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

RISE Working Paper 15-011

"World Productivity Growth: A Model Averaging Approach"

by

Meryem Duygun, Jiaqi Hao, Anders Isaksson and Robin C. Sickles

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

World Productivity Growth: A Model Averaging Approach

Meryem Duygun
Business School
University of Hull
M.Duygun@hull.ac.uk

Jiaqi Hao
Credit Scorecard and Portfolio Management
ATB Financial
Edmonton Alberta
jhao@atb.com

Anders Isaksson
United Nations Industrial Development Organization
Vienna, Austria
A.Isaksson@unido.org

Robin C. Sickles
Rice University/Economics
rsickles@rice.edu

This version May 26, 2015

Abstract

The paper provides a discussion of panel data and productivity analysis in applied economic modeling. We discuss a variety of modeling scenarios and justifications for them based on classical economic theory and on more recent advances in production modeling, which formulate methods to decompose productivity growth based on a Solow-type residual (Solow, 1957) into innovation and catch-up. Methods to combine the various estimates based on different empirical specifications that model and estimate productivity growth are then discussed and these provide the econometric approaches we use to estimate world productivity growth. We also provide a counterfactual analysis of a scenario in which the rise in income inequality since the 1970's in the US is tempered by distributing productivity growth to wage compensation growth as had been the case during the post-WWII years to the early 1970's.

Keywords: Productivity; Panel Data; Economic Growth; .Time Varying Unobservable Effects

JEL Classification Numbers: D24, C23, O47.

1 Introduction

Modeling total factor productivity (TFP) growth is fraught with many challenges. Over the years, approaches towards productivity measurement have been developed to better address these many challenges. While some approaches relax theoretical assumptions such as perfect competition and constant returns to scale, others have evolved as a response to econometric issues. The arsenal of tools available to the researcher today is manifold and technically advanced and invites researchers to provide comparisons of results obtained from applying several methods. Unfortunately, the latter is seldom the case, as analysts tend to resort to one method only. The advantage of having a rich toolkit is, of course, a potential increase in accuracy with which we are able to measure productivity performance. However, a notable disadvantage is that these different measurement methods yield a range of estimates with sometimes very wide dispersion and, in the worst case, conflicting results. Furthermore, all models may be subject to misspecification of unknown form, researchers might have different information sets while presenting their own studies and models may be affected differently by structural breaks caused by institutional change or technological development, to name but a few possible reasons leading to variation in TFP measurement. The World Productivity Database (WPD) developed by the United Nations Industrial Development Organization (UNIDO) has tried to address this by providing productivity analysts across the globe with TFP estimates based on numerous methods, production function specifications, functional forms, different capital stock and labor input measures, and much more. However commendable such work may be, it is still silent on the issue of what the "correct" productivity estimate is.

Our paper demonstrates how one may get closer to, if not the "correct", then at least a "consensus" productivity estimate, which contains information gathered from a set of measurement methods. We contribute to the large productivity literature by demonstrating how results from various approaches can be combined, or averaged, to arrive at this consensus result. We discuss the many technical hurdles of such an exercise and how these can be appropriately handled. Importantly, we show that our results are robust to variations in how to deal with these technical hurdles, in particular that of the appropriate weights given to the different econometric methods and data definitions we utilize. We apply our methods to real data from the WPD. These data, in turn, are based on the Penn World Tables (Heston, Summers and Aten (2006). Our analysis, *inter alia*, includes results at the aggregate world level, compares the performance of six country groups at different stages of development, and decomposes TFP growth into change in technical efficiency and innovation, which provides policy makers with a richer and more detailed basis for policy making. Importantly, the analysis introduces a comparison of our consensus estimates with those provided by common approaches such as growth accounting, pooled and panel regression analysis, and data envelopment analysis. Our consensus estimates fare well in comparison and we conclude that it may be advisable to combine estimates in order to make the best conclusion based on all the available information.

The paper provides a discussion of panel data and productivity analysis in applied economic modeling. We discuss a variety of modeling scenarios and justifications for them based on classical economic theory and on more recent advances in production modeling that formulate methods to decompose productivity growth based on a Solow-type residual (Solow, 1957) into innovation and catch-up, the latter referred to as technical efficiency change in the stochastic frontier literature. We point to a number of innovations contributed to the panel data literature by those working in the stochastic frontier productivity discipline. In that literature the focus has been on the interpretation of relative temporal heterogeneity between production units (firms, countries, etc.) as a measure of relative technical efficiency in the use of the frontier technology. We point out why panel data are needed to identify and measure productive efficiency and innovation, which provides an additional link to the two strands of literatures. Our paper has a more aggregate productivity perspective, focusing on country level productivity, as it better motivates and displays the strong intellectual parallels between the efficiency literature, the economic growth and development literatures, and the literature on panel data econometrics.

The paper is organized in the follow way. We first discuss how productivity growth typically has been measured in classical productivity studies. We then briefly discuss how innovation and catch-up can be distinguished empirically. We next outline methods that have been proposed to measure productivity growth and its two main factors, innovation and catch-up. We show how such methods have particular canonical representations that seamlessly transfer to the panel data literature and briefly discuss competing specifications introduced into the productivity literature. We then point out how such models can be combined to provide consensus model average estimates of innovation and catch-up, utilizing and extending results from Hao (2011) and Sickles, Hao, and Shang (2015) on world productivity growth. Finally we utilize our consensus estimates to discuss and flesh out the patterns of income inequality in the US and provide a counterfactual analysis of scenarios in which rises in income inequality since the 1970's are tempered by a redistribution of productivity growth to wage compensation growth based on pre-1970's historical patterns.

2 Productivity Growth and Its Measurement

2.1 Classical Residual based Partial and Total Factor Productivity Measurement

Measurements of productivity rely on a ratio of some function of outputs (Y_i) to some function of inputs (X_i). To account for changing input mix, modern index number analyses utilize a measure of total factor productivity (TFP) for a single output technology that in its simplest form is a ratio of output to a weighted sum of inputs:

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

Following Jorgenson and Griliches (1972), a total factor productivity index can be constructed as the difference between log output and log input indices. The predominate TFP measure currently in use by the central governments in most countries is a variant of Solow's measure based on the Cobb-Douglas production function with constant returns to scale, $Y = AX_L^{\alpha}X_K^{1-\alpha}$ and leads to the TFP measure:

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

As is well-known, at cost minimizing levels of inputs, the parameter α describes the input expenditure share for labor and TFP growth is the time derivative of TFP:

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

Where multiple outputs exist, TFP can also be described as a ratio of an index number describing aggregate output levels(y_j) divided by an index number describing aggregate input levels(x_i). As a function of index numbers, TFP indexes derive many of their properties based the assumptions of the underlying aggregator functions used. Fisher (1927) laid out a number of desirable properties for these index numbers, many of which are easily achievable and many of which are not (Good, Nadiri, and Sickles, 1997).

3 Sources of Economic Growth, the Neoclassical and New Growth Theory Models

Debates among researchers on the primary sources of economic growth and development centered on two basic explanations that are rooted in the decomposition of economic growth sources: factor-accumulation and productivity-growth components. According to Kim and Lau (1994), Young (1992, 1995) and Krugman (1994), rapid economic growth in such emerging areas as East Asia was largely explained by the mobilization of resources. Alternative explanations to the neoclassical growth model explain economic growth not only in terms of intensive and extensive utilization of input factors but also due to factors that impact the degree to which countries can appropriate the productivity potential of world technical innovations . Again, factors such as governmental industrial policies, trade liberalization policies, and political, religious, and cultural institutions are often viewed as central to the ability of countries to catch-up with a shifting world production possibilities frontier.

Stiroh (2001) provides a coherent treatment of neoclassical theory that frames the problem of measuring sources of TFP growth in the context of the neoclassical production $Y = f(K, L, T)$ where variables are indexed by a time subscript. The production function is typically assumed to have constant returns to scale, positive and diminishing returns with respect to each input, and marginal products of each input that approach zero (infinity) as each input goes to infinity (zero). As noted by Stiroh (and many others), long run per capita output.growth is exogenously determined by technical change. The neoclassical growth model is not a model that explains long-run growth since technical change, which is the sole determinant of productivity growth, is determined outside the system.

A modification of the neoclassical growth model that addresses this problem leads to the so-called "new growth theory, wherein endogenous growth is introduced to weaken the strong neoclassical assumption that long-run productivity growth is only explained by an exogenously driven change in technology. The classic model put forth by Romer (1986), allowed for non-diminishing returns to capital due to external effects, such as research and development, that 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

$$Y = A(R)f(K, L, R)$$

In the "new" growth theory the production frontier is shifted by factors that are endogenous. The sources of the spillover differentiate many contributions to this literature. Arrow (1962) emphasized learning-by-doing, Lucas (1988) modeled A as a function of stock of human capital, Coe and Helpman (1995) introduced trade spillovers and showed that the rate of return on R&D was not limited to performing countries but extended to their trade partners. Diao *et al.* (2005) examined the impacts of both a protectionist alternative and shock liberalization and concluded that reduced openness had a negative impact on the overall growth rate due to reduced learning from the foreign spillover. Blazer and Sickles (2009) pursued the spatial effects of learning spillovers in their analysis of the determinants of rapid gains in the productivity of constructing "liberty ships" during World War II.

As pointed out by Abramovitz (1986), Dowrick and Nguyen (1989), and Nelson and Wright (1992), among many others, 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). The "new growth theory" implicitly recognizes the role of efficiency in production. However, if the explanation for the spillover of endogenously determined technology change is the loosening of constraints on the utilization of that 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 and Cigerli, 2009). One set of papers that provides an explicit efficiency interpretation for this growth process is Hultberg, Nadiri and Sickles (1999, 2004), and Ahn, Good, and Sickles (2000) who introduce inefficiency into the growth process. Of course the standard neoclassical model without explicit treatment of efficiency has been used by many authors in examining growth and convergence. An implication of the endogenous growth model is that if a time trend is added to the standard neoclassical production function then the trend must be stochastic. This clearly has implications for stationarity (Reikard, 2005). Recent work by Kneip, Sickles, and Song (2012) has addressed the estimation issues that are associated with estimating the endogenous technical change in the presence of technical efficiency change.

3.1 Alternative Explanations for Sources of Economic Growth-Explicit Modeling of Technical Efficiency via the Panel Stochastic Frontier Model

The sources of world economic growth using alternatives to the neoclassical model can be estimated by explicitly introducing the role of catch-up due to increases in the level of productive efficiency. Introducing the role of efficiency in production means introducing some form of frontier production process, such as the stochastic frontier production (Aigner, Lovell, and Schmidt, 1977, Meeusen and van Den Broeck, 1977). However, the inability of such cross-sectional approaches to identify efficiency differences in different countries and temporal changes in those measures, or catch-up, is a major drawback of the cross-sectional stochastic frontier. For such a model it is also necessary to specify parametric distributions for the idiosyncratic and inefficiency error terms in the composed error. This led researchers to quickly pursue panel data methods, such as those introduced by Pitt and Lee (1981) and Schmidt and Sickles (1984). These, however, did not allow the country effects to vary over time and the innovations by Cornwell, Schmidt, and Sickles (1990) and Kumbhakar (1990), Battese and Coelli (1992), Lee and Schmidt (1993) addressed this shortcoming. Kim and Lee (2006) generalized the Lee and Schmidt (1993) model by considering different patterns for different groups, thus eliminating their unrealistic restriction that the temporal patterns be the same for all firms. These new methods, allowing for time-varying and country specific efficiency change components as well as technical innovation changes, were applied in a variety of empirical settings. They decomposed total factor growth of 49 countries into technological change and technical efficiency change components and estimated the temporal pattern of productivity changes in certain regions and compared their regional characteristics. The results of their study show that technical efficiency had a significant positive effect on productivity growth. East Asia led the world in total factor productivity growth because technical efficiency gain there was much faster than that of other countries. Kalarajian, Obwona, and Zhao (1996) note that the key determinant of economic growth is not the level of input use but rather the method of application of inputs. They are able not only to rank TFP but also the technical efficiency over 45 countries. Findings from these many studies made it clear that technical efficiency, as a separate and contributing factor to innovation in determining productivity growth, was often a significant determinant of productivity growth and that a failure to properly distinguish between a shift in the technology and a movement to the best-practice of that technology could seriously bias productivity measurement and, subsequently, derived policy making.

The regression-based approaches to estimating sources of time varying and country specific total factor productivity growth utilize panel data methods in specifying time varying technical inefficiency captured by normalized (possibly time-varying) intercepts or fixed effects. Technical inefficiency can also be identified through normalized within residuals from error components models with the technical inefficiency effects. Moreover, parametric distributions can be assumed for such panel random effect models and maximum likelihood can be used. For example, a truncated normal distribution with time varying means can be specified as the one-sided error process for technical efficiency (Battese and Coelli, 1992). Cuesta (2000) generalized Battese and Coelli (1992) by allowing each country to have its own time path of technical inefficiency. The assumption of independence between inputs and technical efficiency is problematic as is the incidental parameters problem of MLE when fixed effects are assumed since the number of parameters increases with the sample size.

Proper specification of the catch-up process and the constraints on its adjustment speeds within a neoclassical growth model context also has been found to require a similar heterogeneous treatment of the catch-up, or technical efficiency growth, process. Hultberg, et al. (1999, 2004) modify the standard neoclassical convergence model to allow for such heterogeneity in the efficiency catch-up rates. In Hultberg et al. (2004) the relationship between growth in labor productivity of manufacturing sectors and transfers of technology from a leading economy to sixteen OECD countries is analyzed. In the standard catch up literature, the greater the gap in per capita income between low and high growth countries the faster the convergence occurs. However, this literature assumes identical technologies across countries. In addition to the existence of an external technology gap the ability to adopt new technology is an important source of growth. They also find that proper control for unobserved production heterogeneities is important in identifying the catching-up effect.

Hultberg et al.'s (1999, 2004) studies also are instructive in that the determinants of efficiency levels are proxied by a set of variables related to economic, political, and social institutions of a country. Their indicator variables are bureaucratic efficiency, which consists of three variables: judiciary system, red tape and bureaucracy, and corruption; political stability, which contains six indicators: political change–institutional, political stability–social, probability of takeover by opposition group, stability of labor, relationship with neighboring countries, and terrorism; economic openness, which consists of two measures of openness, the Sachs and Warner (1995) and Summers and Heston index. Hultberg *et al.* examine a second stage regression of efficiency on these aforementioned institutional variable proxies. Although the significance of individual variables is not widespread since there is often little country specific variation, these factors have an important combined effect in explaining the extent to which efficiency impacts the growth convergence. Upward of 60% of the variation in efficiency was attributed to the combined effects of the institutional constraint proxies.

4 Decomposition of Economic Growth-Innovation and Efficiency Change Identified by Index Numbers

Identifying the sources of TFP growth while imposing minimal parametric structure has obvious appeal on grounds of robustness. Sharpness of inferences may, however, be comprised vis-a-vis parametric structural econometric models. There has been a long standing tradition to utilize index number procedures as well as reduced form or structural econometric estimation to quantify TFP growth and its determinants. The essential difference between the approaches is discussed in Good, *et al.* (1997).

One common approach to decompose TFP into sources due to innovation and efficiency change based on the economic theory of index numbers is presented in a study of productivity growth in the OECD by Färe, Grosskopf, Norris, and Zang (1994) using an innovation and efficiency change decomposition based on the Malmquist index. Their method has been widely used and has many theoretical aspects to it that are quite appealing, although its statistical properties illustrate the difficulties in identifying significant sources of productivity growth while at the same time being sensitive to overly parametric assumptions. For example, utilizing bootstrapping techniques introduced by Simar and Wilson (2000), Jeon and Sickles (2004) found that there was no statistical significance to the productivity decompositions at standard nominal significance levels using the OECD data. Førsund and Hjalmarsson (2008) point out what they consider to be the main problem with the Malmquist index and its decomposition. The Malmquist index blurs the distinction between the ex ante micro function relevant for investments and the short-run production possibilities for the industry as a unit. When estimating technological change and technical efficiency change with the Malmquist index it is assumed that any producing firm may potentially produce at the frontier. According to Førsund and Hjalmarsson (2008), this would be the case only when there are no vintage effects, an assumption that could hold in industries where capital has a minor role, unlike paper, pulp, cement, etc. where the Malmquist index has been used to study productivity growth. In the case of disembodied technical change, wherein the shift in the production function over time is not incorporated into a specific best practice production function, the technical change in principle can only be relevant for existing units and thus the index cannot discriminate between efficiency change and disembodied technical change.

Grosskopf and Self (2006) calculate the Malmquist index and its decomposition into technical and efficiency change. They also provide estimates based on a neoclassical production approach with embodied technical change. In summarizing their findings Grosskopf and Self note that country differences are crucial in developing the proper structural interpretations for what are essentially reduced form correlations between factor accumulation and TFP growth on the one hand and economic growth in the region on the other. They also point out that "*... Growth is complicated; for a set of countries with apparently similar growth patterns, similar geographical location and relatively similar socioeconomic and cultural environments, we find complex and dissimilar explanations for their recent growth...*" (Grosskopf and Selof, 2006, p. 55).

4.1 Index Number Procedures

The Malmquist TFP index number procedure requires panel data to be implemented. Färe et al. (1994), among others, develop the methodology to construct and decompose a TFP measure based on the multi-output distance function $\mathrm{OD}_t(x_t, y_t)$. A special case of the multi-output distance function is the single output production function. An output efficient firm has a distance function score of 1 and it is not possible for the firm to increase its output without increasing one or more of its inputs. Conversely, an output inefficient firm has a distance function score that is less than one. The Malmquist TFP index requires output distance functions calculated between adjacent periods that are then manipulated and decomposed into a pure technological change component (A_{t+1}) and change in efficiency (catch-up) component (E_{t+1}).

$$M(x_{t+1}, y_{t+1}, x_t, y_t) = \frac{\mathrm{OD}_{t+1}(x_{t+1}, y_{t+1})}{\mathrm{OD}_t(x_t, y_t)} * \tag{3}$$

$$\left\{ \frac{\mathrm{OD}_t(x_{t+1}, y_{t+1})}{\mathrm{OD}_{t+1}(x_{t+1}, y_{t+1})} \frac{\mathrm{OD}_t(x_t, y_t)}{\mathrm{OD}_{t+1}(x_t, y_t)} \right\}^{1/2} \tag{4}$$

$$= E_{t+1} * A_{t+1}.$$

This index captures the dynamics of productivity change by incorporating data from two adjacent periods. E_{t+1}reflects changes in relative efficiency. A_{t+1}, reflects changes in technology between t and $t+1$. For the index, a value below 1 indicates productivity decline while a value exceeding 1 indicates growth. For the index components, values below 1 signify a performance decline while values above 1 signify an improvement. There are significant shortcomings of this approach noted by Førsund and Hjalmarsson (2009) due to potential vintage capital effects and/or its lack of any obvious inferential theory (Jeong and Sickles, 2004).

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

Regression based approaches to decompose productivity growth into technical change and 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 or translog or linear in levels generalized-Leontief or quadratic. These constitute the predominant functional forms used in productivity studies. Our many treatments for various forms of unobserved heterogeneity can be motivated with the following classical model for a single output technology estimated with panel data assuming unobserved country (firm) effects:

$$y_{it} = x_{it}\beta + \eta_i(t) + v_{it} \tag{5}$$

where $\eta_i(t)$ represents the country specific fixed effect that may be time varying, x_{it} is a vector of regressors, some of which may be endogenous and correlated with the error v_{it} or the effects $\eta_i(t)$.

We may interpret eq. (5) as the basic regression model that comes from the following transformation of the output distance function, which nests the single output production function that is predominately used in aggregate growth studies. We start 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):

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

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 linear homogeneous in outputs. Thus, after taking logs, adding 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, and redefining a few variables, the distance function can be written as (5) above.

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, where the second-order terms allow for greater flexibility, proper local curvature, and lift the assumed separability of outputs and inputs, can also be framed in this canonical model representation of a linear panel model with country-specific and time-varying heterogeneity. Any distance function that is linear in parameters, such as the translog, generalized Leontief or quadratic can be similarly written as a simple panel model that is linear in parameters and takes on the above canonical form. Of course 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 vehicle for estimating efficiency change using frontier methods that we will explore. 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. It 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 under nonpathological circumstances. 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 is also often proxied by exogenous or stochastic linear time trends (Bai, Kao, and Ng, 2009), which are often identified by nonlinear specifications of time varying inefficiency used in many of the approaches below or by orthogonality conditions.

There are many contributions to the efficiency and productivity literature that offer different ways to estimate this canonical panel model and to decompose TFP growth into a catch-up and innovation component. These include models introduced by Cornwell, Schmidt, and Sickles (1990), Kumbhakar (1990), Battese and Coelli (1992), Lee and Schmidt (1993), Park, Sickles, and Simar (1998, 2003, 2007), Greene (2005a,b), and Kneip, Sickles, and Song (2012). Space limits the possibility of dealing the many other approaches that have been proposed to estimate the panel stochastic frontier and provide a decomposition of TFP growth into innovative and catch-up, or technical efficiency. Some of the more general purpose estimators that have been proposed for panel stochastic frontiers that are also very appropriate to estimate the canonical panel productivity model are the Bayesian Stochastic Frontier Model (Liu, Sickles, and Tsionas, 2015), 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 (2014) and related models of Lee (1996), Lee and Lee (2012), and Orea and Steinbuks (2012), and the "True" Fixed Effects Model of Greene (2005a,b).

Below we use a number of these methods in a model averaging exercise to evaluate world productivity trends from 1970 to 2000.

6 Discussion on Combining Estimates

Combining estimates, or weighting estimates, provides a solution to modeling uncertainty. Sickles (2005) pursued this strategy in his examination of semiparametric and nonparametric panel frontier estimators. As discussed in details by Burnham and Anderson (2002), given that a model is appropriate, from a parametric approach we can use maximum likelihood methods or other methods, depending on how the model is specified, to estimate parameters in some optimal fashions. However, model selection uncertainty needs also be looked at more carefully. It should be considered the same as sources from other type of uncertainties, such as uncertainty due to the limited set of observation or model defect (Hjorth, 1994). Statistical inference based on "post-model-selection estimators" (Leeb and Potscher, 2005) may lead to invalid analysis. Moreover, different selecting criteria may often result in contradictory ranking orders. Also, focusing on one model and dismissing the results of alternative specifications may compromise the information content of the information set. If observed data are conceptualized as random variables, then the sample variability introduces uncertain inference from the particular data set (Burnham and Anderson, 2002). Moreover, due to the non-experimental nature of the data, model specification is very challenging to address in economics. Considering the complexity of economic and social structures, it is often unrealistic to find a correct or true model that fully recovers the underlying data generating process (DGP). In other words, all the existing models are misspecified in one way or another. Analogically, combining different misspecified models in some sense is similar to construct a diversified portfolio. Although each asset has negative price effects triggered by different factors, putting them together in a single basket would provide some benefits, for example, guaranteeing an overall risk-free return. Just as the famous quote attributed to Box, "*essentially, all models are wrong, but some are useful*", it is important to carefully design procedures to approximate the underlying DGP based on all possibly collected information (Box and Draper, 1987, p. 424).

Typical model selection from some encompassing supermodel can be viewed as a special case of weighting models which assigns the entire weight on one model and none on others. We do not pursue this approach in our empirical work below. Instead we utilize insights from economics and from statistics to motivate several canonical methods to combine estimates and forecasts from a variety of potentially misspecified models.

Economic approaches can be gleaned from work on *majority voting*. It is well known in the literature of social choice theory (see Moulin, 1980) that the median will be chosen as an outcome of majority voting. The median is of course a measure of the central tendency of the preferences of each voter and as a function of the stochastic registered preferences (or opinions of experts, as voters more often than not prefer to be thought of as experts) is simply an aggregator. If the voter opinions follow a symmetric distribution then of course the aggregator (or estimate, as it is a function of random variables) will be the simple average. The Tullock (1980) *contest function* also can be shown (Hao, 2011) to devolve in our setting into a decision rule in which expected outcomes are weighted averages of each estimates' R^2.

Insights from statistics come from *model averaging* and *combining forecasts*. Model averaging chooses a weighting scheme to average across various selected estimates. An alternative approach from the same root, model selection, is to select the best model among all available models according to some statistical criteria. However, it is never obvious to argue that any best-performed model is indeed the true model. Statistical inference based on the above "post-model-selection estimators" (Leeb and Potscher, 2005) might lead to invalid analysis. As argued in Buckland, Burnham, and Augustin (1997), the uncertainty of model selection should be incorporated into statistical inference. In analogy to sampling theories, if we consider our models in some sense as a valid random sample from an infinite set of possible models, combining information from different models would give us a more informative idea on the population parameters.

The important question of model averaging is how we can choose reasonable weights for each estimate in the process of combining them. The simplest way is to take an arithmetic mean of all estimates. However, it might not be always reasonable to assume that every model provides the same amount of information. The weights assigned to each model should reflect the extent of it supporting the data. So "goodness-of-fit " is a natural criterion to measure how data are supported by a model.

In the last four decades, many statistical criteria are developed under model selection context: For example, Akaike Information Criterion (Akaike, 1973), hereinafter AIC), Mallows' CP (Mallows, 1973), and the Bayesian Information Criterion (Schwarz, 1978), hereinafter BIC). Buckland, et al. (1997) used two of these information criteria, the Akaike and Schwarz, as weights in their model averaging exercise. There are broad literatures on conditions, limitations and asymptotic properties of each criterion. In addition to the Buckland et al. study, Hansen (2007) showed that the Mallows' Model Average estimator is asymptotically optimal in some cases and more favorable compared to AIC and BIC. Carroll, Midthune, Freedman, and Kipnis (2006) conducted a nutritional epidemiologic study and showed AIC achieved an efficiency gain, whereas BIC had serious issues and was not recommended. Burnham and Anderson (2002) and Claeskens and Hjort (2008) have more detailed discussions of the literature. Hansen and Racine (2012) considered situations in which candidate models are non-nested proposed a jackknife model averaging estimator, which they showed is asymptotically optimal in the sense that it approaches the lowest possible expected squared errors. Simulated comparisons of criteria have also been studied in different disciplines. Parmeter, Wan, and Zhang (2015) have begun to assess the finite sample properties of this estimator via Monte Carlo simulations.

Another interesting observation is that the model averaging with assigning weights according to variances coincides with Meta-Analysis when we regard efficiency as effect size. It is common in meta-analysis to weight the effect size according to inverse variance, which is referred to as the "inverse variance method."

Bayesian Model Averaging (hereinafter BMA) is developed in parallel with model averaging under classical framework. For detailed discussion of the framework and BMA techniques, see Raftery, Matigan, and Hoeting (1997), Hoeting, Matigan, and Raftery (1999), and Koop, Poirier, and Tobias (2007). However, the Bayesian technique is mainly developed to deal with linear models and generalized linear models with variable selection problems. In our situation, independent variables are fixed according to economic theories. Moreover, it is not clear that BMA or Bayesian model selection would perform better than other model averaging methods.

Several common assumptions of applying model averaging and meta-analysis are difficult to be defended empirically. One such assumption is the independence between each pair of studies. It is almost impossible for two researchers in the same field to conduct their studies without any shared resources: information source from Internet, or academia conferences, for instance. The other problem is that researchers have to include hundreds of models or an exhaustive literature review to ensure that their combinations have fully implied the unknown "true model" or the underlying DGP. However, it is still not convincing to us that the model discovered by this way is the underlying true model. In sum, what really matters is if we can efficiently utilize and make a reasonable conclusion based on all the information we have.

The second combining approach is developed in the literature of *combining* time-series *forecast* models. In the literature of forecast, researchers also combine studies for forecast improvement. As mentioned in Newbold and Harvey (2002), Bates and Granger (1969) urged that researchers should consider creating a combined forecast, possibly a weighted average of the individual forecast, when alternative forecasts are available. The importance of combining forecasts may be seen in Diebold and Lopez (1996). They propose that weighting relevant results can be viewed as a key link between short-run, aggregating available information of models we have, and longer-run, ongoing process of model development. This idea of combining forecast is comparable to our idea of aggregating estimates. In addition, one interesting observation is that the forecasts are often not independent because studies have correlated attributes such as having the same data set or the same coauthors, or including the same independent variables. Following the lines of their arguments on combining forecast, we can claim that our weighting criteria are also more optimal than individual estimate while viewing our estimates as "in-sample forecast".

Bates and Granger (1969) introduce the methodology of forecast combination. In their paper, clearly the results they attempt to combine are correlated since the outcomes are obtained by two different forecast methods but on the same data set. In their first weighting method, if the forecast errors σ_1^2, σ_2^2 from the two models are uncorrelated, to minimize the total error, the weights should be assigned as $\sigma_2^2/(\sigma_1^2+\sigma_2^2)$ and $\sigma_1^2/(\sigma_1^2+\sigma_2^2)$. The weighting will be a little bit more complicated if correlation is considered: weight for forecast 1 will be $(\sigma_2^2 - \rho\sigma_1\sigma_2)/(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)$. If the weights are decided as above, the variance of forecast error is no greater than the smaller values of the two variances. It is obvious that the bigger error variance result will receive smaller weights. If only two results are combined, the weights trivially are the same as one of our weighting criteria which we assign $1/\sigma_1^2$ to estimate 1 and $1/\sigma_1^2$ to estimate 2. The method they applied to forecast model can be used in our study since it is to minimize combined error, whether it is an out-of-sample forecast error or in-sample error. Generally speaking, all the weight selecting methods are based on some types of loss function which in turn rely on the differences between the realized outcome and the forecast outcome, such as a Mean Squared Error (hereinafter MSE) or Mean Squared Percent Error (hereinafter MSPE). If we choose the loss function as typical square of error, it would be perfectly reasonable to use "goodness-of-fit" criteria. The implementation will be identical to the model averaging we discussed earlier. New developments of choosing combining weights by applying automatic machine learning algorithms and methods to solve missing data problem have been made in recent literature, for details, see Lahiri, Peng and Zhao (2015).

Two relevant points are raised in both Clemen (1989) and Timmermann (2006). First, lower sums of mean squared error can be actually achieved by weights according to simpler assumptions, for example, by ignoring the correlation between models. Without correlation, weighting formula would be simplified in combining two studies, and it would be possible for cases involved more than two estimators. The second interesting observation also appears in many forecast combination studies: simple averaging are reported doing as well as other more complicated weight selecting methods in many empirical studies, even compared with most recent developed techniques (Lahiri, Peng and Zhao, 2015). Based on the two methodologies and empirical findings above, we combine our estimates by simple arithmetic average, R-square, RSS, AIC and BIC.

Empirical measures of forecast uncertainty have a significant role in macroeconomics and the monetary policy making process (Lahiri and Shang (2010). However, constructing measures of forecast uncertainty involves challenging methodological problems. In fact there is still no well-established theory to measure forecast uncertainty. Zarnowitz and Lambros (1992) define "consensus" as the degree of agreement among corresponding point predictions by different individuals, "uncertainty" as the diffuseness of the probability distributions attached by the same individuals to their predictions, and disagreement among individual as a proxy for uncertainty. Many studies have shed light on how to measure forecast uncertainty empirically by using different proxies. One example is a series of studies by Lahiri and Sheng (2010), Lahiri, Peng and Sheng (2015), and Lahiri, Peng and Zhao (2015). Following the method developed by Davies and Lahiri (1995), they obtain a panel data of multi-horizon forecasts from all individuals in different periods, then decompose forecast errors into two uncorrelated components: forecaster-specific idiosyncratic errors and aggregate shocks. In the studies they show that forecast uncertainty can be expressed as the sum of disagreement among forecasters according to their private information and the perceived variability of common aggregate shocks. They also point out that when a combined forecast is based on a criteria that minimizes the risk of the combined forecast, that from the standpoint of a policy maker the key uncertainty is not the variance of the average forecast but the variances of each of the individual forecasts. The most commonly used dispersion of an alternative forecast from the consensus forecast could underestimate uncertainty since it may fail to account for the variance of the aggregate shocks. However, they also show that during many situations where forecast environments are stable, disagreement is found empirically to be a reliable estimate for forecast uncertainty.

As for our combining estimation strategy, the studies mentioned above provide us an alternative perspective from which to consider uncertainty. The aggregate forecast uncertainty and the uncertainty derived from the sampling theory in the model averaging literature are interestingly linked, since both need to consider variations within each model and among models. Zarnowitz and Lambros (1992) provide several empirical arguments on why correlations across alternative forecasts should be considered and why they should not be considered. Both Zarnowitz and Lambros (1992) and Lahiri, Teigland, and Zaporowski, (1988) illustrate that the average variance of individual estimators represents a true measure of uncertainty. Lahiri and Sheng (2010) also give an interesting interpretation to the estimation of uncertainty without consideration of the correlations among experts. They note that the average of the individual forecast error variance should be used as the confidence an outside observer will have in a random drawn of a typical individual forecast from the panel of forecasters. They thus provide a rationale for presenting estimation results of the variances of combined estimators without addressing the correlation of individual forecasts.

7 Modeling World Economic Growth with the UNIDO Data

In order to better understand existing and upcoming patterns of world income levels, growth in per/capita income, political stability, and international trade flows, it is important to correctly measure countries' productivity growth. In addition, when gauging crucial economic figures such as these it is important to use methods which are robust to misspecification error. The following section deals with productivity growth measurement's robustness using a number of economic methodologies, and estimators consistent with them, to elucidate productivity growth. Our model averaging productivity procedures use the United Nations Industrial Development Organization (UNIDO) data over the period between 1970 and 2000. The issue of country heterogeneity is dealt with by separately analyzing countries grouped by their development features and using different panel data methods.

7.1 UNIDO Data Description

Information on measures of the level and growth of TFP based on 12 different empirical approaches across 112 countries over the period $1960-2000$ are provided by The World Productivity Database (WPD). The Penn World Tables version 6.1 (PWT, Heston, Summers and Aten, 2002) is the primary source of data. The Penn World Tables version 6.1 was used to obtain (chain weighted) GDP and investment in power purchasing parity 1996 US dollars. Data on employment and hours worked were taken from the Groningen Growth and Development Centre (GGDC, 2005) and Asian Development Bank (ADB, various issues).. The International Labor Organization (ILO) Yearbook 2003 was the source of unemployment figures, schooling data was taken from the ILO's Key Indicators of the Labour Market and ADB (various issues), Barro and Lee (2000) and the health indicators -life expectancy and adult mortality rates- were obtained from the World Development Indicators (World Bank, 2004). These are documented in the technical appendix to the WPD (Isaksson, 2007)

Capital input measurement is the most complicated. It can be argued that Capital (K) is the hardest feature to measure, which is why the WPD has made 4 different approaches available based on various computations for the initial capital stock, depreciation rates, depreciation schedules, and the lifetime of the asset. The various capital measures are labelled K06, K13, Ks and Ke_. The first three capital measures all have capital which is expected to depreciate at a constant rate over time. The first two capital stocks are only different with respect to their presumed depreciation proportions (6% and 13.3%, correspondingly, which resemble approximately 12 and 6 year asset lives). The different depreciation schedules highlight the significance of either the initial capital or the influence of recent investments. K06 and K13 are built on the assumption that ten years of investment act as a sufficient proxy for the initial capital stock K0. An additional popular method of computing the initial capital stock is to presume that the country has reached its steady state capital-output ratio, which results in a steady-state capital stock (Ks) whose anticipated depreciation rate is 6% per annum. Another method of gauging capital exmaines the profile of capital productivity and uses a time-varying depreciation rate. The productive efficiency of the asset falls accumulatively as the asset ages which results in Keff.

Two types of labor utilization rates for which labor force can be adjusted are involved in *Labor input* measurement: differences in employee numbers and working hours. Employment (EMP) is the first measure for labor force (LF) and is a direct measure of employment. The second is achieved by using unemployment rates on LF data, resulting in derived employment (DEMP).

7.2 Empirical Findings

We adopted the approach taken by Hulten and Isaksson (2007) who divided all 112 countries in the WPD into six mutually exclusive groups, in accordance with the World Bank classification by income per capital. The group of Low Income countries (hereinafter LOW) is comprised of 40 countries, 22 countries in the group of Lower-Middle Income countries (hereinafter LOW-MID), 17 countries in the Upper-Middle Income countries (hereinafter UPPER-MID), 24 High-Income countries (hereinafter HIGH), 4 Old Tigers (the original Asian Four Tigers) and 5 New Tigers. The countries are:

Low Income Countries: Angola, Bangladesh, Benin, Bolivia, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Comoros, Congo, Cote d'Ivoire, Democratic Republic of the Congo, Ethiopia, Gambia, Ghana, Guinea, Guinea Bissau, Haiti, Honduras, Kenya, Lesotho, Madagascar, Malawi, Mali, Mauritania, Mozambique, Nepal, Nicaragua, Niger, Nigeria, Papua New Guinea, Rwanda, Senegal, Sierra Leone, Tanzania, Togo, Uganda, Zambia, and Zimbabwe.

Lower-Middle Countries: Algeria, Cape Verde, Colombia, Costa Rica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Fiji, Guatemala, Guyana, Iran, Jamaica, Jordan, Morocco, Namibia, Pakistan, Paraguay, Peru, Philippines, and Sri Lanka.

Upper-Middle Countries: Argentina, Barbados, Botswana, Brazil, Chile, Gabon, Mauritius, Mexico, Panama, Seychelles, South Africa, Syria, Trinidad and Tobago, Tunisia, Turkey, Uruguay, and Venezuela.

High-Income Countries: Australia, Austria, Belgium, Canada, Cyprus, Denmark, Finland, France, Greece, Iceland, Ireland, Israel, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, and USA.

Old Tigers: Hong Kong (SAR of China), Republic of Korea, Singapore, and Taiwan (Province of China).

New Tigers: China, India, Indonesia, Malaysia, and Thailand.

K06, K13 and Keff were selected as the different capital inputs used in our analyses. As a result of limitations in the availability of data, we used LF as labor input. Thus, every country group contains 3 input combinations. Ten estimating models were used: Cornwell et al. (1990) efficient instrumental variables estimator (EIV) and time varying random effects estimator (CSSG), the Battese and Coelli (1992) time varying random effects estimator (BC), a set of Park, Sickles, and Simar (1996) semi-nonparametric efficient estimators (PSS1), the Park, Sickles, and Simar (1999) semi-nonparametric efficient estimator with serially correlated errors (PSS2W) and (PSS2G), standard fixed effects with no temporal variability in the effects (FIX1), standard random effects with no temporal variability in the effects (RND1), comparable estimates with a quadratic and linear time trend for productivity growth (FIX2) and (RND2). The period of observation was between 1970 to 2000. Detailed estimation results used in the model averaging are in the Appendix.

TFP was split into technical efficiency change and innovation change. Figures $1-6$ display the outcomes of technical efficiency variations. Efficiency growth outcomes are presented in each figure based on three different weighting systems: arithmetic, geometric, GDP share. According to the graphs, all the models concur that LOW had substantial efficiency enhancements from 1970 to the early 1980s. At this point the growth rates diminished, the indicators of which differed depending on the models used. At the end of the sample period technical efficiency was showing a decline. LOW-MID has efficiency change patterns akin to LOW, the only difference being that the efficiency decreased at the turn of the 1980s. In comparison to the last two groups, although the magnitudes are less significant, UPPER-MID has longer-lasting yearly progresses and the waning of efficiency improvements occurred up to the middle of the 1990s . HIGH has minor degrees of efficiency development until the end of the 1980s. At this point the models differ in terms of the direction of efficiency variation change, nonetheless the degrees for all the models were reduced. The efficiency estimates for Old Tigers show a greater level of diversity. A reduction in efficiency was detected at the beginning of the sample period, whilst a pattern of increasing efficiency was visible towards the end of the sample period. The efficiency trends for the New Tigers is akin to LOW, but are smaller in magnitude.

Figures $7-12$ shows the outcomes for technical innovation differences in all of the groups. The Old Tigers hold the greatest average innovation growth amongst the 6 groups at 3.46%. The New Tigers are close with a 2.63% yearly growth rate in innovation. The UPPER-MID and HIGH income nations have average technical growth rates of 0.69% and 0.64% per annum. LOW-MID nations display very little if any improvements and LOW income nations show a slight technological regress of -0.47% per year.

In our analyses, TFP growth is based on the contributions of the growth in technical efficiency and the growth in technical innovation. Figure 13 shows that the Old Tigers have the highest TFP improvements over the observation period due to their high level of technical development. The New Tigers are next, again as a result of their innovation developments, which overshadow the measured efficiency contributions. Figure 14 shows the average TFP growth rates by level of development. HIGH and UPPER-MID income nations have modest TFP growths annually; however the figures show a pattern of slight decline. LOW-MID and LOW income nations display downward trends as a result of a decline both in technical efficiency growth and growth in innovation. The projected accumulated TFP growth between 1970 and 2000 is 15.4% for LOW income nations, 10.1% for LOW-MID income nations, 27.7% for UPPER-MID income nations, 17.2% for HIGH income nations, 199.7% for the New Tigers, and 239.4% for the Old Tigers. We combined all six groups to establish the world's TFP growth. This can be seen in Figure 15, where there is a visible decline in the world's TFP growth rate.

In addition, we compared our model averaged TFP results (using GDP share weights) with results presented on the UNIDO WPD website. Five approaches were selected to match against our findings. These are labelled as Growth Accounting, Hicks Neutral, Pooled Regression, Panel Regression, Stochastic Frontier Model, and Data Envelopment Analysis (DEA). We selected identical nations as well as identical input groupings. Figures $16-21$ display TFP growth rates for nations in the various developmental classifications. Our model averaging results indicate smoother patterns as a result of the averaging. Four out of our average TFP growth rates surpassed the ones projected from the five approaches used on the UNIDO WPD website.

The last results we discuss are the combined estimates and their distribution within the differing country groupings we use in our analysis. As discussed above, the motivation for employing the model averaging methodology is to obtain consensus results based on all the modeling and data information at hand. The annual changes of technical efficiency, technical innovation and TFP for each individual group based on different weighting methods are shown in Table 1. The most crucial component the model averaging exercise is the assignment of weights to each set of estimates. The simplest averaging is to take the arithmetic mean of all estimates, which implicitly assumes equal importance of all models. Besides simple averaging, we use other four statistical criteria that have been developed to measure model fit and to assess model specification in order to assign weights. These are the (adjusted) R^2, the inverse of the residual sum of squares (RSS), the Akaike's information criterion (AIC), and the Bayes Information Criteria (BIC). Weights based on arithmetic averages for annual changes of technical efficiency, technical innovation and TFP are $0.64\%, -0.17\%$ and 0.47% for LOW, $0.35\%, -0.04\%$ and 0.32% for MID-LOW, $0.15\%, 0.64\%$ and 0.79% for UPPER-MID, $-0.06\%, 0.57\%$ and 0.51% for HIGH, $0.30\%, 3.63\%$ and 3.93% for the Old Tigers, and $-0.09\%, 2.95\%$ and 2.86% for the New Tigers. Since (adjusted) $R^{2\prime}s$ for our different specifications are rather close to each other, the weighted results are quite similar. The combined estimates display a relatively stable pattern during the 31 years covered in the sample period. Old Tigers lead the world in TFP improvement due mainly to substantial technological innovations with technical efficiency change having a small positive contribution. New Tigers were next with quite respectable TFP growth to innovation change with no economically significant contribution by technical efficiency change. LOW income countries had the greatest efficiency gains among all groups but their poor innovation growth leads to poor overall TFP performance. LOW-MID and UPPER-MID countries have essential no progress in either technical efficiency or innovation. HIGH income countries face little growth in innovation as well as a slight regress in their levels of technical efficiency.

The variances of the combined estimates can also be calculated under the model averaging framework and thus the statistical significance of the consensus estimates can also be determined. In this way such consensus estimates provide an advantage over the index number measures that are usually presented simply as a point estimate. Burnham and Anderson (2002), and more recently Huang and Lai (2012), have provided discussions on how to such variances can be computed. The difficult components to estimate are the correlations between each pair of estimators. For example, in our case each estimate of TFP growth per period results from each statistical model with one combination of inputs. Bootstrap methods have been suggested. However, bootstrapping of data might not be valid in many situations. In situations where correlations cannot be estimated directly, an upper bound on the variance can still be obtained assuming all correlations are 1.0 (Huang and Lai, 2012). However, in practice estimating sample correlations between each pair of estimates is not difficult in our study because of our panel-data setting. We can calculate sample correlations between each pair of TFP estimates directly as we have estimates in each period.[1] The estimated variances and variance bounds for the model averaged results are also presented on Table 1.

Data Envelopment Analysis (DEA), a non-parametric approach utilizing mathematical programming, is an alternative to the regression-based models we have employed to estimate TFP growth and its decomposition into growth in innovation and growth in efficiency. One significant advantage of DEA compared to regression-based methods is that DEA does not need to specify a functional form of production technology. However, a shortcoming of DEA is that the constructed production frontier is biased, which in turn results in downward biased efficiency estimates. Badunenko *et al.* (2013) use bootstrap procedures (Simar and Wilson, 2000; Jeon and Sickles, 2004) to construct an unbiased production frontier. Following Henderson and Russell (2005), Badunenko *et al.* decompose the productivity growth into 4 components: change in efficiency (technical efficiency change), technological change (innovational progress), capital deepening (Kumar and Russell, 2002) and human capital accumulation. The output, capital and labor data are derived from the Penn World Tables version 6.2 (Heston *et al.,* 2006) and their human capital data are the education data obtained from Cohen and Soto (2007).

Although the country groupings differ between our study and those of Badunenko et al. and Henderson and Russell, we can highlight similarities and differences in our study theirs based on the alternative DEA methodology[2]. Between 1965 and 2000, they report total productivity (output per worker) increases of 114.8% (annualized at 2.10%) in Asia (notice their included countries in this region are significantly different from ours), 26.6% (annualized at 0.66%) in Latin America, and 110.2% (annualized at 2.10%) in the OECD. Even though the two studies employ different methodologies on different data sets, they both find that countries in Asia have the highest TFP growth, those that constitute the OECD have the second highest TFP growth rates, and countries in Latin (and South) America have the lowest TFP growth rates. Another common finding is that technical innovation contributes significantly more than the efficiency gains to TFP growth in Asia and the OECD. That is, shifts in the production possibility frontier outweigh the catch-up effect in the generation of TFP growth. Since TFP growth is much smaller than the growth of GDP per capita through the sample period, other factors such as such as physical and human capital accumulation play a relatively larger role in world economic growth.

8 TFP Growth and Inequality

We have found that productivity growth has differential patterns across countries with different levels of economic development. Such heterogeneity in the growth of the wealth of nations appears to be related systematically to initial levels of development and points to a potential leveling of between-country standards of economic well-being. Whether or not differential TFP growth rates among developing and developed countries will continue at the rates we have observed during our sample period is not clear. Future world economic growth dynamics will be determined by many factors that our study is not able to capture. However, one important factor on which most economists agree is that increased income inequality within a country will have important implications for long term growth and of course for political stability. While our reduced form models and statistical treatments can only provide a consensus estimate of TFP growth by level of development, and are not designed to conduct counterfactual analyses, we can design interesting future growth scenarios that are calibrated in part by our consensus estimates.

One such analysis is considered in this section and addresses a counterfactual question involving

[1] Correlations cannot be computed for models with time-variant TFP estimates. In order to provide a conservative variance estimate set the correlations involving these time-invariant estimates at 1 in generating the results reported in Table 1.

[2] In Asian countries, they have India, Indonesia, Iran, Jordan, Malaysia, Nepal, Syria and Thailand. In Latin America, they have all the 11 countries we have in addition to 9 other countries. In OECD countries, Iceland and Luxembourg are not included in their data set, Mexico is not included in our data set.

the decoupling of aggregate TFP growth and the growth in real wage compensation, a relationship that at least in the US had been relatively stable during the post WWII years until the mid to late 1970's, at which time the growth in wage compensation began to substantially lag behind aggregate TFP growth. Such a decoupling of wage compensation and wealth creation has clear implications for the substantial changes that have occurred in the wealth distributions in the US and in many other countries in the developed and developing world. A number of recent contributions to the literature on income equality (Piketty, 2013, Stiglitz, 2012, among others) have pointed to the reasons for such a decoupling and the implications of such. In the U.S. a recent report from the Congressional Budget Office (CBO) (2014) identified rising inequality in the US as unsustainable. Real income in the bottom 90% of households changed very little in the last three and a half decades but consumption growth did not. This was largely financed by excessive borrowing, which of course further constrained spending. The CBO report indicated that the deleveraging due to excessive borrowing by this group that accompanied the 2007-2009 recession explains much for the slow recovery. Using the Current Population survey from 1976 to 2012, Cortes, Jaimovich, Nekarda, and Siu (2014) showed how the recession accelerated the displacement of the mid-wage workers engaged in jobs requiring repetitive tasks by computers, robots and other machines. These sorts of permanent job displacements have differentially impacted young, less educated men. Autor and Dorn (2013) note that routine tasks following well-defined procedures, jobs that typically are mid-wage jobs, are particularly vulnerable to competition from computers and non-routine jobs requiring interpersonal skills, social IQ, flexibility, and problem-solving talents. At least in the US it has been argued that changing demographics, off-shoring, unions and the minimum wage have played an important role in the loss of mid-wage jobs. However, only 1/3 of the job losses could be attributed to these factors, the rest are due to the decline in routine jobs. Brynjolfsson and McAfee (2014) point out that such 'non-routine' jobs as elder care, wealth management, and even art are being replaced by machines. It is necessary to distinguish between jobs that are viewed as routine and thus able to be replaced with technology and those that have non-routine aspects to them that may protect in some way from a complete substitution by technological innovation.

Income inequality that may have its cause in the disparate control and ownership of wage and capital services continues to grow and shows no sign of abating. One of the primary sources of this increase, lack of availability and poor quality of education for those that cannot afford it, persists and the persistence is systematic and unequal (United Nations Human Development Report (UNHDR), 2014). Although overall inequality, measured by the Inequality-Adjusted Human Development Index (IHDI), has declined slightly in most regions, this mainly has been a function of health care, in the form of better hygiene, other public health improvements, and nutrition. Recent estimates of the United Nations development Programme (UNDP) Multidimensional Poverty Index are that 1.5 billion people in 91 developing countries are living in poverty with overlapping deprivations in health, education and living standards. Of course many of these developing countries have at best rather frail democratic institutions and little in the way of democratic traditions. However, one could ask why voters are not angrier about economic inequality? An answer to this question is provided by researchers at the University of Hannover. Using data from the International Social Survey Programme, in which respondents were asked to locate their relative income status on a scale of 1 to 10, Engelhardt and Wagener (2014) built a measure of perceived inequality, defined as the gap between the median income and the average income of the population. In every one of the 26 nations they studied, most of them in the developed world, people believe that the income gap is substantially smaller than it really is. To some extent these issues provide a backdrop to the analysis we provide in this section.

What is the income distribution for the US relative to other countries? The distribution of Gini coefficients across the world is represented in the Figure 22. The position of different countries in the distribution may shift depending on the data source but the shape is generally quite robust. In 2007 the CIA estimated a Gini of 45 for the US, while the OECD estimated it to be about 40 at roughly the same time period. Countries in the left tail include Austria, Norway, and Sweden. Countries in the right tail include Namibia, South Africa, and Botswana. Countries close to the US are (Islamic Republic of) Iran, Cambodia, Bulgaria, and Jamaica. A quite relevant question involves how such remarkable inequality came about and how might the US return to an income distribution that is more in line with the sentiments expressed by it populace. For example, According to Norton and Ariely (2011), people in the US prefer an income distribution much like Sweden.

Consider the following counterfactual thought experiment that links the increasingly skewed US income distribution and the resulting increase in its Gini coefficient to the distribution of productivity growth and to wage compensation. To do this we first calculate the difference in the growth rates of TFP and wage compensation in the group of nations whose level of development is comparable to U.S., the OECD. We use the growth rates of TFP calculated above. Wage compensation in the OECD includes: (i) Wage rates (basic wages, cost-of-living allowances, and other guaranteed and regularly paid allowances), (ii) Overtime payments, (iii) Bonuses and gratuities regularly paid, (iv) Remuneration for time not worked, (v) Payments in kind. Data ranges from 1971 to 2009.[3] We then generate a comparable series for the Gini coefficient. These data range from 1985 to 2005. All the data are GDP weighted. The weighted averages are constructed in constant prices (using exchange rate and PPPs). For the Euro area countries, the data in national currency for all years are calculated using the fixed conversion rates against the Euro. We next examine the relationship between the growth in TFP minus the growth in real wage composition and the level of the Gini coefficient for the OECD countries. The coefficient of such a simple regression is 0.23 ($R^2 = 0.52$) and is significant at the 0.0025% level.

We next examine the differential growth rates of TFP and wage compensation based on Economic Policy Institute Figure 23. These are based on U. S. Bureau of Labor Statistics (BLS) and U. S. Bureau of Economic Analysis (BEA) data. We then calculate the implied change in the Gini coefficient were the differential growth rates not to have decoupled in the mid-$1970\prime s$. We find that the U.S. Gini coefficient–calculated after taxes and transfers– of about 0.40 in 2009 would have been reduced by about 0.17 in 2009 had the growth rates of TFP and wage compensation been comparable, as they had been from 1947 to the $\text{mid}-1970\prime s$. This calculation is based on taking the approximately 72% difference in growth rates between TFP and average hourly compensation between 1980 and 2009 and multiplying the difference by the regression coefficient 0.23. The U. S. income distribution would be comparable to Sweden's Gini coefficient of about 22 had wage compensation growth and TFP growth continued to trend from the mid/late-1970's to first decade of new millennium as it had during the years after WWII to the mid-1970's.

Similar analyses can be conducted over all countries in our sample and is a research project that we anticipate pursuing in the near future.

9 Conclusions and Suggestions for Future Research

Estimating productivity growth and its components is a difficult endeavor. Over time, many approaches to such measurement have been developed, all with their pros and cons but none that may be argued to be "correct". While UNIDO's World Productivity Database (WPD) is trying to resolve this by providing TFP estimates from a wide range of methods, specifications and functional forms, to name but a few, an alternative that has recently evolved is to employ the model average methodology to obtain a consensus result.

In this paper we have focused on the role that panel data econometrics plays in formulating and estimating the most important contributors to productivity growth: innovation and catch-up. We have explained different theories on economic growth and productivity measurement and the econometric specifications they imply. Various index number and regression-based approaches to measuring productivity growth and its innovation and catch-up components have been discussed in detail. We have also discussed methods that can be used to combine results from the many different perspectives on how economic growth is modelled and estimated, focusing on methods used in model averaging and in the combination of forecasts. We have utilized various panel data and model averaging methods in an analysis of world productivity growth using the WPD and have

[3] Data for Netherlands and Greece are not available and thus are countries are excluded. Data for some other countries are not available in earlier years. From 1971-1980, Australia, Iceland, Korea, Luxembourg, New Zealand, Portugal, Spain, Turkey are excluded. Luxembourg is included since 1980; Spain is included since 1981; Australia is included since 1984, Turkey is included since 1988; New Zealand is included since 1989; Korea is included since 1992; Portugal is included since 2000; Iceland is included since 2005. Australia's data at 2008 and 2009 are estimated.

analyzed as well the changes in the US income distribution that would have resulted by relinking the growth of TFP with the growth in wage compensation, which were connected so strongly in the post-WWII to early $1970\prime s$ era.

The motivation for employing the model averaging methodology is to arrive at a consensus result based on all the modeling and data information at hand. To this end, we started by creating country groups based on income levels (LOW, LOWER-MID, UPPER-MID and HIGH), however singling out two fast-growing groups of developing countries-Old and New Tigers-producing a total of six groups of countries. Based on results from ten estimating models and three different capital definitions (a total of 30 different sets of estimates), we found that Old Tigers have the highest TFP improvements thanks to the group's relatively large level of technical progress. New Tigers were the second best performer, despite some small deterioration in technical efficiency over the sample period. Both HIGH and UPPER-MID display negative trends in TFP growth, but still manage moderately positive TFP growths every year. The worst outcome is shown for LOW-MID and LOW because of lack of technological innovation and decline in technical efficiency. When aggregating these results to obtain a world average we found a declining trend in TFP growth rates.

In a second step, we compared our average estimate with those obtained from growth accounting (Hicks-neutral), Pooled Regression, Panel Regression, Stochastic Frontier Model and Data Envelopment Analysis. Compared to these models, our estimates display smoother trends, with four of the ten estimating models showing higher TFP growth than the five comparators. It is reassuring that our results are robust to different approaches to weighing each set of estimates, which is the most crucial component for combining estimates.

While our paper is an important step in the direction of obtaining a consensus TFP result much work remains. For example, in this paper we have used then estimating models but why not use more models? Secondly, in this paper we have not experimented much with other sources of TFP growth variation such as additional production factors (e.g., human capital, health capital and land) or functional forms (e.g., Cobb-Douglas or CES). Given that different countries are at various stages of development, applying different properties of the production function might be an important future step; at least it would be good to see an attempt being made in that direction. Thirdly, our choice of country groups is not cut in stone and there may be scope for experimenting with groupings based on other criteria than income alone, For example, it is likely that countries significantly based on natural resources such as the Gulf States or countries in transition (Eastern Europe) might display alternative performance patterns. Similarly, as more country data and longer time series become available, opportunities for even richer and nuanced analysis may arise. It is in the direction of exploiting the next generation of WPD with significantly richer data that we are heading towards next.

10 References

Abramovitz, M., "Catching Up, Forging Ahead, and Falling Behind,"*Journal of Economic History,* 46, 385–406, 1986.

Ahn, S. C., Good, D. H., and R. C. Sickles, "Estimation of Long-run Inefficiency Levels: A Dynamic Frontier Approach," *Econometric Reviews,* 19, 461–492, 2000.

Aigner, D., Lovell, C. A. K., and P. Schmidt, "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics* 6, 21–37, 1977.

Akaike, H., "Information Theory and an Extension of the Maximum Likelihood Principle," in B. N. Petrov and F. Csaki, eds., *Second International Symposium on Information Theory,* 267-281, Akademiai Kiadó, Budapest, 1973.

Almanidis, P., Qian, J. and R. C. Sickles, "Stochastic Frontiers with Bounded Inefficiency," *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications,* R. C. Sickles and W. C. Horrace, eds., 47-82, Springer Science & Business Media, New Yorki, 2014.

Arrow, K. J., "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 29, 155–173, 1962.

Autor, D. and D. Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market," *American Economic Review*, 103, 1553–1597, 2013.

Bai, J., Kao, C. and S. Ng, "Panel Cointegration with Global Stochastic Trends," *Journal of Econometrics*, 149, 82–99, 2009

Balk, B. M., *Price and Quantity Index Numbers Models for Measuring Aggregate Change and Difference*, Cambridge University Press, Cambridge, MA, 2008.

Baltagi, B. H. and J. M. Griffin, "A General Index of Technical Change," *Journal of Political Economy*, 96, 20–41, 1988

Barrow, R. J. and J.-W. Lee, "International Data on Educational Attainment Updates and Implications," *National Bureau of Economic Research Working Paper No. 7911*, Cambridge, MA, 2000.

Bates, J. M. and C. W. Granger, "The Combination of Forecasts," *Operations Research Quarterly,* 20, 451–468, 1969.

Badunenko, O., Henderson, D. J. and R. R. Russell, "Polarization of the Worldwide Distribution of Productivity," *Journal of Productivity Analysis*, 40, 153–171, 2013.

Blazek, D. and R. C., "Knowledge and Geographical Spillovers in Wartime Shipbuilding," *Economic Modeling*, 27, 1484–1497, 2009.

Box, G. P. and N. R. Draper, *Empirical Model-Building and Response Surfaces*, Wiley, New York, 1987.

Brynjolfsson. E. and A. McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W.W. Norton & Company, Inc., New York, 2014.

Buckland, S. T., Burnham, K. P., and N. H. Augustin, Model Selection: An Integral Part of Inference, *Biometrics*, 53, 603-618, 1997.

Burnham, K. P. and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, Springer Verlag, New York, 2002.

Carroll, R. J., Midthune, D., Freedman, L. S. and V. Kipnis, V., "Seemingly Unrelated Measurement Error Models, With Application to Nutritional Epidemiology," *Biometrics*, 62, 75–84, 2006.

Caves, D. W., Christensen, L. R. and W. E. Diewert, "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica*, 6, 1393–1414, 1982.

Charnes, A., Cooper, W. W. and E. Rhodes, "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, 2, 429–444, 1978.

Claeskens, G. and N. L. Hjort, *Model Selection and Model Averaging (Cambridge Series in Statistical and Probabilistic Mathematics)*,Cambridge University Press, New York, 2008.

Clemen, R. T.,"Combining Forecasts: A Review and Annotated Bibliography," *International Journal of Forecasting,* 5, 559–583, 1989.

Coe, D. T. and E. Helpman, "International R&D Spillovers," European Economic Review, 39, 859–887, 1995.

Coelli, T., "Econometric Estimation Of The Distance Function Representation of A Production Technology," CEPA, University of New England, 2000.

Coelli, T. and S. Perelman, "Efficiency Measurement, Multiple Output Technologies and Distance Functions: With Application to European Railways," *European Journal of Operational Research,*117, 326-339, 1996.

Cohen, D. and M. Soto, "The Marginal Product of Capital,"*Journal of Economic Growth*,12, 51-76, 2007.

Congressional Budget Office, *The Budget and Economic Outlook: 2014-2024*, CBO,Washington, DC, 2014

Cornwell, C., Schmidt, P. and R. C. Sickles, "Production Frontiers with Cross-sectional and Time-series Variation in Efficiency Levels," *Journal of Econometrics*, 46, 185–200, 1990.

Cortes, G. M., Jaimovich, N., Nekarda, C. J., and H. E. Siu, "The Micro and Macro of Disappearing Routine Jobs: A Flows Approach, *National Bureau of Economic Research Working Paper No. 20307*, Cambridge, MA, 2014.

Cuesta, R. A., "A Production Model With Firm-Specific Temporal Variation in Technical Inefficiency: With Application to Spanish Dairy Farms," *Journal of Productivity Analysis*,13, 139–158, 2000.

Davies, A. and K. Lahiri, "A New Framework for Analyzing Survey Forecasts using Three-dimensional Panel Data," *Journal of Econometrics*, 68, 205–227, 1995.

Diao, X., Rattsø, J. and H. E. Stokke, " International Spillovers, Productivity Growth and Openness in Thailand: An Intertemporal General Equilibrium Analysis," *Journal of Development Economics*,76, 429–450, 2005.

Diebold, F. X. and J. A. Lopez, "Forecast Evaluation and Combination," *National Bureau of Economic Research Working Paper No. 192*, Cambrisdge, MA, 1996.

Dowrick, S. and D.-T. Nguyen, "OECD Comparative Economic Growth 1950-85: Catch-up and Convergence," *American Economic Review,* 79, 1010–1030, 1989.

Engelhardt, C. and A. Wagener, "Biased Perceptions of Income Inequality and Redistribution," Discussion Paper, Wirtschaftswissenschaftliche Fakultät,Leibniz Universität Hannover, No. 526, 2014

Färe, R., Grosskopf, S., Norris, M. and Z. Zhang, "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries," *American Economic Review,* 84, 66–83, 1994.

Fisher, I., *The Making of Index Numbers: A Study of Their Varieties,Tests, and Reliability,* Houghton Mifflin Company, Boston, 1927.

Førsund, F. R. and L. Hjalmarsson, "Dynamic Analysis of Structural Change and Productivity Measurement," Technical Report, Memorandum, Department of Economics, University of Oslo, 2008.

Good, D. H., Nadiri, M. I. and R. C. Sickles, "Index Number and Factor Demand Approaches to the Estimation of Productivity," in M. H. Pesaran and P. Schmidt, eds., *Handbook of Applied Economics, Volume II-Microeconometrics*, Basil Blackwell,Oxford, 14-80, 1996.

Greene, W., "Fixed and Random Effects in Stochastic Frontier Models," *Journal of Productivity Analysis*, 23, 7–32, 2005a.

Greene, W., "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model," *Journal of Econometrics,* 126, 269–303, 2005b.

Groningen Growth and Development Centre, *GGDC Data Bases,* University of Groningen, Netherlands, 2005.

Grosskopf, S. and S. Self, "Factor Accumulation or TFP? A Reassessment of Growth in Southeast Asia," *Pacific Economic Review,* 11, 39–58, 2006.

Hansen, B. E., "Least Squares Model Averaging," *Econometrica,* 75, 1175–1189, 2007.

Hansen, B. E. and J. S. Racine, "Jackknife Model Averaging," *Journal of Econometrics*, 167, 38-46, 2012.

Hao, J., *Essays on Productivity Analysis*, Unpublished Ph.D. Dissertation, Rice University, Houston, 2011.

Henderson, D.J. and R. R., Russell, "Human Capital and Convergence: A Production-Frontier Approach," *International Economic Review*, 46, 1167-1205, 2005.

Heston A., Summers R. and B. Aten, *Penn World Table Version 6.1*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, Philadelphia, 2002.

Heston A., Summers R. and B. Aten, *Penn World Table Version 6.2*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, Philadelphia, 2006.

Hjorth, J. U., *Computer Intensive Statistical Methods: Validation Model Selection and Bootstrap*, Chapman & Hall/CRC, Boca Raton, FL., 1994.

Hoeting, J. A., Madigan, D., Raftery, A. E. and C. T. Volinsky, "Bayesian Model Averaging: A Tutorial," *Statistical Science*, 14, 382–401, 1999.

Huang, C. J. and H.-P. Lai, "Estimation of Stochastic Frontier Models Based on Multimodel Inference," *Journal of Productivity Analysis,* 38, 273–284, 2012.

Hultberg, P. T., Nadiri, M. I., and R. C. Sickles, "An International. Comparison of Technology. Adoption and Efficiency: A Dynamic. Panel Model," *Annales d'Economie et de Statistique,* 55-56, 449–474, 1999.

Hultberg, P. T., Nadiri, M. I. and R. C. Sickles, "Cross-country Catch-up in the Manufacturing Sector: Impacts of Heterogeneity On Convergence and Technology Adoption," *Empirical Economics*, 29, 753–768, 2004.

Hulten C. R. and A. Isaksson A., "Why Development Levels Differ: The Sources of Differential Economic Growth in a Panel of High and Low Income Countries," *National Bureau of Economic Research Working Paper No. 13469*, Cambridge, MA, 2007.

Isaksson, A., "World Productivity Database: A Technical Description," *Research and Statistics Staff Working Paper No. 10*, United Nations Industrial Development Organization, Vienna, 2007.

Jeon, B. M. and R. C. Sickles, "The Role of Environmental Factors in Growth Accounting," *Journal of Applied Econometrics,* 19, 567–591, 2004.

Jorgenson, D. and Z. Griliches, "Issues in Growth Accounting: A Reply to Edward F. Denison," *Survey of Current Business*, 52, 65-94, 1972.

Kalirajan, K. P., Obwona, M. B. and S. Zhao, "A Decomposition of Total Factor Productivity Growth: The Case of Chinese Agricultural Growth Before and After Reforms," *American Journal of Agricultural Economics,* 78, 331–338, 1996.

Kim, J.-I. and L. J. Lau, "The Sources of Economic Growth of the East Asian Newly Industrialized Countries," *Journal of the Japanese and International Economies*, 8, 235–271, 1994.

Kim, S. and Y. H. Lee, "The Productivity Debate of East Asia Revisited: A Stochastic Frontier Approach," *Applied Economics*, 38, 1697–1706, 2006.

Klein, L., *A Textbook of Econometrics*, Row, Peterson & Company, Evanston, IL, 1953.

Kneip, A., Sickles, R. and W.-H. Song, "A New Panel Data Treatment for Heterogeneity in Time Trends," *Econometric Theory*, 28, 590-628, 2012.

Koop, G., Piorier, D.J. and J. L. Tobias, *Bayesian Econometric Methods*, Cambridge University Press, New York, 2007.

Krugman, P., "The Myth of Asia's Miracle," *Foreign Affairs* Nov/Dec, 62–78, 1994.

Kumar S. and R. R. Russell, "Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence," *American Economic Review*, 92, 527-548, 2002.

Kumbhakar, S. C., "Production Frontiers, Panel Data, and Time-varying Technical Inefficiency," *Journal of Econometrics*, 46, 201–211, 1990.

Lahiri, K., Peng, H. and X. Sheng, "Measuring Aggregate Uncertainty in a Panel of Forecasts and a New Test for Forecast Heterogeneity," University of Albany-SUNY, 2015.

Lahiri, K., Peng, H. and Y. Zhao, "On-line Learning and Forecast Combination in Unbalanced Panels," *Econometric Reviews,* forthcoming, 2015.

Lahiri, K. and X. Sheng, "Measuring Forecast Uncertainty by Disagreement: The Missing Link," *Journal of Applied Econometrics*, 25, 514–538, 2010.

Lahiri, K., Teigland, C. and M. Zaporowski, "Interest Rates and the Subjective Probability Distribution of Inflation Forecasts," *Journal of Money, Credit and Banking*, 20, 233–248, 1988.

Lee, Y. H., "Tail Truncated Stochastic Frontier Models," *Journal of Economic Theory and Econometrics*, 2, 137–152, 1996.

Lee, Y. H. and S. Lee, "Stochastic Frontier Models with Threshold Efficiency," *Journal of Productivity Analysis*, 42, 45-54, 2012.

Lee, Y. H. and P. Schmidt, "A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency," in H. O. Fried, C. A. K. Lovell, and S. S. Schmidt, eds., *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press, New York, 237–255, 1993.

Leeb, H. and B. M. Pötscher, "Model Selection and Inference: Facts and Fiction," *Econometric Theory*, 21, 21–59, 2005

Liu, J., Sickles, R. C., and E. Tsionas, "Bayesian Treatment to Panel Data Models with Time-varying Heterogeneity," Rice University, Houston, 2015.

Lucas Jr, R. E., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3–42, 1988.

Mallows, C. L., "Some Comments on Cp," *Technometrics*, 15, 661–675, 1973.

Meeusen, W. and J. van Den Broeck, "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error," *International Economic Review*, 18, 435–444, 1977.

Moulin, H. "On Strategy-proofness and Single Peakedness," *Public Choice*, 35, 437–455, 1980.

Nelson, R. R. and G. Wright, "The Rise and Fall of American Technological Leadership: The Postwar Era in Historical Perspective," *Journal of Economic Literature*, 30, 1931–1964, 1992.

Newbold, P. and D. I. Harvey, "Forecast Combination and Encompassing," in M. P. Clements and D. F. Hendry, eds., *A Companion to Economic Forecasting,* John Wiley & Sons, New York, 268–283, 2002.

Norton M. I. and D. Ariely, "Building a Better America–One Wealth Quintile at a Time," *Perspectives on Psychological Science*, 6, 9-12, 2011.

Orea, L. and J. Steinbuks, "Estimating Market Power in Homogenous Product Markets Using a Composed Error Model: Application to the California Electricity Market," *Cambridge Working Paper in Economics 1220, EPRG Working Paper 1210*, University of Cambridge, 2012.

Park, B. U., Sickles, R. C. and L. Simar, "Stochastic Panel Frontiers: A Semiparametric Approach," *Journal of Econometrics*, 84, 273–301, 1998.

Park, B. U., Sickles, R. C. and L. Simar, "Semiparametric-efficient Estimation of AR (1) Panel Data Models," *Journal of Econometrics*, 117, 279–309, 2003.

Park, B. U., Sickles, R. C. and L. Simar, "Semiparametric-efficient Estimation of Dynamic Panel Data Models," *Journal of Econometrics*,136, 281–301, 2007.

Parmeter, C., Wan, A., and X. Zhang, "A Model Averaging Stochastic Frontier Estimator," University of Miami, Coral Gables, 2015.

Piketty, T., *Capital in the Twenty-First Century,* Harvard University Press, Cambridge, 2013.

Pitt, M. M. and L.-F. Lee, "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," *Journal of Development Economics*, 9, 43–64, 1981.

Raftery, A. E., Madigan, D. and J. A. Hoeting, "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association*, 92, 179–191, 1997.

Reikard, G., "Endogenous Technical Advance and the Stochastic Trend in Output: A Neoclassical Approach," *Research Policy*, 34, 1476–1490, 2005.

Romer, P. M., "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94, 1002–1037, 1986.

Sachs, J. D., Warner, A., Åslund, A. and S. Fischer, "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity*, 1, 1–118, 1995.

Schmidt, P. and R. C. Sickles, "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics*, 2, 367–374, 1984.

Schwarz, G., "Estimating the Dimension of A Model," *Annals of Statistics,* 6, 461–464, 1978.

Sickles, R. C., "Panel Estimators and the Identification of Firm-Specific Efficiency Levels in Parametric, Semiparametric and. Nonparametric Settings," *Journal of Econometrics,* 126, 305–324, 2005.

Sickles, R. C. and B. Cigerli, "Krugman and Young Revisited: A Survey of the Sources of Productivity Growth in a World with Less Constraints," *Seoul Journal of Economics*, 22, 29–54, 2009.

Sickles, R. C., Hao, J., and C. Shang, "Panel Data and Productivity Measurement," in Badi Baltagi, ed., *Oxford Handbook of Panel Data*, Oxford University Press, New York, 517-547, 2015.

Simar, L. and P. W. Wilson, "Statistical Inference in Nonparametric Frontier Models: The State of the Art," J*ournal of Productivity Analysis,* 13, 49–78, 2000.

Smolny, W., "Sources of Productivity Growth: An Empirical Analysis with German Sectoral Data," *Applied Economics*, 32, 305–314, 2000.

Solow, R. M., "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, 39, 312–320, 1957.

Stiglitz, J., *The Price of Inequality*, W.W. Norton & Company, New York, 2011.

Stiroh, K., "Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?," *Federal Reserve Bank of New York Staff Report No. 115*, 2001.

Timmermann, A., "Forecast Combinations," *Handbook of Economic Forecasting*,1, 135–196, 2006.

Tsionas, E. G., "Inference in Dynamic Stochastic Frontier Models," *Journal of Applied Econometrics*, 21, 669–676, 2006.

Tullock, G., "Efficient Rent Seeking," in J. M. Buchanan, R. D. Tollison, and G. Tullock, eds., *Toward a Theory of the Rent-Seeking Society*, Texas A&M University Press, College Station, 97-112, 1980.

United Nations Human Development Report, "Sustaining Human Progress: Reducing Vulnerabilities and Building Resilience," United Nations Development Programme, Washington, DC, 2014

Van den Broeck, J., Koop, G., Osiewalski, J. and M. F. Steel, "Stochastic Frontier Models: A Bayesian Perspective," *Journal of Econometrics*, 61, 273–303, 1994.

World Bank, *2004 World Bank Indicators*, International Bank for Reconstruction and Development/The World Bank, Washington, DC, 2004.

Young, A., "A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong and Singapore," in *NBER Macroeconomics Annual 1992, Volume 7*, MIT Press, Cambridge, 13–64.1992.

Young, A., "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *The Quarterly Journal of Economics*, 110, 641–680, 1995.

Zarnowitz, V. and L. A. Lambros, "Consensus and Uncertainty in Economic Prediction," in *Business Cycles: Theory, History, Indicators, and Forecasting,* National Bureau of Economic Research, Inc., Cambridge, MA, 1992.