Stokke, 2005). However, i*f the explanation for the spillover that endogenously determines technology change is the loosening of constraints on the utilization of the technology, then this is just a another way of saying that TFP growth is primarily determined by the efficiency with which the existing technology (inclusive of innovations) is utilized* (Sickles, et al. 2015).

We will take a reduced form approach in much of what we discuss below. The literature on structural modeling of productivity models is quite dense and, outside the scope of our study. The broader structural modeling of static and dynamic productivity models (see for example, Olley and Pakes, 1996) speaks to other issues than those we focus on herein. These issues involve, among other things, the role of errors-in-variables, weak instrument bias, index construction, and stability in panel data modeling of production processes. They have been taken up by a number of researchers. The NBER is particularly well-represented. Studies by Griliches and Hausman (1986), Stoker, Berndt, Ellerman, Schennach (2005), Griliches and Mairesse (1990, 1998), and Griliches and Pakes (1984), Diewert (2002, 2004a,b) are but a few in this extensive literature.

3 Decomposition of Economic Growth-Innovation and Efficiency Change Identified by Regression

Regression based approaches to decompose productivity growth into technical change and catch-up (efficiency change) components can be based on the following generic model. Assume that the multiple output / multiple input technology can be estimated parametrically using the output distance function (Caves, Christensen and Diewert, 1982; Coelli and Perelman, 1996). We consider distance or single output production functions that are linear in parameters, such as the linear in logs Cobb-Douglas, translog, generalized-Leontief and quadratic. These constitute the predominant functional forms used in productivity studies. We begin with a relatively simple representation of the output distance function as an $m-$output, $n-$input deterministic distance function $D_o(Y,X)$ given by the Young index, described in Balk (2008):

$$D_o(Y,X) = \frac{\prod_{j=1}^{m} Y_{it}^{\gamma_j}}{\prod_{k=1}^{n} X_{it}^{\delta_k}} \leq 1. \tag{5}$$

The output-distance function $D_o(Y,X)$ is non-decreasing, homogeneous, and convex in Y and non-increasing and quasi-convex in X. The output distance function is linearly homogeneous in outputs. Take logs, add a disturbance term v_{it} to account for nonsystematic error in observations, functional form, etc. and a technical efficiency term $\eta_i(t)$ to reflect the nonnegative difference between the upper bound of unity for the distance function and the observed value of the distance function for country i at time t.Then we can write the distance function as:

$$-y_{1,it} = \sum_{j=2}^{m} \gamma_j y^*_{jit} + \sum_{k=1}^{n} \delta_k x^*_{kit} + \eta_i(t) + u_{it} \tag{6}$$

where $y^*_{jit,j=2,\ldots,m} = ln(Y_{jit}/Y_{1it})$ and $x^*_{kit} = \ln(X_{kit})$.

After redefining a few variables the distance function can be written as

$$y_{it} = x_{it}\beta + \eta_i(t) + v_{it}. \tag{7}$$

The Cobb-Douglas specification of the distance function (Klein, 1953) has been criticized for its assumption of separability of outputs and inputs and for incorrect curvature as the production possibility frontier is convex instead of concave. However, as pointed out by Coelli (2000), the Cobb-Douglas remains a reasonable and parsimonious first-order local approximation to the true function.

The translog output distance function introduces second-order terms that allow for greater flexibility without sacrificing the possibility of proper local curvature and lifts the assumption that outputs and inputs are separable. The translog output distance function also can be framed in this canonical model representation of a linear panel model with country-specific and time-varying heterogeneity. If the translog technology is applied, the distance function takes the form:

$$-y_{1it} = \sum_{j=2}^{m} \gamma_j y^*_{jit} + \frac{1}{2}\sum_{j=2}^{m}\sum_{l=1}^{m} \gamma_{jl} y^*_{jit} y^*_{lit} + \sum_{k=1}^{n} \delta_k x^*_{kit}$$
$$+\frac{1}{2}\sum_{k=1}^{n}\sum_{p=1}^{n} \delta_{kp} x^*_{kit} x^*_{pit} + \sum_{j=2}^{m}\sum_{k=1}^{n} \theta_{jk} y^*_{jit} x^*_{lit} + \eta_{it} + u_{it} \quad (8)$$

Since the model is linear in parameters then after redefining a few variables the translog distance function also can be written as $y_{it} = x_{it}\beta + \eta_i(t) + v_{it}$.

A similar transparent reparametrization of any distance function that is linear in parameters can be used to estimate other linear in parameters distance or production functions such as the generalized Leontief or quadratic. If the technology involves multiple outputs, then the right hand side endogenous variables must be instrumented. Whether or not the effects need to be instrumented depends on their orthogonality with all or a subset of the regressors. This is the generic model for estimating efficiency change using panel data (frontier) methods that we will explore below. If we assume that innovations are available to all firms and that country- or firm-specific idiosyncratic errors are due to relative inefficiencies, then we can decompose sources of TFP growth in a variety of ways. The overall level of innovation change (innovation is assumed to be equally appropriable by all countries) can be measured directly by such factors as a distributed lag of R&D expenditures, or patent activity, or some such direct measure of innovation. The overall level of innovation change also can be proxied by the time index approach of Baltagi and Griffin (1988), linear time trends, or some other type of time variable. Innovation measured in any of these ways would be identified in most empirical settings. Direct measures are identified of course by the assumption that the matrix of regressors has full column rank, and the indirect measures by functional form assumptions. For example, the index number approach used in Baltagi and Griffin is identified by its nonlinear construction. Innovation also is often proxied by exogenous or stochastic linear time trends (Bai, Kao, and Ng, 2009).

3.1 What is the Correct Model?

One can explore a number of regression-based methods introduced into the literature to measure productivity growth and its decomposition into innovation and catch-up, or efficiency change (Sickles, et al, 2015). We will examine estimates from these various specifications using the generic linear panel data model with time-varying and cross-sectionally varying effects that is given above. The generic panel data model $y_{it} = x_{it}\beta + \eta_i(t) + v_{it}$ nests all multi-output/multi-input panel models that are linear in parameters and can be used to estimate productivity growth and decompose it into innovation and catch-up. We assume that we have a balanced panel although this is done more for notational convenience than for substantive reasons. The generic model of course nests all models that we introduce below for which there is no temporal change in technical efficiency, that is, the usual fixed or random effects stochastic panel frontier models introduced by Pitt and Lee (1981) and Schmidt and Sickles (1984).

We first discuss the most common estimators in use and those that have been introduced rather recently and how they can be implemented in empirical applications. We then show how these methods can be used in a model averaging exercise to evaluate world productivity trends.

3.1.1 The Cornwell Schmidt and Sickles (1990) Panel Stochastic Frontier Model

Extensions of the panel data model by Cornwell, Schmidt, and Sickles (CSS) (1990) generalized the model in which heterogeneity was only allowed in the interecept by allowing for heterogeneity in slopes as well and this permitted researchers to estimate productivity change that was specific to the cross-sectional unit (firms, industries, countries) that could change over time.

A particular parameterization of the CSS model that accomplishes this objective is based on the assumption that in the generic model above $(y_{it} = x_{it}\beta + \eta_i(t) + v_{it})$ the heterogeneity term $\eta_i(t)$ is given by

$$\eta_i(t) = W_{it}\delta_i + v_{it}.$$

The coefficients in the vector δ_i depend on the different cross-sectional units i and represent heterogeneity in slopes. In their application to the US commercial airline industry CSS specified $W_{it} = (1, t, t^2)$ although this was just a parsimonious parameterization useful for their application. The CSS estimator does not in general limit the effects to be quadratic in time but does restrict the effects to be linear in the parameters of the variables whose slopes vary by cross-sectional units. Three different estimators were derived based on differing assumptions made in regard to the correlation of the efficiency effects and the regressors, specifically relating to the correlation between the error term u and regressors X and W. These estimators are the *within* (FE) estimator, which allows for correlation between all of the regressors and the effects, the *gls* estimator, which is consistent when no correlation exists between the technical efficiency term and the regressors (Pitt and Lee, 1981; Kumbhakar, 1990), and the *efficient instrumental variables* estimator, which can be obtained by assuming orthogonality of some of the regressors with the technical efficiency effects. The explicit formulas for deriving each estimator and methods for estimating the δ_i parameters are provided in the Cornwell et al. (1990). Relative efficiencies, normalized by the consistent estimate of the order statistics identifying the most efficient cross-sectional unit, are calculated as:

$$\widehat{\eta}(t) = \max_j [\widehat{\eta}_j(t)]$$

and

$$RE_i(t) = \widehat{\eta}(t) - \widehat{\eta}_i(t).$$

Here $RE_i(t)$ is the relative efficiency of the ith cross-sectional unit at time t. For this class of models the regressors X contain a time trend interpreted as the overall level of innovation. When it is combined with the efficiency term $\widehat{\eta}_j(t)$ we have a decomposition of TFP into innovation and catch-up. When the time trend and the efficiency term both enter the model linearly then the decomposition is not identified using the within estimator. The composition is identified for the gls and for selected variants of the efficient IV model, such as those used in the Cornwell et al. (1990) airline study. In our empirical illustration using the UNIDO data to estimate world productivity growth that follows, we utilize the gls version of the CSS estimator (labelled CSSG) and the efficiency IV estimator (labelled EIV).

3.1.2 The Kumbhakar (1990) Panel Stochastic Frontier Model

Consider the linear in log production function:

$$y_{it} = x_{it}\beta + \eta_i(t) + v_{it}$$

$\eta_i(t) = \gamma(t)\tau_i$,where v_{it} is assumed i.i.d. with distribution $N(0,\sigma_v^2)$. $\eta_i(t)$ is the inefficiency term with a time-varying factor $\gamma(t)$ and time-invariant characteristics τ_i. τ_i is assumed to be distributed as $i.i.d.$ half-normal and $\gamma(t)$ is specified as the logistic function

$$\gamma(t) = (1+\exp(bt+ct^2))^{-1}.$$

Here $\gamma(t)$ is bounded between $(0,1)$ and accommodates increasing, decreasing or time-invariant inefficiency behavior as the parameters b and c vary. Although the Kumbhakar model also estimates allocative efficiency from side conditions implied by cost-minimization (Schmidt and Lovell, 1979) we will only examine the portion of his model that directly pertains to the technical inefficiency/innovation decomposition of productivity change. Parametric maximum likelihood is used for estimation of the main parameters of the model. The inefficiency term is estimated by analogue methods based on the population first moment of $\tau_i|\theta_i$.The best predictor of technical efficiency is then given by $E(\exp\{\gamma(t)\tau_i|\theta_i\})$ and efficiency for each unit is given by $\widehat{\eta}_i(t) = \gamma(t)\widehat{\tau}_i$.

3.1.3 The Battese and Coelli Model (1992, 1995)

The production function is given by the generic model

$$y_{it} = x_{it}\beta + \eta_i(t) + v_{it}. \tag{9}$$

The effects are specified as

$$\eta_i(t) = -\{\exp[-\eta(t-T)]\}u_i,$$

where v_{it} is assumed to be an $i.i.d.$ $N(0,\sigma_v^2)$ random variable, u_{it} is assumed to follow an $i.i.d.$ non-negative truncated $N(\mu,\sigma^2)$ distribution, η is a scalar and the temporal movement of the technical efficiency effects depends on the sign of η. Time invariant technical efficiency corresponds to $\eta = 0$. A richer temporal path for firm efficiency effects can be obtained by specifying $\eta(t-T)$ as

$$\eta_t(t-T) = 1 + a(t-T) + b(t-T)^2.$$

This permits the temporal pattern of technical efficiency effects to be convex or concave rather than simply increasing or decreasing at a constant rate. The model is estimated by parametric mle and the minimum-mean-squared-error predictor of the efficiency for unit i at time t is

$$E[\exp(-u_{it})|\varepsilon_i] = \{\frac{1-\Phi[\eta_{it}\sigma_i^* - (\mu_i^*)/\sigma_i^*]}{1-\Phi(-\mu_i^*/\sigma_i^*)}\}\exp[-\eta_{it}\mu_i^* + \frac{1}{2}\eta_{it}^2\sigma_i^{*2}].$$

Estimates of technical change due to innovation are based on the coefficient of a time trend in the regression. The effect of innovation as distinct from catch-up is identified by the nonlinear time effects in the linear technical efficiency term and thus the decomposition of TFP growth into a technological change and efficiency change component is quite natural with this estimator. Cuesta (2000) generalized Battese and Coelli (1992) by allowing each country (firm, etc.) to have its own time path of technical inefficiency. Extensions of the Battese and Coelli model that allow for technical inefficiency to be determined by a set of environmental factors that differ from those that determine the frontier itself are given in Battese and Coelli (1995). These were also addressed by Reifschneider and Stevenson (1991) and by Good, Roeller, and Sickles (1995). Environmental factors that were allowed to partially determine the level of inefficiency and productivity were introduced in Cornwell et al. (1990) and in Good, Nadiri, Roeller, and Sickles (1993).

3.1.4 Alternatives to the Classical Parametric Stochastic Panel Frontier Approaches

Many other variations in the basic panel model treatment of inefficiency have been considered in the literature. We do not pursue those in this here but direct the reader to the work of Park, Sickles and Simar (PSS; 1998, 2003, 2006), who considered linear stochastic frontier panel models in which the distribution of country specific technical efficiency effects is estimated nonparametrically. The latent class models of Orea and Kumbhakar (2004), Tsionas and Kumbhakar (2004), and Greene (2005b) relate to work on production heterogeneity by Mundlak (1961, 1978) and Griliches (1979), among others. Kneip, Sickles, and Song (2012) assume a linear semiparametric panel frontier that allows for an arbitrary pattern of technical change $\eta_i(t)$ based on a general factor model set-up. Their specification of the effects is more flexible than parametric methods and the multiplicative effects models of Lee and Schmidt (1993), Ahn, Lee, and Schmidt (2007), Bai (2009), and Bai and Ng (2011). Ahn, Lee and Schmidt (2013) generalize Ahn, Lee, and Schmidt (2007) and consider a panel data model with multiple individual effects that also change over time and focus on large N and finite T asymptotics. Addtional estimators that have been proposed for panel stochastic frontiers and that are also quite appropriate for general panel data problems are the *Bayesian Stochastic Frontier Model* (Liu, Sickles, and Tsionas, 2013), which builds on earlier work by Van den Broeck, Koop, Osiewalski, and Steel (1994) and Tsionas (2006), the *Bounded Inefficiency Model* of Almanidis, Qian, and Sickles (2013) and related models of Lee (1996), Lee and Lee (2012), and Orea and Steinbuks (2012) as well as the *"True" Fixed Effects Model* of Greene (2005a,b). Kumbhakar, Parmeter, and Tsionas (2013) considered a semiparametric smooth coefficient model to estimate the TFP growth of certain production technologies that addresses the *Skewness Problem* in classical SFA modeling considered by Feng, Horrace and Wu (2013), Almanidis and Sickles (2012) and Almanidis et al. (2013). The *Spatial Stochastic Frontier* shows great promise and has been pursued in recent work by Glass, Kenjegalieva, and Sickles (2013 a,b) based on the original contribution by Druska and Horrace (2004). Work on productivity measurement in the presence of spatial heterogeneity has also recently been pursued Mastromarco and Shin (2013), Entur and Musolesi (2013), and Demetrescu and Homm (2013). These are alternatives to less structured approaches to address cross-sectional dependence in panel data models using methods such as those developed by Pesaran (2007). *Factor Models* continue to be pursued in the context of productivity modeling in panel data contexts and the space for such approaches is getting quite dense as pointed out by Kneip and Sickles (2012).

4 Can We Combine Model Estimates Instead of Choosing the Best?

Discovering the true model might not be possible. Statistical inference based on the "post-model-selection estimators" (Leeb and Potscher, 2005) might lead to invalid analysis. Different selecting criteria might give contradicted ranking orders and focusing on one model and dismissing the results of alternative specifications may compromise the information content of the information set. As discussed in Burnham and Anderson (2002), if observed data are conceptualized as random variables, the sample variability introduces uncertain inference from the particular data set. Model selection is a special case of weighting models in which one model is given the entire weight. Combining model estimates and forecasts can be motivated on the basis of economic theory based on models of majority voting and the Tullock contest function. It can also be motivated on the basis of statistical theory via model averaging and forecast conbination theory.

Most model averaging methods take on a Bayesian perspective, although many recent studies have a frequentist interpretation. In the frequentist literature, the weights are usually based on AIC or BIC criterion. What we will employ here is the method proposed in Hansen (2007), in which the weights are chosen by minimizing a Mallows criterion. It was shown that the resulting estimator can asymptotically achieve the lowest squared error among a finite number of model averaging estimators. For this application our unrestricted model is

$$y_{it} = \sum_{j=1}^{p} \beta_j x_{itj} + \sum_{r=0}^{\infty} \delta_{ir} t^r + \varepsilon_{it}, \qquad i = 1, \ldots, n; t = 1, \ldots, T.$$

Here ε_{it} is the standard noise component assumed to be *iid* $N(0, \sigma)$. Let $z_{it} = (x_{it1}, \ldots, x_{itp}, 1, t, t^2, \ldots)$ be the vector of all regressors, and $\gamma = (\beta_1, \ldots, \beta_p, \delta_0, \delta_1, \ldots)$ be the associated parameter vector. Then the model can be compactly expressed as

$$y_{it} = \sum_{j=1}^{\infty} z_{itj} \gamma_j + \varepsilon_{it}, \tag{10}$$

where z_{itj} is a vector with countably infinite entries and can include regressors that are terms of a series expansion that is linear in parameters.

Denote $\mu_{it} = \sum_{j=1}^{\infty} z_{itj} \gamma_j$, it is assumed $E[\mu_{it}^2] < \infty$ and μ_{it} converges to the mean square. Consider a set of M nested models, and assume the mth model uses the first k_m variables, with $p < k_1 < k_2 < \cdots < k_M$. Put in matrix form, the mth model is

$$Y = Z_m \Gamma_m + \varepsilon. \tag{11}$$

Let $\hat{\Gamma}_m$ be the estimate of the coefficients in the mth model, and let $w = (w_1, \ldots, w_M)$ be the associated weight vector for each model, where $w_m \in [0, 1]$ and $\sum_{m=1}^{M} w_m = 1$. Then the model average estimator of the coefficients for the unrestricted model is

$$\hat{\Gamma} = \sum_{m=1}^{M} w_m \begin{pmatrix} \hat{\Gamma}_m \\ 0 \end{pmatrix}. \tag{12}$$

We further define $k(w) \equiv \sum_{m=1}^{M} w_m k_m$, then the Mallows criterion can be stated as

$$C(w) = (Y - Z_M \hat{\Gamma})'(Y - Z_M \hat{\Gamma}) + 2\sigma^2 k(w) \tag{13}$$

and an optimal weight is obtained by numerically minimizing $C(w)$.

The unrestricted model we consider is from Kneip, Sickles and Song (2012) and is specified as

$$y_{it} = \beta_0(t) + \sum_{j=1}^{p} \beta_j x_{itj} + u_{it} + \varepsilon_{it}, \quad i = 1, \ldots, n; t = 1, \ldots, T. \tag{14}$$

The u_{it}'s are assumed to be smooth time-varying individual effects where $\sum_i^n u_{it} = 0$, $t = 1, \ldots, T$. The individual effects are assumed to be affected by a set of underlying factors and are model by linear combinations of some basis functions.

$$u_{it} = \sum_{r=1}^{L} \delta_{ir} v_r(t) \quad i = 1, \dots, n \tag{15}$$

where $\beta_0(t)$ is some average function and can be eliminated by transforming the model to the centered form:

$$y_{it} - \bar{y}_t = \sum_{j=1}^{p} \beta_j (x_{itj} - \bar{x}_{tj}) + \sum_{r=1}^{L} \delta_{ir} v_r(t) + \varepsilon_{it} - \bar{\varepsilon}_i, \qquad i = 1, \dots, n; t = 1, \dots, T, \tag{16}$$

where $\bar{y}_t = \frac{1}{n}\sum_i y_{it}$, $\bar{x}_{tj} = \frac{1}{n}\sum_i x_{itj}$ and $\bar{\varepsilon}_i = \frac{1}{n}\sum_i \varepsilon_{it}$. Denote $\tilde{y}_{it} = y_{it} - \bar{y}_t$ and $\tilde{x}_{itj} = x_{itj} - \bar{x}_{tj}$.

The functional form used for estimation can be written as

$$\tilde{y}_{it} = \sum_{j=1}^{p} \beta_j \tilde{x}_{itj} + \sum_{r=1}^{L} \delta_{ir} v_r(t) + \tilde{\varepsilon}_{it}, \qquad i = 1, \dots, n; t = 1, \dots, T \tag{17}$$

This model nests several specifications in stochastic frontier analysis. When $v_r(t) = t^{r-1}$ and $L = 3$ we have the Cornwell, Schmidt and Sickles (1990) model discussed above. To show how Kumbhakar (1990) is nested in the general model consider a translog production function that is linear in parameters and can be expressed as

$$y_{it} = X_{it}'\beta + u_{it} + \varepsilon_{it}, \tag{18}$$

where the $u_{it}{\prime}s$ represent the individual effects and given by $u_{it} = v(t)\theta_i = (1 + exp(bt + ct^2))^{-1}\theta_i$.Taking a Taylor expansion of $v(t)$ at $t = 0$, the individual effects can be expressed as

$$\begin{aligned} u_{it} &= \sum_{r=0}^{\infty} \frac{v^{(r)}(0)}{r!} t^r \theta_i \\ &= \theta_i \left(\frac{1}{2} - \frac{1}{4}bt + \frac{1}{4}ct^2 + \cdots\right). \end{aligned} \tag{19}$$

With a finite time period under study, the exponential time-varying path can be closely approximated by a polynomial function of finite degree L_1. Thus the model can be written as

$$y_{it} = X_{it}'\beta + \sum_{r=0}^{L_1} \delta_{ir} t^r + \varepsilon_{it} \tag{20}$$

with $\delta_{ir} = \theta_i \frac{v^{(r)}(0)}{r!}$.

The Battese and Coelli (1992) model can also be nested in the KSS general model using a Taylor expansion. The basic setting is the same, while the individual effects are assumed to follow a different time-varying path.

$$u_{it} = -\eta_{it} u_i = -\{exp[-\eta(t - T)]\} u_i.$$

Taking a Taylor expansion of this function, we have

$$u_{it} = -\sum_{r=0}^{\infty} \frac{\eta^{(r)}(0)}{r!} t^r u_i$$
$$= -e^{\eta T} u_i + \eta e^{\eta T} u_i t - \frac{1}{2}\eta^2 e^{\eta T} u_i + \cdots \quad (21)$$

This exponential function can be sufficiently well approximated by a polynomial function of finite degree L_2 in empirical studies. Setting $\delta_{ir} = -u_i \frac{\eta^{(r)}(0)}{r!}$, the model can be written as

$$y_{it} = X_{it}'\beta + \sum_{r=0}^{L_2} \delta_{ir} t^r + \varepsilon_{it} \quad (22)$$

The KSS model also nests the traditional **random** and **fixed effects** estimators as these are special cases of the CSS estimator.

In the next section we utilize our model averaging methods on these various nested special cases of the general KSS specification and analyze average productivity growth rates and their decomposition into efficiency change and innovation change across various countries in the world economy. The section illustrates the feasibility of the approach and the potential gains that researchers can derive from bringing different models and different assumptions on which they are based to bear in analyzing an important determinant of economic growth and long term economic welfare of a country and of the world economy.

5 Taking Model Averaging to the Data-Some Preliminary Results in a Study of World Productivity (1960-2010)

5.1 UNIDO Data Description

The World Productivity Database (WPD) provides information on measures of the level and growth of TFP based on 12 different empirical methods across 112 countries over the period 1960 -2010. The principal data source is the Penn World Tables from which (chain weighted) GDP and investment are obtained, both in purchasing power parity (1996) US dollars. From the Groningen Growth and Development Centre and Asian Development Bank (ADB, various issues), data on employment and hours worked were also obtained. Unemployment rates and key indicators of the labor market were collected from the International Labor Organization (ILO) Yearbook, and ADB (various issues). Various capital input measures were also constructed. Capital (K) is arguably the most difficult production factor to measure. The WPD presents 4 different approaches based on: (1) different computations for the initial capital stock (2) the depreciation rate (3) schedule for depreciation, and (4) the lifetime of the asset. The different capital measures are labeled K06, K13, Ks and Keff. Common to the first three capital measures is that capital is assumed to depreciate at a constant rate over time. The first two capital stocks differ only in terms of their assumed depreciation rates (6% and 13.3%, respectively, which correspond to about 12 and 6 year asset lives). The different depreciation rates emphasize the importance of the either the initial capital or the effect of recent investments. K06 and K13 are based on assumption that ten years of investment serve as an adequate proxy for the initial capital stock K0. Another common way of computing the initial capital stock is to assume that the country is at its steady state capital-output ratio. This leads to a level of steady-state capital service flows (Ks) from a capital stock whose assumed depreciation rate is 6% per year. A different way of measuring capital focuses on the profile of capital productivity and utilizes a time-varying depreciation rate. As the asset ages, its capital services decline at an increasing rate. This leads to the measure labelled Keff. There also are two kinds of labor utilization rates for which labor force can be adjusted in our analysis. One is based on variations in the numbers employed and one is based on variations in hours worked. The first alternative to labor force (LF) is employment (EMP), which is obtained as a direct measure of employment. The second is derived by applying unemployment rates to LF data which leads to derived employment (DEMP).

We apply our model averaging methodology to the OECD (24 countries) based on UNIDO data from 1960 to 2010. The countries in the **OECD** are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Greece, Iceland, Ireland, Italy, Japan, Republic of Korea, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, UK, and USA.

5.1.1 Summary of Preliminary TFP Findings for the OECD

We choose K06, K13 and Keff as the capital inputs. We use EMP as the labor input. The observation periods are from 1960 to 2010. We use the CSS, K, BC, RE, FE specifications using EMP and three different capital measures. Thus, for this exercise we have 15 different sets of estimates. We aggregate the results by country, by time, to construct aggregate summary measures of technical innovation and technical efficiency growth over the 50 years in our sample of OECD countries. Aggregation is based on utilize geometric means using exchange and ppp weighted gdp shares by each country as the weights. Results suggest that the impact of catch-up relative to tehnical innovation is marginal. Preliminary results based on from our model averaging exercise yields

OECD
1960-2010

TFP growth = 1.04% (innovation) + .09% (catch-up)
= 1.13%.

6 Conclusion

We have discussed different theories on economic growth and productivity measurement and the econometric specifications they imply. We develop a variety of methodologies to combine the results from different models. Our methodologies are illustrated with date from the World Productivity Database gathered by UNIDO. TFP growth is decomposed to two components: technical efficiency change and technological change. We aggregate growth rates of different efficiency measures using model averaging criteria. We find out that in the time period between 1960 and 2010, OECD countries averaged about 1% TFP growth. Innovation that expanded the production possibility frontier plays a much more significant role than catch-up in improving TFP.

References

1. Arrow K J (1962) The economic implications of learning by doing. The Review of Economic Studies 29(3):155–173
2. Bai J (2009) Panel data models with interactive fixed effects. Econometrica 77(4):1229–1279
3. Balk B M (2008) Price and quantity index numbers models for measuring aggregate change and difference. Cambridge University Press, Cambridge
4. Coelli T (2000) Econometric estimation Of the distance function representation of a production technology. CEPA, University of New England
5. Cuesta R A (2000) A production model with firm-specific temporal variation in technical inefficiency: with application to Spanish dairy farms. Journal of Productivity Analysis 13(2):139–158
6. Diewert E (2004a) Preface, (with Hill P, Armknecht P); Chapter 15, basic index number theory; Chapter 16, the axiomatic and stochastic approaches to index number theory; Chapter 17, the economic approach to index number theory: the single household case; Chapter 18, the economic approach to index number theory: the many household case; Chapter 19, price indices using an artificial data set; Chapter 20, elementary indices; Chapter 22, the treatment of seasonal products. In: Consumer price index manual: theory and practice. International Labour Organization, Geneva
7. Diewert E (2004b) Preface, (with Hill P, Armknecht P); Chapter 15, basic index number theory; Chapter 16, the axiomatic and stochastic approaches to index number theory; Chapter 17, economic approach; Chapter 18, aggregation issues; Chapter 19, price indices using an artificial data set; Chapter 20, elementary indices; Chapter 21, quality change and hedonics; Chapter 22, treatment of seasonal products. In: Producer price index manual: theory and practice. International Monetary Fund, Washington DC
8. Fisher I (1927) The making of index numbers: a study of their varieties, tests, and reliability. Houghton Mifflin company, Boston and New York
9. Greene W (2005a) fixed and random effects in stochastic frontier models. Journal of Productivity Analysis 23(1):7–32
10. Greene W (2005b) reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics 126(2):269–303
11. Griliches Z (1957) Hybrid corn: an exploration in the economics of technological change. Econometrica: Journal of the Econometric Society 25(4):501-522
12. Griliches Z (1979) Issues in assessing the contribution of research and development to productivity growth. The Bell Journal of Economics 92–116
13. Hansen B E (2007) Least squares model averaging. Econometrica 75(4):1175–1189

14. Kendrick J W (1961) Front matter, productivity trends in the United States. Productivity Trends in the United States 52–0. NBER
15. Klein L (1953) A textbook of econometrics. Evanston [etc.], Peterson
16. Krugman P (1994) The myth of Asia's miracle. Foreign Affairs 62–78
17. Kumbhakar S C (1990) Production frontiers, panel data, and time-varying technical inefficiency. Journal of Econometrics 46(1):201–211
18. Lee Y (1996) Tail truncated stochastic frontier models. Journal of Economic Theory and Econometrics 2:137–152
19. Lucas R E Jr (1988) On the mechanics of economic development. Journal of Monetary Economics 22(1):3–42
20. Mansfield E (1961) Technical change and the rate of imitation. Econometrica 29(4):741-766
21. Mundlak Y (1961) Empirical production function free of management bias. Journal of Farm Economics 43(1):44–56
22. Mundlak Y (1978) On the pooling of time series and cross section data. Econometrica: Journal of the Econometric Society 69–85
23. Pesaran M H (2007) A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics 22:265-312
24. Romer P M (1986) Increasing returns and long-run growth. The Journal of Political Economy 1002–1037
25. Scherer F M (1971) Industrial market structure and economic performance. Rand McNally, Chicago
26. Shang C (2015) Reduced form analysis of world productivity growth: a model averaging approach. Mimeo.
27. Smolny W (2000) Sources of productivity growth: an empirical analysis with German sectoral data. Applied Economics 32(3):305–314
28. Solow R M (1957) Technical change and the aggregate production function. The Review of Economics and Statistics 39(3):312–320
29. Tsionas E G (2006) Inference in dynamic stochastic frontier models. Journal of Applied Econometrics 21(5):669–676
30. Young A (1992) A tale of two cities: factor accumulation and technical change in Hong Kong and Singapore. In: NBER macroeconomics annual 1992, Vol7. MIT press
31. Young A (1995) The tyranny of numbers: confronting the statistical realities of the east Asian growth experience. The Quarterly Journal of Economics 110(3):641–680
32. Almanidis P, Sickles R C (2012) The skewness problem in stochastic frontier models: fact or fiction? In: Van Keilegom I, Wilson P (ed) Exploring research frontiers in contemporary statistics and econometrics: a festschrift in honor of Leopold Simar. Springer Publishing, New York
33. Bai J, Ng S (2011) Principal components estimation and identification of the factors. Department of Economics, Columbia University
34. Baltagi B H, Griffin J M (1988) A general index of technical change. The Journal of Political Economy 20–41
35. Battese G E, Coelli T J (1992) Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. Journal of Productivity Analysis 3(1-2):153–169
36. Battese G E, Coelli T J (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical Economics 20:325-332
37. Burnham K P, Anderson D R (2002) Model selection and multi-model inference: a practical information-theoretic approach. Springer Verlag
38. Coe D T, Helpman E (1995) International R&D spillovers. European Economic Review 39(5):859–887
39. Coelli T, Perelman S (1996) Efficiency measurement, multiple output technologies and distance functions: with application to European railways. European Journal of Operational Research 117:326-339

40. Diewert E, Deaton A (2002) Conceptual foundations of price and cost of living indexes. In: Schultze C L, Mackie C (ed) At what price? Conceptualizing and measuring cost of living and price indexes. National Academy Press, Washington, DC
41. Demetrescu M, Homm U (2013) A directed test of no cross-sectional error correlation in large-N panel data models. Mimeo
42. Druska V, Horrace W C (2004) Generalized moments estimation for spatial panel data: Indonesian rice farming. American Journal of Agricultural Economics 86: 185-198
43. Entur C, Musolesi A (2013) Weak and strong cross sectional dependence: a panel data analysis of international technology diffusion. Mimeo
44. Griliches Z, Hausman J A (1986) Errors in variables in panel data. Journal of Econometrics 31:93-118
45. Griliches Z, Mairesse J (1990) Heterogeneity in panel data: are there stable production functions? Essays in Honor of Edmond Malinvaud. Vol3:192-231. MIT Press, Cambridge
46. Griliches Z, Mairesse J (1998) Production functions: the search for identification. In: Ström S (ed) Econometrics and Economic Theory in the 20th Century: The Ragnar Frish Centennial Symposium. Cambridge University Press
47. Griliches Z, Pakes A (1984) Distributed lags in short panels with an application to the specification of depreciation patterns and capital stock constructs. The Review of Economic Studies 51(2): 1175-1189
48. Jorgenson D, Griliches Z (1972) Issues in growth accounting: a reply to Edward F. Denison. Survey of Current Business 52.5, Part II
49. Kim J I, Lau L J (1994) The sources of economic growth of the east Asian newly industrialized countries. Journal of the Japanese and International Economies 8(3):235–271
50. Kneip A, Sickles R C (2012) Panel data, factor models, and the solow residual. In: Van Keilegom I, Wilson P (ed) Exploring research frontiers in contemporary statistics and econometrics: a festschrift in honor of Leopold Simar. Springer Publishing, New York
51. Lee Y, Lee S (2012) Stochastic frontier models with threshold efficiency. Mimeo
52. Lee Y H, Schmidt P (1993) A production frontier model with flexible temporal variation in technical efficiency. The Measurement of Productive Efficiency: Techniques and Applications 237–255
53. Leeb H, Pótscher B M (2005) Model selection and inference: facts and fiction. Econometric Theory 21(1):21–59
54. Mastromarco C, Shin Y (2013) Modelling technical efficiency in cross sectionally dependent panels. Mimeo
55. Olley G S, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. Econometrica 64(6):1263-1297
56. Orea L, Kumbhakar S C (2004) Efficiency measurement using a latent class stochastic frontier model. Empirical Economics 29(1):169–183
57. Orea L, Steinbuks J (2012) Estimating market power in homogenous product markets using a composed error model: application to the California electricity market. University of Cambridge, Faculty of Economics
58. Pitt M M, Lee L F (1981) The Measurement and sources of technical inefficiency in the Indonesian weaving industry. Journal of Development Economics 9(1):43–64
59. Reifschneider D, Stevenson R (1991) Systematic departures from the frontier: a framework for the analysis of firm inefficiency. International Economic Review 32:715-723
60. Schmidt P, Lovell C K (1979) Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. Journal of Econometrics 9(3):343–366
61. Schmidt P, Sickles R C (1984) Production frontiers and panel data. Journal of Business & Economic Statistics 2(4):367-374
62. Tsionas E G, Kumbhakar S C (2004) Markov switching stochastic frontier model. Econometrics Journal, Royal Economic Society 7(2):398-425
63. Ahn S C, Lee Y H, Schmidt P (2007) Stochastic frontier models with multiple time-varying individual Effects. Journal of Productivity Analysis 27(1):1-12
64. Ahn S C, Lee Y H, Schmidt P (2013) Panel data models with multiple time-varying individual effects. Journal of Econometrics 174(1):1-14

65. Almanidis P, Qian J, Sickles R C (2013) Bounded stochastic frontiers. Mimeo
66. Bai J, Kao C, Ng S (2009) Panel cointegration with global stochastic trends. Journal of Econometrics (149):82–99
67. Caves D W, Christensen L R, Diewert W E (1982) The economic theory of index numbers and the measurement of input, output, and productivity. Econometrica: Journal of the Econometric Society 1393–1414
68. Coe D T, Helpman E, Hoffmaister A (1997) North-south R&D spillovers. The Economic Journal 107(440):134–149
69. Cornwell C, Schmidt P, Sickles R C (1990) Production frontiers with cross-sectional and time-series variation in efficiency levels. Journal of Econometrics 46(1):185–200
70. Diao X, Rattsø J, Stokke H E (2005) International spillovers, productivity growth and openness in Thailand: an intertemporal general equilibrium analysis. Journal of Development Economics 76(2):429–450
71. Feng Q, Horrace W, Wu G L (2012) Wrong skewness and finite sample correction in parametric stochastic frontier models. Mimeo
72. Glass A, Kenjegalieva K, Sickles R C (2013a) A spatial autoregressive production frontier model for panel data: with an application to European countries. Mimeo
73. Glass A, Kenjegalieva K, Sickles R C (2013b) Estimating efficiency spillovers with state level evidence for manufacturing in the U.S.. Mimeo
74. Good D H, Nadiri M I, Sickles R C (1997) Index number and factor demand approaches to the estimation of productivity. In: Pesaran M H, Schmidt P (ed) Handbook of Applied Economics, Volume II-Microeconometrics. Oxford, Basil Blackwell
75. Good D H, Roeller L H, Sickles R C (1995) Airline efficiency differences between Europe and the U.S.: implications for the pace of E.C. integration and domestic regulation. European Journal of Operational Research 80:508-518
76. Hultberg P T, Nadiri M I, Sickles R C (1999) An international comparison of technology adoption and efficiency: a dynamic panel model. Annales d'Economie et de Statistique 449–474
77. Hultberg P T, Nadiri M I, Sickles R C (2004) Cross-country catch-up in the manufacturing sector: impacts of heterogeneity on convergence and technology adoption. Empirical Economics 29(4):753–768
78. Kneip A, Sickles R C, Song W (2012) A new panel data treatment for heterogeneity in time trends. Econometric Theory 28: 590-628
79. Kumbhakar S C, Parmeter C, Tsionas E G (2013) A zero inefficiency stochastic frontier model. Journal of Econometrics 172(1):66-76
80. Liu J, Sickles R C, Tsionas E (2013) Bayesian treatment to panel data models with time-varying heterogeneity. Mimeo
81. Park B U, Sickles R C, Simar L (1998) Stochastic panel frontiers: a semiparametric approach. Journal of Econometrics 84(2):273–301
82. Park B U, Sickles R C, Simar L (2003) Semiparametric-efficient estimation of AR (1) panel data models. Journal of Econometrics 117(2):279–309
83. Park B U, Sickles R C, Simar L (2006) Semiparametric efficient estimation of dynamic panel data models. Journal of Econometrics 136(1):281–301
84. Sickles R C, Hao J, Shang C (2014) Panel data and productivity measurement: an analysis of Asian productivity trends. Journal of Chinese Economic and Business Studies 12(3):211-231
85. Sickles R C, Hao J, Shang C (2015) Productivity and panel data. In: Oxford handbook of panal data. Oxford University Press, New York.
86. Good D H, Nadiri M I, Roeller L H, Sickles R C (1993) Efficiency and productivity growth comparisons of European and U.S. air carriers: a first look at the data. In: Mairesse J, Griliches Z (ed) Journal of Productivity Analysis Special Issue 4:115-125
87. Stoker T M, Berndt E R, Ellerman A D, Schennach S M (2005) Panel data analysis of U.S. coal productivity. Journal of Econometrics 127:131-164
88. Van den Broeck J, Koop G, Osiewalski J, Steel M F (1994) Stochastic frontier models: a bayesian perspective. Journal of Econometrics 61(2):273–303