Performance of Countries in the Post-Crisis Era

Paul W. Wilson Shirong Zhao *

Feb 2020

Preliminary Draft - Not for Citation

Abstract

This paper examines the performance of 144 countries in the world before, during, and after the 2007–2012 global financial crisis. Fully nonparametric methods are used to estimate technical efficiencies. Recently-developed statistical results are used to test for changes in efficiencies as well as productivity over time, and to test for changes in technology over time. We also test for these differences between developing and developed countries. We find evidence of the non-convexity of countries' production set. The data revealed that technical efficiency declined at the start of the global financial crisis (2006–2008), but recovered in the years later (2008–2014), ending higher in 2014 than in 2004. We also find that mean productivity continued decreasing from 2004 to 2010. Moreover, productivity in 2004 stochastic dominants in the first order that in 2014. Statistical tests indicate that the frontier continued shifting downward from 2004 to 2010, and then continued shifting upward from 2010 to 2014. Overall, the technology has shifted downward from 2004 to 2014. Finally, we provide evidence that developing economies have lower technical efficiency but higher productivity than developed economies over this period.

Keywords: technical efficiency, productivity, technology, non-parametric efficiency estimators

---

*Wilson: Department of Economics and School of Computing, Division of Computer Science, Clemson University, Clemson, South Carolina 29634–1309, USA; email pww@clemson.edu.

Zhao: Department of Economics, Clemson University, Clemson, South Carolina 29634–1309, USA; email shironz@g.clemson.edu.

We thank the Cyber Infrastructure Technology Integration group at Clemson University for operating the Palmetto cluster used for our computations. All remaining errors are our own. JEL classification nos.: O33 O47 O50 C12 C13 C14.

1 Introduction

The global financial crisis of 2007–2012 was the worst economic disaster since the 1929 great depression. It began with the subprime crisis of the housing mortgage markets in the U.S. Since securities linked to subprime loans were accumulated in all the banks and all the global financial market, the subprime crisis quickly caused inter-banking crisis. As early as on 14 September 2007, Britain's Northern Rock Bank received a liquidity support facility from the Bank of England, which led to the UK's first bank run in 150 years. In 2008, several other depressed financial institutions were purchased by others (Bear Stearns by JP Morgan Chase, Merrill Lynch by Bank of America), nationalized (Fannie Mae, Freddie Mac, and American International Group), or bankrupted (Washington Mutual, Lehman Brothers). Even though the governments of many countries provided liquidity and enacted large fiscal programs, bank failures led to a credit shortage and blocked the investment, plunging the global economy into a deep recession. Given this disruptive period 2004–2014, it is reasonable to ask what happened to the global economy following 2004, some years before the global financial crisis. Especially, we are interested in the evolution of productivity, efficiency, and technology of the countries during this period.

The traditional approach to the analysis of productivity using non-frontier analysis assumes that all the countries lie in the frontier and are perfectly efficient so that the growth of productivity is purely interpreted as the movement of the frontier or technology. However, the non-frontier analysis does not incorporate the inefficiency part resulting from the constraints or low efficient management, hence the estimation of productivity and technical progress would be biased. On the other hand, the frontier analysis could directly model the inefficiency behavior of the countries when estimating productivity, efficiency, and technical progress.

There exist some literature using parametric and nonparametric frontier methods to estimate technical efficiency and productivity of the countries in the world. Färe et al. (1994) may be the first to use nonparametric frontier method to analyze productivity growth in 17 OECD countries over the period 1979–1988. They decompose Malmquist productivity growth into changes in technical efficiency which measures catching-up effect, and shifts in technology which measures the innovation. They find that U.S. productivity growth is slightly higher than average. Ray and Desli (1997) propose an alternative decomposition of Malmquist productivity growth. Maudos et al. (1999) add human capital as another input to estimate Malmquist productivity and they find that the inclusion of human capital has a significant effect on the accurate measurement of total factor productivity. Chang and Luh (1999) use the same method as in Färe et al. (1994) to check for Asian countries and they find that the United States is not the sole innovator among the 19 APEC member economies. Rather, Hong Kong and Singapore have shown their capability to shift the grand frontier of the APEC economies. Arcelus and Arocena (2000) use nonparametric frontier method to estimate technical efficiency and scale efficiency for 14 OECD countries over 1970–1990 period. They find evidence of convergence, even if at quite different speeds, for total industry, manufacturing, and services. Emrouznejad (2003) proposes a dynamic nonparametric efficiency model to estimate efficiency, and compare the results with the static efficiency model proposed by Färe et al. (1994). They find that dynamic models capture efficiency better than static models. Han et al. (2004) use a varying coefficients frontier production (parametric method) to estimate total factor productivity, and then decompose it into technical efficiency change and technical change for a sample of 45 developed and developing countries over the period 1970–1990. They find that East Asian economies are not outliers in terms of total factor productivity growth. Salinas-Jiménez and Salinas-Jiménez (2007) use nonparametric frontier approach to estimate productivity, technical efficiency and then check the effects of corruption on these two measures for OECD countries. They find that corruption negatively influences the possibilities of growth. Wang et al. (2012) estimate Malmquist productivity index with and without defense expenditure, and then compare the difference of Malmquist productivity over time to check the effects of defense expenditure on economic productivity in OECD countries. They find that average Malmquist productivity with defense expenditure is higher than that without defense expenditure.

Among the nonparametric methods, Data Envelopment Analysis (DEA) estimators which impose convexity assumption of the production set, have been widely applied to estimate productivity and efficiency in these literature. However, they did not test the convexity of the production set, nor do they test constant returns to scale (CRS) versus variable returns to scale (VRS). Moreover, some of these studies simply report efficiency estimates without any inference and compare the mean efficiency of two groups without correcting the bias of the estimated efficiency. Of course, point estimates without inference are largely meaningless. Hence, the results of these studies are dubious. In addition, these studies only have a few observations and hence they naturally encounter the "curse of dimensionality", which is a serious problem in nonparametric efficiency estimation.[1] Hence, dimension reduction is needed in the context of nonparametric efficiency estimation. Recently, Kneip et al. (2016), using the central limit theorem results from Kneip et al. (2015), develop hypotheses testing the model structure. They provide tests of the convexity of the production set, returns to scale and differences in mean efficiency across groups of producers. Also, Wilson (2018) proposes a new dimension reduction technique that is advantageous in terms of reducing estimation error. Results suggest that Free Disposal Hull (FDH) estimator which does not impose convexity assumption is a viable, attractive alternative to the VRS-DEA in many cases when dimension reduction is used. We are the first to use recently-developed statistical methods to assess what can be learned about efficiency change, productivity growth and technical progress of the countries in the world from the data.

This paper provides evidence on the performance of the countries in the world before, during and after the 2007–2012 global financial crisis. The approach is fully non-parametric and exploits recently developed theoretical results. Estimates of technical efficiency and productivity at two-year intervals from 2004 to 2014 are examined in a statistical paradigm permitting inference and hypothesis testing. Therefore, this paper both (i) contributes to the literature by shedding light on the evolution of efficiency and productivity of the countries in the world before, during and after the 2007–2012 global financial crisis, and (ii) fills the gap between point estimates and inference in the empirical research on countries' technical efficiency and productivity.

The rest of paper is organized as follows. Estimators of technical efficiency and their properties are discussed in Section 2. Section 3 discusses various statistical results needed for testing hypotheses about model features. Section 4 describes the data we used. Section 5 discusses the empirical results of the tests. Major conclusions and directions for future works are presented in Section 6.

[1] Curse of dimensionality means the convergence rates of nonparametric estimators decrease with increasing dimensions (number of inputs and outputs).

2 The Statistical Model

To establish notation, let $X \in \mathbb{R}_+^p$ and $Y \in \mathbb{R}_+^q$ denote (random) vectors of input and output quantities, respectively. Similarly, let $x \in \mathbb{R}_+^p$ and $y \in \mathbb{R}_+^q$ denote fixed, nonstochastic vectors of input and output quantities. The production set

$$\Psi := \{(x, y) \mid x \text{ can produce } y\} \tag{2.1}$$

gives the set of feasible combinations of inputs and outputs. Several assumptions on Ψ are common in the literature. The assumptions of Shephard (1970) and Färe (1988) are typical in microeconomic theory of the firm and are used here.

Assumption 2.1. Ψ *is closed.*

Assumption 2.2. $(x, y) \notin \Psi$ *if* $x = 0,\ y \geq 0,\ y \neq 0$*; i.e., all production requires use of some inputs.*

Assumption 2.3. *Both inputs and outputs are strongly disposable, i.e.,* $\forall\ (x, y) \in \Psi$*, (i)* $\widetilde{x} \geq x \Rightarrow (\widetilde{x}, y) \in \Psi$ *and (ii)* $\widetilde{y} \leq y \Rightarrow (x, \widetilde{y}) \in \Psi$.

Here and throughout, inequalities involving vectors are defined on an element-by-element basis, as is standard. Assumption 2.1 ensures that the *efficient frontier* (or *technology*) Ψ^∂

$$\Psi^\partial := \left\{(x, y) \mid (x, y) \in \Psi,\ (\gamma^{-1}x, \gamma y) \notin \Psi \text{ for any } \alpha \in (1, \infty)\right\} \tag{2.2}$$

is the set of extreme points of Ψ and is contained in Ψ. Assumption 2.2 means that production of any output quantities greater than 0 requires use of some inputs so that there can be no free lunches. Assumption 2.3 imposes weak monotonicity on the frontier.

The Farrell (1957) input efficiency measure

$$\theta(x, y \mid \Psi) := \inf \{\theta \mid (\theta x, y) \in \Psi\} \tag{2.3}$$

gives the amount by which input levels can feasibly be scaled downward, proportionately by the same factor, without reducing output levels. The Farrell (1957) output efficiency measure gives the feasible, proportionate expansion of output quantities and is defined by

$$\lambda(x, y \mid \Psi) := \sup \{\lambda \mid (x, \lambda y) \in \Psi\}. \tag{2.4}$$

Both (2.3) and (2.4) provide *radial* measures of efficiency since all input or output quantities are scaled by the same factor θ or λ, holding output or input quantities fixed (respectively). Clearly, $\theta(x, y \mid \Psi) \leq 1$ and $\lambda(x, y \mid \Psi) \geq 1$ for all $(x, y) \in \Psi$.

Alternatively, Färe et al. (1985) provide a hyperbolic, graph measure of efficiency defined by

$$\gamma(x, y \mid \Psi) := \inf \left\{ \gamma > 0 \mid (\gamma x, \gamma^{-1} y) \in \Psi \right\}. \tag{2.5}$$

By construction, $\gamma(x, y \mid \Psi) \leq 1$ for $(x, y) \in \Psi$. Just as the measures $\theta(x, y \mid \Psi)$ and $\lambda(x, y \mid \Psi)$ provide measures of the *technical efficiency* of a firm operating at a point $(x, y) \in \Psi$, so does $\gamma(x, y \mid \Psi)$, but along the hyperbolic path from (x, y) to the frontier of Ψ. The measure $\gamma(x, y \mid \Psi)$ gives the amount by which input levels can be feasibly, proportionately scaled downward while simultaneously scaling output levels upward by the same proportion.

All of the quantities and model features defined so far are unobservable, and therefore must be estimated. The set Ψ can be estimated using the free-disposal hull (FDH) estimator introduced by Deprins et al. (1984) or either the variable returns to scale (VRS) or constant returns to scale (CRS) versions of the data envelopment analysis (DEA) estimator proposed by Farrell (1957). But, inference is needed in order to know what might be learned from data, and inference requires a well-defined statistical model.

3 Estimation and Inference

Let $\mathcal{S}_n = \{(X_i, Y_i)\}_{i=1}^n$ be a random input-output pairs sample drawn from the density f introduced in Assumption A.1. Given a random sample $\mathcal{S}_n = \{(X_i, Y_i)\}$, the production set Ψ can be estimated by the free disposal hull of the sample observations in $\mathcal{S}_n$,

$$\widehat{\Psi}_{\mathrm{FDH},n} := \bigcup_{(X_i, Y_i) \in \mathcal{S}_n} \left\{ (x, y) \in \mathbb{R}_+^{p+q} \mid x \geq X_i,\ y \leq Y_i \right\}, \tag{3.1}$$

proposed by Deprins et al. (1984). Alternatively, Ψ can be estimated by the convex hull of $\widehat{\Psi}_{\mathrm{FDH},n}$, i.e., by

$$\widehat{\Psi}_{\mathrm{VRS},n} := \left\{ (x, y) \in \mathbb{R}^{p+q} \mid y \leq \boldsymbol{Y}\boldsymbol{\omega},\ x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n \right\}, \tag{3.2}$$

where $\boldsymbol{X} = \left(X_1,\ \ldots,\ X_n\right)$ and $\boldsymbol{Y} = \left(Y_1,\ \ldots,\ Y_n\right)$ are $(p \times n)$ and $(q \times n)$ matrices of input and output vectors, respectively; $\boldsymbol{i}_n$ is an $(n \times 1)$ vector of ones, and $\boldsymbol{\omega}$ is a $(n \times 1)$ vector of weights. The estimator $\widehat{\Psi}_{\text{VRS},n}$ imposes convexity, but allows for VRS. This is the VRS (DEA) estimator of Ψ proposed by Farrell (1957) and popularized by Banker et al. (1984). The CRS (DEA) estimator $\widehat{\Psi}_{\text{CRS},n}$ of Ψ is obtained by dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (3.2). FDH, VRS or CRS estimators of $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ defined in Section 2 are obtained by substituting $\widehat{\Psi}_{\text{FDH},n}$, $\widehat{\Psi}_{\text{VRS},n}$ or $\widehat{\Psi}_{\text{CRS},n}$ for Ψ in (2.3)–(2.5) (respectively). In the case of VRS estimators, this results in

$$\widehat{\theta}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \min_{\theta,\boldsymbol{\omega}} \left\{ \theta \mid y \leq \boldsymbol{Y}\boldsymbol{\omega},\ \theta x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n \right\}, \tag{3.3}$$

$$\widehat{\lambda}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \max_{\lambda,\boldsymbol{\omega}} \left\{ \lambda \mid \lambda y \leq \boldsymbol{Y}\boldsymbol{\omega},\ x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n \right\} \tag{3.4}$$

and

$$\widehat{\gamma}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \min_{\gamma,\boldsymbol{\omega}} \left\{ \gamma \mid \gamma^{-1} y \leq \boldsymbol{Y}\boldsymbol{\omega},\ \gamma x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n \right\}. \tag{3.5}$$

The corresponding CRS estimators $\widehat{\theta}_{\text{CRS}}(x, y \mid \mathcal{S}_n)$, $\widehat{\lambda}_{\text{CRS}}(x, y \mid \mathcal{S}_n)$ and $\widehat{\gamma}_{\text{CRS}}(x, y \mid \mathcal{S}_n)$ are obtained by dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (3.3)–(3.5). The estimators in (3.3)–(3.4) can be computed using linear programming methods, but the hyperbolic estimator in (3.5) is a non-linear program. Nonetheless, estimates can be computed easily using the numerical algorithm developed by Wilson (2011). Substituting $\widehat{\Psi}_{\text{FDH},n}$ into (2.3)–(2.5) (respectively) will yield FDH estimators $\widehat{\theta}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$, $\widehat{\lambda}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$ and $\widehat{\gamma}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$. However, this leads to integer programming problems, but the estimators can be computed using simple numerical methods.[2]

The statistical properties of these efficiency estimators are well-developed. Kneip et al. (1998) derive the rate of convergence of the input-oriented VRS estimator, while Kneip et al. (2008) derive its limiting distribution. Park et al. (2010) derive the rate of convergence of the input-oriented CRS estimator and establish its limiting distribution. Park et al. (2000) and Daouia et al. (2017) derive both the rate of convergence and limiting distribution of the input-oriented FDH estimator. These results extend trivially to the output orientation after straightforward (but perhaps tedious) changes in notation. Wheelock and Wilson (2008) extend these results to the hyperbolic FDH estimator, and Wilson (2011) extends the results to the hyperbolic DEA estimator.

Kneip et al. (2015) derive moment properties of both the input-oriented FDH, VRS and CRS estimators and establish central limit theorem (CLT) results for mean input-oriented

[2]For details, see Kneip et al. (2015) and Wilson (2011).

efficiency after showing that the usual CLT results (e.g., the Lindeberg-Feller CLT) do not hold unless $(p+q) < 4$ in the CRS case, $(p+q) < 3$ in the VRS case, or unless $p+q < 2$ in the FDH case.[3] Kneip et al. (2016) use these CLT results to establish asymptotically normal test statistics for testing differences in mean efficiency across two groups, convexity versus non-convexity of Ψ, and CRS versus VRS (provided Ψ is weakly convex).[4] All of these results extend trivially (but again, tediously) to the output-oriented FDH, VRS and CRS estimators. These results could also be extended to the hyperbolic VRS and CRS estimators following Wilson (2011). The hyperbolic FDH estimator can be viewed as an input-oriented FDH estimator in a transformed space, hence moment results for the hyperbolic FDH estimator could also be extended trivially (but again, tediously) from the input-oriented FDH estimator. The new CLT results of Kneip et al. (2015) as well as the results from Kneip et al. (2016) on tests of differences in means, convexity versus non-convexity of Ψ, and CRS versus VRS carry over to the hyperbolic FDH estimator.

To summarize, in all cases, the FDH, VRS and CRS estimators are consistent, converge at rate n^{κ} (where $\kappa = 1/(p+q)$ for the FDH estimators, $\kappa = 2/(p+q+1)$ for the VRS estimators and $\kappa = 2/(p+q)$ for the CRS estimators) and possess non-degenerate limiting distributions under the appropriate set of assumptions. In addition, the bias of each of the three estimators is of order $O\left(n^{-\kappa}\right)$. Bootstrap methods proposed by Kneip et al. (2008, 2011) and Simar and Wilson (2011) provide consistent inference about $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ for a fixed point $(x, y) \in \Psi$, and Kneip et al. (2015) provide new CLT results enabling inference about the expected values of these measures over the random variables (X, Y).

Additional technical assumptions required for moment properties and central limit theorem results of means of FDH, VRS and CRS estimates, established by Kneip et al. (2015) and used below are given in Appendix A.

[3] In other words, standard CLT results hold in the FDH case if and only if $p = 1$ and output is fixed and constant, or $q = 1$ and input is fixed and constant.

[4] If Ψ^{∂} is globally CRS, then the VRS estimator attains the faster convergence rate of the CRS estimator due to the Theorem 3.1 of Kneip et al. (2016).

4 Data and Variable Specification

We calculate the efficiency and productivity for 144 countries over 2004–2014 using data from version 9.0 of the Penn World Table, PWT9.0 (Feenstra et al. (2015)).[5] Following Glass et al. (2016), the output Y is output-side real GDP at chained PPPs (in million 2011 US$, *rgdpo*). As recommended in the documentation, we use *rgdpo* to measure productive capacity across countries rather than expenditure-side real GDP (*rgdpe*) or real GDP at constant 2011 national prices (*rgdpna*). The first input X_1 is the labor input, measured by the number of persons engaged (in millions, *emp*). The second input X_2 is the capital stock at current PPPs (in million 2011 US$, *ck*). Maudos et al. (1999) mentioned that the inclusion of human capital has a significant effect on the accurate measurement of productivity. Therefore, the third input X_3 is human capital stock (*hc*), one index based on years of schooling and returns to education (see Human capital in PWT9.0). The input-output specification is standard (Färe et al. (1994)), reflecting the basic production process of countries. Table 1 shows the summary statistics for year 2014.

We assume that all countries operate in the same production set Ψ defined by (2.1), and therefore they face the same frontier in the four-dimensional input-output space. The model described in Section 2 is fully non-parametric, and hence quite flexible. The assumptions listed in Section 2 impose only minimal restrictions involving free-disposability, continuity, and some smoothness of the frontier, etc. Note that there is no assumption of convexity of Ψ, which is tested below in Section 5.

The flexibility of the non-parametric model specified in Section 2 comes with a price, however, in terms of the well-known "curse of dimensionality". The convergence rate of non-parametric efficiency estimators decreases with increasing inputs and outputs. The number of observations in each period that we consider is 144. With $p = 3$ and $q = 1$, the effective parametric sample size defined by Wilson (2018) is $144^{\frac{2}{4}} \approx 12$ for FDH estimators, $144^{\frac{4}{5}} \approx 53$ for VRS estimators and $144^{\frac{4}{4}} \approx 144$ for CRS estimators. With the sample size of 144 and the highest converge rate of $n^{\frac{2}{4}}$, non-parametric estimators should be expected to result in estimation error of order no better than that one would achieve with 144 observations in a typical parametric estimator. Given the relatively small sample size and the four dimensions,

[5]For more information about this data, see https://www.rug.nl/ggdc/productivity/pwt/.

it is not surprising that the estimated efficiency for many countries is equal to 1 and hence is not reliable.

To address this, the dimension reduction technique proposed by Wilson (2018) is applied. Considering the $(n \times p)$ and $(n \times q)$ matrices $\mathbf{X}$ and $\mathbf{Y}$ of observed non-negative inputs and outputs, we compute the $(n \times 1)$ vectors of principle components $X^* = \boldsymbol{X}\Lambda_x$ and $Y^* = \boldsymbol{Y}\Lambda_y$, where Λ_x and Λ_y are the $(p \times 1)$ and $(q \times 1)$ eigenvectors corresponding to the largest eigenvalues of $\mathbf{X}'\mathbf{X}$ and $\mathbf{Y}'\mathbf{Y}$, respectively. The dimensions of both $\mathbf{X}$ and $\mathbf{Y}$ are then reduced to only one dimension. However, we need to examine R_x and R_y, which are the ratios of the largest eigenvalue of the moment matrices $\mathbf{X}'\mathbf{X}$ and $\mathbf{Y}'\mathbf{Y}$ to the corresponding sums of the eigenvalues for these moment matrices. Wilson (2018) mentions that R_x and R_y provide measures of how close the corresponding moment matrices are to rank-one, regardless of the joint distributions of inputs and outputs.

The eigensystem analysis of input moment matrix is shown in Table 2. Since we only have one output, there is no need of dimension reduction for output $\mathbf{Y}$. The columns give the values of R_x from 2004 to 2014. The results are quite similar across years. An eigensystem analysis on the full data yields $R_x \geq 87.43$ percent for all years. It is clear that X^* contains most of the independent information of $\boldsymbol{X}$. Wilson (2018) shows that in many cases, but not in general, this dimension reduction method is advantageous in terms of reducing efficiency estimation error. In addition, dimension reduction could significantly increase the convergence rate of non-parametric efficiency estimators and lead to a more accurate estimation of efficiency. After dimension reduction applied, the convergence rates for FDH, VRS, and CRS are $n^{\frac{1}{2}}$, $n^{\frac{2}{3}}$ and n respectively. The tradeoff is that a small amount of information may be lost, but the mean squared error is reduced. All estimation in the following is done using X^* and Y.

5 Empirical Results

5.1 Efficiency and Productivity Evolution

As a robustness check to the need for dimension reduction, we estimate the hyperbolic efficiency for each year first using full data with four dimensions, and then using reduced data with only two dimensions. The FDH, VRS and CRS estimators are applied for both cases. Table 3 shows the number of observations with estimates equal to one. For each year, the FDH estimator produces more estimates equal to 1 than the VRS estimator, which produces more estimates equal to 1 than the CRS estimator. This is expected since there are more restrictions for the CRS estimator than the VRS estimator, which has more restrictions than the FDH estimator. More importantly, when using the full data, the FDH estimator results in 70.14 percent to 77.08 percent of the observations in a given year with estimates equal to one. The proportions for the VRS are between 15.28 percent and 17.36 percent and for the CRS are between 6.94 percent and 10.42 percent. This is clear evidence of too many dimensions for the given sample size. With dimension reduction, Table 3 shows that when using either estimator for any given year, the number of observations with estimates equal to 1 is much smaller than that without dimension reduction. In addition, the numbers using the FDH estimator are at least 3 times those using the VRS estimator, suggesting that the production set Ψ may be non-convex. In addition to large values of R_x discussed in Section 4, Table 3 provides another piece of evidence that dimension reduction likely reduces estimation error relative to what would be obtained when using the full data without dimension reduction. Therefore, the principal component for the three inputs X^* and the single output Y described in Section 4 are used for obtaining all the following results.

The next question is to determine which estimator we should use. As discussed in Section 2, in decreasing order of restrictions and rates of convergence lies the CRS, VRS, and FDH estimators. Using the test developed by Kneip et al. (2016), we use the principal component X^* and Y to test the null hypothesis of convexity of the production set Ψ versus the alternative hypothesis that Ψ is non-convex. Two randomly splitted subsamples for a given year are needed for this test. The first subsample of size $n_1 = \lfloor n/2 \rfloor$ is used for computing VRS estimates, and the second subsample of size $n_2 = n - n_1$ is used for computing FDH estimates. As Daraio et al. (2018) suggest, we first randomly shuffle the data using their randomization algorithm, and then take the first n_1 observations as the first subsample for computing VRS estimates, and the remaining n_2 observations as the second subsample for computing FDH estimates. The test statistic given in equation (50) of Kneip et al. (2016) involves the difference of the means of these two sets of estimates, with generalized jackknife estimates of biases and corresponding sample variances, and is asymptotically normally distributed with mean zero and unit variance. The test is a one-sided test since under the null the two means should be roughly similar, but should diverge with increasing departures from the null resulting in the mean of the FDH estimates exceeding the mean of the VRS estimates. The statistic given in equation (50) of Kneip et al. (2016) is defined in terms of input-oriented estimators but extends trivially to output-oriented and hyperbolic estimators. The tests are one-sided and we define the statistics so that "large" positive values indicate rejection of the null hypothesis.

The results of the convexity tests for each year are shown in Table 4. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.01. The results reveal that convexity is overwhelmingly rejected except the one in the input direction of 2004. Hence, the results in Table 4 provide strong evidence of the non-convexity of the production set Ψ. Also, even if the production set is convex, FDH estimator is still consistent. However, if the production set is non-convex, the VRS estimator is not consistent anymore. Consequently, the FDH estimators are applied for the remainder of the analysis. Our results cast doubts on the results of previous literature, which use DEA estimators to estimate the Malmquist productivity and technical efficiency.

Table 5 presents summary statistics of the FDH technical efficiency estimates in the input, output, and hyperbolic orientations. To compare with the input-oriented and hyperbolic-oriented estimates, we report the statistics of the reciprocals of the output-oriented estimates. For each orientation, the closer the estimates are to 1, the more technically efficient the countries. As might be expected, the hyperbolic estimates are more conservative on average, with mean efficiencies ranging from 0.7799 to 0.8252. By contrast, the means of the input-oriented estimates range from 0.6113 to 0.7139, while the means of the output-oriented estimates range from 0.6499 to 0.7146. These differences are due to the geometry of the efficiency measures as discussed by Wilson (2011). The mean efficiency in hyperbolic orientation increased from 2004 to 2010 and then decreased from 2010 to 2014. The pattern of mean efficiency in the input orientation appears to be the same, while the pattern in the output orientation is a little different. The mean efficiency in the output orientation increased from 2004 to 2008 and then decreased from 2008 to 2014.

We use the test described by Kneip et al. (2016, Section 3.1.1) to test for significant differences between the means reported in Table 5 from one year to the next, as well as from the first year to the last year. As discussed in Kneip et al. (2015, 2016), even with the reduced dimensionality so that $p + q = 2$, the usual CLT results (e.g., the Lindeberg-Feller CLT) do not hold for means of FDH efficiency estimates. As with the convexity test discussed above, the test statistic given by equation (18) of Kneip et al. (2016) involves not only the difference in sample means of efficiency estimates in a pair of years, but also the corresponding difference in generalized jackknife estimates of bias. The test extends trivially to the output-orientation and the hyperbolic orientation. In each case, the statistic used here is defined so that a positive value indicates that efficiency increases from year 1 to year 2, while a negative value indicates that efficiency decreases from year 1 to year 2.[6] As shown by Kneip et al. (2016), the test statistics are asymptotically normal with zero mean and unit variance. Since our data is balanced panel, there may exist time correlation, which violates the independent assumption of the test for differences of mean efficiency. The technical details dealing with time correlations is given in Appendix B Section B.1.

Table 6 gives the results of the tests of significant differences in mean efficiency over time. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.10. The result from 2004 to 2006 is mixed, where one case (input-oriented) shows that mean efficiency increased, while another case (output-oriented) shows that mean efficiency decreased. From 2006 to 2008, while there was no change of mean efficiency in the input and output orientation, mean efficiency declined significantly in the hyperbolic orientation. This possibly reflects the negative effect of the global financial crisis. The tests provide clear evidence that mean efficiency increased from 2008 to 2010 while there was no change of mean efficiency from 2010 to 2012. From 2012 to 2014, we see that mean efficiency started increasing again, showing that the global economy finally recovered from the global financial crisis. Overall, from 2004 to 2014, even though the statistic in the hyperbolic orientation is not significant, mean efficiency increased significantly in the other two orientations. Therefore, we find that the technical efficiency declined at the start of the global financial crisis (2006–2008), but recovered in the years later (2008–2014), ending higher in 2014 than in 2004.

In order to measure productivity, note that with the dimension reduction to $(p + q) = 2$ dimensions using the principal components X_i^*, Y_i as described in Section 4, productivity can be measured by Y_i/X_i^* for country i. Summary statistics for this measure is displayed

[6]Consequently, the statistic we use for the output orientation is the negative of the statistic appearing in equation (18) of Kneip et al. (2016).

in Table 7. Mean productivity continuously decreased from 2004 to 2010, then continuously increased from 2010 to 2014. The results show that the pattern of mean productivity is similar to the pattern of mean efficiency. Since productivity is measured by a simple ratio that does not involve estimators of efficiency, standard CLT results can be used to test for significant changes in means over time.[7] The results of these tests are shown in Table 8. Cells in column 7 are shaded whenever p-value is less than 0.01. It shows that mean productivity continued decreasing significantly from 2004 to 2010 and there was no change from 2010 to 2012 and from 2012 to 2014. Overall, from 2004 to 2014, the data reveals that there was a significant decrease in mean productivity.

To learn more about the difference of distributions of productivity in the two years interval, we use the stochastic dominance tests developed by Linton et al. (2005). Their method is based on subsampling and allows for the observations to be serially dependent. The outcome for the first-order stochastic dominance test is shown in Table 9. Cells in columns 3 and 5 are shaded whenever we could not reject the null hypothesis. Denote stochastic dominance in the first and second-order as SD1 and SD2, respectively. The p-value for the null hypothesis that productivity in 2004 SD1 that in 2006, is 0.999, which suggests that 2004 SD1 2006. The p-value for the null hypothesis that productivity in 2006 SD1 that in 2004, is close to 0, which suggests that 2006 does not SD1 2004 and hence the possibility that 2004 and 2006 have the same distribution is ruled out. Combining these two tests, we find that 2004 SD1 2006. Similarly, we also find that 2006 SD1 2008, 2008 SD1 2010 and 2004 SD1 2014. However, we do not find any SD1 over the periods 2010–2012 and 2012–2014. We also test whether there exists any SD2 in the two years interval. The outcome for the second-order stochastic dominance test is shown in Table 10. Cells in columns 3 and 5 are shaded whenever we could not reject the null hypothesis. Since SD1 implies SD2, it is not surprising that 2004 SD2 2006, 2006 SD2 2008, 2008 SD2 2010, and 2004 SD2 2014. Moreover, there did not exist SD2 over the periods 2010–2012 and 2012–2014.

The results presented so far provide clear evidence of changes in mean technical efficiency and productivity over the years represented in the sample. To gain further insight, we test whether the frontiers change over time. This involves the test of "separability" developed by Daraio et al. (2018), in which time is treated as a binary "environmental" variable. We

[7]However, we need to deal with time correlation. Please see Appendix B Section B.2 for technical details.

examine it using pairs of years 2004–2006, 2006–2008, ..., 2012–2014 as well as 2004–2014.

Implementation of the separability test of Daraio et al. (2018) involves pooling the data for two periods and then randomly shuffling the observations using the randomization algorithm presented by Daraio et al.. Then the pooled, randomly shuffled observations are split into two subsamples of equal size (or, if the combined number of observations is odd, one subsample will have one more observation than the other). Using the first subsample, efficiency is estimated as usual for each observation, ignoring which period a particular observation comes from, and the sample mean of the efficiency estimate is computed. The second subsample is split into the set of observations from period 1 and the set of observations from period 2. Efficiency is estimated for the period 1 observations using only the observations from period 1, while efficiency for the period 2 observations is estimated using only those observations from period 2. Then the sample mean of these two sets of efficiency estimates from the two sub-subsamples (of the second subsample) is computed. The resulting test statistic involves differences in the two subsample means as well as differences in the corresponding generalized jackknife estimates of bias. See Daraio et al. (2018) for discussion and details.

Results of the separability tests are shown in Table 11. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.10. In every case of periods 2004–2006, 2008–2010, 2010–2012 and 2012–2014, separability is rejected with p-value less than 0.01, and in most cases well less than 0.01. From 2006 to 2008, two statistics are significant at the 1 percent level, while the remaining statistic is not significant at the 10 percent level. Therefore the data provides moderate evidence that the technology changed from 2006 to 2008. Overall, from 2004 to 2014, two statistics are significant at the 10 percent level, while the remaining statistic is not significant at all. Therefore the evidence shows that the technology changed from 2004 to 2014.

In order to learn something about the *direction* in which technology may have shifted, we use new results from Simar and Wilson (2018) who provide CLT results for components of productivity changed measured by Malmquist indices. Simar and Wilson define the Malmquist index in terms of hyperbolic distances, and then consider various decompositions that can be used to identify components of productivity change. In particular, let Ψ^t represent the production set at time $t \in \{1,\ 2\}$ and let $Z_i^t = (X_i^t, Y_i^t)$ denote the i-th firm's observed input-output pair at time t. Then technical change relative to firm i's position at times 1 and 2 is measured by

$$\mathcal{T}_i = \left[\frac{\gamma(Z_i^2 \mid \Psi^1)}{\gamma(Z_i^2 \mid \Psi^2)} \times \frac{\gamma(Z_i^1 \mid \Psi^1)}{\gamma(Z_i^1 \mid \Psi^2)} \right]^{1/2}. \tag{5.1}$$

This is the hyperbolic analog of the output-oriented technical-change index that appears in the decompositions of Ray and Desli (1997), Gilbert and Wilson (1998), Simar and Wilson (1998) and Wheelock and Wilson (1999). The first ratio inside the brackets in (5.1) measures technical change relative to firm i's position at time 2, while the second ratio measures technical change relative to the firm's position at time 1. The measure $\mathcal{T}_i$ is the geometric mean of these two ratios. Values greater than 1 indicate an upward shift in the technology, while values less than 1 indicate a downward shift (a value of 1 indicates no change from time 1 to time 2).

Estimates $\widehat{\mathcal{T}}_i$ are obtained by substituting the hyperbolic FDH estimator for each term in (5.1). Simar and Wilson (2018) develop CLT results for geometric means $\widehat{T}^{1,2}$ of $\mathcal{T}_i$ over firms $i = 1, \ldots, n$, for periods 1 and 2, and these results can be used to test significant differences of the geometric means from 1. Table 12 shows the results of tests of technology change for each two-year interval as well as for the entire period 2004–2014. Cells in column 7 are shaded whenever p-value is less than 0.01. All the statistics are significant at the 1 percent level. The geometric mean $\widehat{T}^{1,2}$ is smaller than 1 for each two-year interval from 2004 to 2010, and greater than 1 for each two-year interval from 2010 to 2014. This suggests continuing downward shifts of the technology from 2004 to 2010 and continuing upward shifts of the technology from 2010 to 2014. Over the full period 2004–2014, $\widehat{T}^{1,2}$ is less than 1 significantly at the 1 percent level, suggesting that the technology shifted downward over this period.

5.2 Developing Versus Developed Countries

The convergence theory, also known as the catch-up effect, implies that developing countries will tend to grow at faster rates than developed countries. Therefore, developing countries should have higher productivity and efficiency. Our tests developed in Section 2 and 3 could be used to examine this hypothesis.

According to *the International Monetary Fund's World Economic Outlook Database, October 2018*, the following are considered as developed economies (or advanced economies):

United States, Japan, Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Portugal, Slovak Republic, Slovenia, Spain, Canada, United Kindom, Australia, Korea, Singapore, Czech Republic, Macao SAR, Sweden, Denmark, New Zealand, Switzerland, Hongkong SAR, Norway,Taiwan Province of China, Iceland, Puerto Rico, Israel, San Marino. However, our data do not cover Puerto Rico and San Marino. All the other remaining countries are considered as developing economies (or emerging markets and developing economies). Our sample covers 37 developed economies and 107 developing economies. Table 13 shows the summary statistics of developing and developed economies for the year 2014. As it is expected, developing economies averagely have lower GDP, capital and human capital but higher labor than developed economies. The annual growth rates of labor, capital, human capital, and GDP are shown in Table 14. During the global financial crisis (2008–2009), the total output of the world economy decreased by about 1.12 percent, which mainly caused by developed countries (decreased 4 percent). Over each period from 2004 to 2014, the developing countries almost always had higher annual growth rates of labor, capital, human capital, and GDP. Overall, from 2004 to 2014, developing economies doubled their GDP, while the GDP of developed economies only increased by about 41 percent.

Table 15 shows the results of tests of the difference in mean technical efficiency between developing and developed economies. Cells in columns 5, 7 and 9 are shaded whenever p-value is less than 0.01. All the statistics are negative and significant at the 1 percent significance level. This suggests that developing economies had lower technical efficiency than developed economies over each period covered by our data. Table 16 provides the results of tests of the difference in productivity between developing and developed economies. Cells in column 7 are shaded whenever p-value is less than 0.01. All the statistics are positive and significant at the 1 percent significance level. This result shows that developing economies had higher productivity than developed economies over each period. Taken together, we find that over 2004–2014, even though developing economies had lower technical efficiency, they had higher productivity than developed economies. This suggests that some developing economies have not fully adopted the current technologies. Our results confirms the convergence theory.

6 Summary and Conclusions

Among studies that use either FDH or DEA estimators to estimate efficiency and benchmark the performances of countries, the vast majority use VRS (DEA) estimators which impose convexity on the production set. The test of convexity versus non-convexity of Ψ developed by Kneip et al. (2016) allows researchers to let the data tell them whether DEA estimators are appropriate in a given setting. Here, in the context of countries, convexity is strongly rejected.

Because we reject convexity of the production set, we use FDH estimators which remain consistent when Ψ is not convex, whereas DEA estimators do not. We exploit collinearity in the data to reduce inputs to their first principle components, resulting in a two-dimensional problem. Results from Wilson (2018) indicate that this substantially reduces mean square error of efficiency estimates. Moreover, the simulation evidence provided by Wilson (2018) suggests that when production sets are convex, FDH estimates often have less mean square error than DEA estimators after dimension reduction.

By rigorously comparing estimates and testing differences across the years represented in our data, we find that technical efficiency of 144 countries in the world declined at the start of the global financial crisis (2006–2008) but recovered in the years later (2008–2014). Overall, there was an increase in mean efficiency from 2004 to 2014. The data revealed that productivity continued decreasing from 2004 to 2010. Overall, there was a significant decrease in mean productivity from 2004 to 2014. We also find that the frontier continued shifting downward from 2004 to 2010, and then continued shifting upward from 2010 to 2014. However, the technology had shifted downward from 2004 to 2014. Finally, the data revealed that developing economies had lower technical efficiency but higher productivity than developed economies over this period.

The 2008 global financial crisis indeed had an influential negative effect on efficiency, productivity, and technology of the global economy. Even though the global economy recovered in the years later, however, until 2014, 6 years after the crisis, our results show that the productivity and technology of the global economy had not fully recovered yet. Over this period, developing economies performed better than developed economies in terms of productivity, however, they need to improve their technical efficiency.

Appendix A Additional Assumptions

The assumptions that follow are similar to Assumptions 3.1–3.4 of Kneip et al. (2015) and complete the statistical model. The first two assumptions that follow are needed for both FDH and VRS estimators.

Assumption A.1. *(i) The random variables (X, Y) possess a joint density f with support $\mathcal{D} \subset \Psi$; and (ii) f is continuously differentiable on $\mathcal{D}$.*

Assumption A.2. *(i) $\mathcal{D}^* := \{(\theta(x, y \mid \Psi)x,\ y) \mid (x, y) \in \mathcal{D}\} = \{(x,\ \lambda(x, y \mid \Psi)y) \mid (x, y) \in \mathcal{D}\} = \{(\gamma(x, y \mid \Psi)x,\ \gamma(x, y \mid \Psi)^{-1}y) \mid (x, y) \in \mathcal{D}\} \subset \mathcal{D}$; (ii) $\mathcal{D}^*$ is compact; and (iii) $f(\theta(x, y \mid \Psi)x, y) > 0$ for all $(x, y) \in \mathcal{D}$.*

The next two assumptions are needed when VRS estimators are used. Assumption A.3 imposes some smoothness on the frontier. Kneip et al. (2008) required only two-times differentiability to establish the existence of a limiting distribution for VRS estimators, by the stronger assumption that follows is needed to establish results on moments of the VRS estimators.

Assumption A.3. *$\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ are three times continuously differentiable on $\mathcal{D}$.*

Recalling that the strong (i.e., free) disposability assumed in Assumption 2.3 implies that the frontier is weakly monotone, the next assumption strengthens this by requiring the frontier to be strictly monotone with no constant segments. This is also needed to establish properties of moments of the VRS estimators.

Assumption A.4. *$\mathcal{D}$ is* almost strictly convex*; i.e., for any $(x, y), (\widetilde{x}, \widetilde{y}) \in \mathcal{D}$ with $(\frac{x}{\|x\|}, y) \neq (\frac{\widetilde{x}}{\|\widetilde{x}\|}, \widetilde{y})$, the set $\{(x^*, y^*) \mid (x^*, y^*) = (x, y) + \alpha((\widetilde{x}, \widetilde{y}) - (x, y))$ for some $0 < \alpha < 1\}$ is a subset of the* interior *of $\mathcal{D}$.*

Alternatively, when FDH estimators are used, Assumptions A.3 and A.4 can be replaced by the following assumption.

Assumption A.5. *(i) $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ are twice continuously differentiable on $\mathcal{D}$; and (ii) all the first-order partial derivatives of $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ with respect to x and y are nonzero at any point $(x, y) \in \mathcal{D}$.*

Assumption A.5 strengthens strong disposability in the assumption 2.3 by requiring that the frontier is strictly monotone and does not possess constant segments (which might be the case, for example, if outputs are discrete as opposed to continuous, as in the case of ships produced by shipyards). Finally, part (i) of Assumption A.5 is weaker than Assumption A.3; here the frontier is required to be smooth, but not as smooth as required by Assumption A.3.[8] Assumptions 2.1–A.2 and Assumption A.5 comprise a statistical model appropriate for use of FDH estimators of technical efficiency, while Assumptions 2.1–A.4 comprise a statistical model appropriate for use of VRS estimators of technical efficiency.[9] These assumptions are sufficient for establishing consistency of the corresponding estimators. The stronger assumptions here are needed for results on moments and central limit theorems of the corresponding estimators.

[8]Assumption A.5 is slightly stronger, but much simpler than assumptions AII–AIII in Park et al. (2000).

[9]Additional assumptions are needed for CRS efficiency estimators. See Kneip et al. (2015) for additional discussion.

Appendix B Time Correlation in Testing Means

B.1 Correlation in Efficiency

Notice that there may exist time correlation when we testing the differences in mean efficiency over time, which violates the independent assumption of test for differences of mean efficiency in Kneip et al. (2016). Hence we take the following method to deal with time correlation.

let n_0 be the number of observations existing in both periods, n_1 be the number of observations existing in period 1 but not in period 2 and n_2 be the number of observations existing in period 2 but not in period 1. We then use the randomization algorithm in Daraio et al. (2018) to randomly shuffle these n_0 observations. For period 1, we combine the first half $n_{01} = \lfloor n_0/2 \rfloor$ of n_0 observations with these n_1 observations to form the sample, denoted as S_1. Similarly, for period 2, we combine the second half $n_{02} = n_0 - \lfloor n_0/2 \rfloor$ of n_0 observations with these n_2 observations to form the sample, denoted as S_2. By construction, S_1 and S_2 are independent, and now we can use the tests for differences in mean efficiency in Kneip et al. (2016).

B.2 Correlation in Productivity

let n_0 be the number of observations existing in both periods, n_1 be the number of observations existing in period 1 but not in period 2 and n_2 be the number of observations existing in period 2 but not in period 1. Productivity is calculated by the ratio of output over input.

Then to test H_0: $\mu_1 = \mu_2$, versus H_1: $\mu_1 \neq \mu_2$, we can use statistics

$$\hat{T} = \frac{\hat{\mu_2} - \hat{\mu_1}}{\frac{\hat{\sigma}_1^2}{(n_1+n_0)} + \frac{\hat{\sigma}_2^2}{(n_2+n_0)} + 2n_0 \frac{\hat{\sigma}_{12}}{(n_1+n_0)(n_2+n_0)}} \sim N(0,1) \tag{2.2}$$

Where $\hat{\mu_i}$, $i \in \{1,2\}$, is the sample mean for all the observations in period i, $\hat{\sigma_i^2}$ is the sample variance for all the observations in period i and $\hat{\sigma}_{12}$ is the sample covariance for all the n_0 observations exsting in both periods.

References

Arcelus, F. J. and P. Arocena (2000), Convergence and productive efficiency in fourteen OECD countries: A non-parametric frontier approach, *International Journal of Production Economics* , 13.

Banker, R. D., A. Charnes, and W. W. Cooper (1984), Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 30, 1078–1092.

Chang, C.-C. and Y.-H. Luh (1999), Efficiency change and growth in productivity: the Asian growth experience, *Journal of Asian Economics* 10, 551–570.

Daouia, A., L. Simar, and P. W. Wilson (2017), Measuring firm performance using nonparametric quantile-type distances, *Econometric Reviews* 36, 156–181.

Daraio, C., L. Simar, and P. W. Wilson (2018), Central limit theorems for conditional efficiency measures and tests of the 'separability condition' in non-parametric, two-stage models of production, *Econometrics Journal* 21, 170–191.

Deprins, D., L. Simar, and H. Tulkens (1984), Measuring labor inefficiency in post offices, in M. M. P. Pestieau and H. Tulkens, eds., *The Performance of Public Enterprises: Concepts and Measurements*, Amsterdam: North-Holland, pp. 243–267.

Emrouznejad, A. (2003), An alternative DEA measure: a case of OECD countries, *Applied Economics Letters* 10, 779–782.

Färe, R. (1988), *Fundamentals of Production Theory*, Berlin: Springer-Verlag.

Färe, R., S. Grosskopf, and C. A. K. Lovell (1985), *The Measurement of Efficiency of Production*, Boston: Kluwer-Nijhoff Publishing.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang (1994), Productivity growth, technical progress, and efficiency change in industrialized countries, *American Economic Review* 84, 66–83.

Farrell, M. J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society A* 120, 253–281.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015), The Next Generation of the Penn World Table, *American Economic Review* 105, 3150–3182.

Gilbert, A. and P. W. Wilson (1998), Effects of deregulation on the productivity of korean banks, *Journal of Economics and Business* 50, 133–155.

Glass, A. J., K. Kenjegalieva, and R. C. Sickles (2016), A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers, *Journal of Econometrics* 190, 289–300.

Han, G., K. Kalirajan, and N. Singh (2004), Productivity, Efficiency and Economic Growth: East Asia and the Rest of the World, *The Journal of Developing Areas* 37, 99–118.

Kneip, A., B. Park, and L. Simar (1998), A note on the convergence of nonparametric DEA efficiency measures, *Econometric Theory* 14, 783–793.

Kneip, A., L. Simar, and P. W. Wilson (2008), Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models, *Econometric Theory* 24, 1663–1697.

— (2011), A computationally efficient, consistent bootstrap for inference with non-parametric DEA estimators, *Computational Economics* 38, 483–515.

— (2015), When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores, *Econometric Theory* 31, 394–422.

— (2016), Testing hypotheses in nonparametric models of production, *Journal of Business and Economic Statistics* 34, 435–456.

Linton, O., E. Maasoumi, and Y.-J. Whang (2005), Consistent Testing for Stochastic Dominance under General Sampling Schemes, *Review of Economic Studies* 72, 735–765.

Maudos, J., J. M. Pastor, and L. Serrano (1999), Total factor productivity measurement and human capital in OECD countries, *Economics Letters* 63, 39–44.

Park, B. U., S.-O. Jeong, and L. Simar (2010), Asymptotic distribution of conical-hull estimators of directional edges, *Annals of Statistics* 38, 1320–1340.

Park, B. U., L. Simar, and C. Weiner (2000), FDH efficiency scores from a stochastic point of view, *Econometric Theory* 16, 855–877.

Ray, S. C. and E. Desli (1997), Productivity growth, technical progress, and efficiency change in industrialized countries: Comment, *American Economic Review* 87, 1033–1039.

Salinas-Jiménez, M. d. M. and J. Salinas-Jiménez (2007), Corruption, efficiency and productivity in OECD countries, *Journal of Policy Modeling* 29, 903–915.

Shephard, R. W. (1970), *Theory of Cost and Production Functions*, Princeton: Princeton University Press.

Simar, L. and P. W. Wilson (1998), Productivity growth in industrialized countries. Discussion Paper #9810, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

— (2011), Inference by the m out of n bootstrap in nonparametric frontier models, *Journal of Productivity Analysis* 36, 33–53.

— (2018), Central limit theorems and inference for sources of productivity change measured by nonparametric Malmquist indices. Discussion Paper 2018/21, Institut de Statistique Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Wang, T.-P., S. H.-P. Shyu, and H.-C. Chou (2012), The impact of defense expenditure on economic productivity in OECD countries, *Economic Modelling* 29, 2104–2114.

Wheelock, D. C. and P. W. Wilson (1999), Technical progress, inefficiency, and productivity change in U. S. banking, 1984–1993, *Journal of Money, Credit, and Banking* 31, 212–234.

— (2008), Non-parametric, unconditional quantile estimation for efficiency analysis with an application to Federal Reserve check processing operations, *Journal of Econometrics* 145, 209–225.

Wilson, P. W. (2011), Asymptotic properties of some non-parametric hyperbolic efficiency estimators, in I. Van Keilegom and P. W. Wilson, eds., *Exploring Research Frontiers in Contemporary Statistics and Econometrics*, Berlin: Springer-Verlag, pp. 115–150.

— (2018), Dimension reduction in nonparametric models of production, *European Journal of Operational Research* 267, 349–367.

Table 1: Summary Statistics for Year 2014

Variable	Min	Q1	Median	Mean	Q3	Max
GDP (Y)	$2.7830 \times 10^{+03}$	$3.2340 \times 10^{+04}$	$1.1820 \times 10^{+05}$	$7.1440 \times 10^{+05}$	$4.6660 \times 10^{+05}$	$1.7140 \times 10^{+07}$
Labor (L)	1.2370×10^{-01}	$1.8400 \times 10^{+00}$	$4.9120 \times 10^{+00}$	$2.1850 \times 10^{+01}$	$1.4460 \times 10^{+01}$	$7.9840 \times 10^{+02}$
Capital (K)	$5.1750 \times 10^{+03}$	$1.1140 \times 10^{+05}$	$3.9030 \times 10^{+05}$	$2.6430 \times 10^{+06}$	$1.8000 \times 10^{+06}$	$6.9380 \times 10^{+07}$
Human Capital (H)	$1.1930 \times 10^{+00}$	$2.0180 \times 10^{+00}$	$2.6620 \times 10^{+00}$	$2.5950 \times 10^{+00}$	$3.1560 \times 10^{+00}$	$3.7340 \times 10^{+00}$

Note: GDP is output-side real GDP at chained PPPs (in million 2011 US$); Labor is number of persons engaged (in millions); Capital is capital stock at current PPPs (in million 2011 US$); Human Capital is one index based on years of schooling and returns to education.

Table 2: Eigensystem Analysis by Year

Year	$R_x(\%)$
2004	87.43
2006	87.70
2008	88.02
2010	88.24
2012	88.53
2014	88.76

Table 3: Numbers of Observations With Estimated Hyperbolic Technical Efficiency Equal to 1 in Each Year

		Without — Dimension Reduction —			With — Dimension Reduction —		
Year	n	FDH	VRS	CRS	FDH	VRS	CRS
2004	144	105	24	15	27	7	1
2006	144	105	23	11	29	8	1
2008	144	105	22	13	32	8	1
2010	144	111	25	13	36	6	1
2012	144	101	22	10	29	6	1
2014	144	109	25	11	27	6	1

Table 4: Results of Convexity Tests (with Dimension Reduction, $p = q = 1$)

	— Input —		**— Output —**		**— Hyperbolic —**	
Year	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004	1.4720	7.05×10^{-02}	5.0490	2.22×10^{-07}	3.7056	1.05×10^{-04}
2006	6.2603	1.92×10^{-10}	9.0609	6.47×10^{-20}	7.0398	9.62×10^{-13}
2008	3.5480	1.94×10^{-04}	4.8906	5.03×10^{-07}	3.5020	2.31×10^{-04}
2010	8.4456	1.51×10^{-17}	9.6505	2.45×10^{-22}	9.1008	4.48×10^{-20}
2012	7.5268	2.60×10^{-14}	8.2257	9.71×10^{-17}	8.8921	3.00×10^{-19}
2014	3.0405	1.18×10^{-03}	3.1071	9.45×10^{-04}	3.8103	6.94×10^{-05}

NOTE: The numerator of statistics is the difference of estimated mean VRS estimates minus mean FDH estimates.

Table 5: Summary Statistics for FDH Technical Efficiency Estimates (with Dimension Reduction, $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max
			— Input Orientation —			
2004	0.1139	0.3785	0.6303	0.6113	0.8899	1.0000
2006	0.1972	0.4616	0.6438	0.6609	0.9033	1.0000
2008	0.2097	0.5135	0.6727	0.6975	0.9557	1.0000
2010	0.2084	0.5239	0.6984	0.7139	0.9953	1.0000
2012	0.2150	0.5038	0.6624	0.6742	0.9218	1.0000
2014	0.1858	0.4150	0.5896	0.6338	0.8684	1.0000
			— Output Orientation —			
2004	0.1284	0.4391	0.6477	0.6499	0.9217	1.0000
2006	0.1245	0.5108	0.6970	0.6917	0.9323	1.0000
2008	0.1705	0.5104	0.7415	0.7146	0.9623	1.0000
2010	0.2009	0.5088	0.6706	0.7014	0.9815	1.0000
2012	0.2255	0.4649	0.6501	0.6802	0.9297	1.0000
2014	0.1859	0.4158	0.6589	0.6500	0.8961	1.0000
			— Hyperbolic Orientation —			
2004	0.3516	0.6371	0.7840	0.7799	0.9437	1.0000
2006	0.4053	0.6863	0.8300	0.8079	0.9470	1.0000
2008	0.4781	0.6930	0.8312	0.8211	0.9819	1.0000
2010	0.4919	0.6914	0.8278	0.8252	0.9953	1.0000
2012	0.4539	0.6633	0.8129	0.8051	0.9670	1.0000
2014	0.4082	0.6783	0.7840	0.7923	0.9403	1.0000

NOTE: Statistics for the reciprocals of the output efficiency estimates are given to facilitate comparison with the input-oriented and hyperbolic estimates.

Table 6: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Time (with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004–2006	2.0874	3.69×10^{-02}	-2.6490	8.07×10^{-03}	-0.6728	5.01×10^{-01}
2006–2008	0.1634	8.70×10^{-01}	0.6564	5.12×10^{-01}	-2.5097	1.21×10^{-02}
2008–2010	2.2908	2.20×10^{-02}	1.5509	1.21×10^{-01}	2.3430	1.91×10^{-02}
2010–2012	-0.4524	6.51×10^{-01}	-0.7692	4.42×10^{-01}	-0.1788	8.58×10^{-01}
2012–2014	2.8697	4.11×10^{-03}	3.5162	4.38×10^{-04}	3.9343	8.34×10^{-05}
2004–2014	1.8283	6.75×10^{-02}	2.4703	1.35×10^{-02}	1.3329	1.83×10^{-01}

NOTE: The numerator of statistics for each period is the difference of estimated mean efficiency of the second year minus the first year.

Table 7: Summary Statistics for Productivity (with Dimension Reduction, $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max
2004	1.5300	2.7583	3.3940	4.0248	4.5488	18.3875
2006	1.4066	2.5806	3.0827	3.6276	4.2816	13.5363
2008	1.3057	2.3642	2.9638	3.3147	3.7893	9.4944
2010	1.2835	2.2572	2.7112	3.0851	3.7210	8.7854
2012	1.2213	2.1914	2.8185	3.1122	3.7039	7.9003
2014	1.2955	2.2069	2.7949	3.1541	3.7377	11.2910

NOTE: Productivity for observation i is defined as Y_i/X_i^*.

Table 8: Tests of Differences in Means for Productivity Estimates with Respect to Time (with Dimension Reduction, $p = q = 1$)

Period	n_1	n_2	Mean1	Mean2	Statistic	p-value
2004–2006	144	144	4.0248	3.6276	-5.2576	1.46×10^{-07}
2006–2008	144	144	3.6276	3.3147	-5.0137	5.34×10^{-07}
2008–2010	144	144	3.3147	3.0851	-4.6205	3.83×10^{-06}
2010–2012	144	144	3.0851	3.1122	0.5699	5.69×10^{-01}
2012–2014	144	144	3.1122	3.1541	1.0478	2.95×10^{-01}
2004–2014	144	144	4.0248	3.1541	-4.6766	2.92×10^{-06}

NOTE: The numerator of statistics for each period is the difference of estimated mean productivity of the second year minus the first year.

Table 9: First Order Stochastic Dominance Test for Productivity with Respect to Time

	— Year 1 SD1 Year 2 —		— Year 2 SD1 Year 1 —	
Period	Statistic	p-value	Statistic	p-value
2004–2006	0.0000	9.99×10^{-01}	0.1250	9.99×10^{-04}
2006–2008	-0.0069	9.99×10^{-01}	0.0972	6.99×10^{-03}
2008–2010	0.0000	9.99×10^{-01}	0.1250	$0.00 \times 10^{+00}$
2010–2012	0.0625	9.19×10^{-02}	0.0417	5.54×10^{-01}
2012–2014	0.0417	2.76×10^{-01}	0.0278	7.60×10^{-01}
2004–2014	-0.0208	9.99×10^{-01}	0.2361	$0.00 \times 10^{+00}$

NOTE: The null hypothesis is that there exists first order stochastic dominance.

Table 10: Second Order Stochastic Dominance Test for Productivity with Respect to Time

	— **Year 1 SD2 Year 2** —		— **Year 2 SD2 Year 1** —	
Period	Statistic	p-value	Statistic	p-value
2004–2006	-0.0013	9.84×10^{-01}	0.3105	$0.00 \times 10^{+00}$
2006–2008	-0.0010	9.92×10^{-01}	0.2490	$0.00 \times 10^{+00}$
2008–2010	-0.0002	9.80×10^{-01}	0.1937	$0.00 \times 10^{+00}$
2010–2012	0.0229	4.10×10^{-01}	0.0195	4.21×10^{-01}
2012–2014	0.0130	4.35×10^{-01}	0.0017	7.77×10^{-01}
2004–2014	-0.0018	9.86×10^{-01}	0.7166	$0.00 \times 10^{+00}$

NOTE: The null hypothesis is that there exists second order stochastic dominance.

Table 11: Tests for Separability with Respect to Time (with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004–2006	5.0232	2.54×10^{-07}	2.4078	8.03×10^{-03}	5.4435	2.61×10^{-08}
2006–2008	1.2211	1.11×10^{-01}	5.5903	1.13×10^{-08}	2.9588	1.54×10^{-03}
2008–2010	3.3544	3.98×10^{-04}	4.8669	5.67×10^{-07}	3.7419	9.13×10^{-05}
2010–2012	4.3189	7.84×10^{-06}	5.9273	1.54×10^{-09}	3.1343	8.61×10^{-04}
2012–2014	2.8978	1.88×10^{-03}	4.2091	1.28×10^{-05}	2.7058	3.41×10^{-03}
2004–2014	1.4604	7.21×10^{-02}	−2.0830	9.81×10^{-01}	1.3939	8.17×10^{-02}

NOTE: The numerator of the statistics is the difference of the conditional mean estimates minus the unconditional mean estimates.

Table 12: Tests for Technology Change with Respect to Time (with Dimension Reduction, $p = q = 1$)

Period	n_1	n_2	n	$\widehat{T}^{1,2}$	Var	p-value
2004–2006	144	144	144	0.9348	0.0247	1.69×10^{-07}
2006–2008	144	144	144	0.9508	0.0148	1.49×10^{-09}
2008–2010	144	144	144	0.9670	0.0158	8.99×10^{-03}
2010–2012	144	144	144	1.0349	0.0174	9.18×10^{-11}
2012–2014	144	144	144	1.0283	0.0188	1.07×10^{-06}
2004–2014	144	144	144	0.9162	0.0216	4.78×10^{-10}

NOTE: For each period, the number of observations in the first year is n_1, while the number of observations in the second year is n_2. The number of observations existing in both years is n. Mean of the technology ratio $\widehat{T}^{1,2}$ is greater than 1 if and only if the technology shifts upward.

Table 13: Summary Statistics of Developing and Developed Economies for Year 2014

Variable	Min	Q1	Median	Mean	Q3	Max
			— Developing Economies —			
GDP (Y)	$2.7830 \times 10^{+03}$	$2.7320 \times 10^{+04}$	$7.7750 \times 10^{+04}$	$5.4090 \times 10^{+05}$	$4.0300 \times 10^{+05}$	$1.7140 \times 10^{+07}$
Labor (L)	1.2370×10^{-01}	$1.8390 \times 10^{+00}$	$5.5490 \times 10^{+00}$	$2.4780 \times 10^{+01}$	$1.4550 \times 10^{+01}$	$7.9840 \times 10^{+02}$
Capital (K)	$5.1750 \times 10^{+03}$	$8.5640 \times 10^{+04}$	$2.2450 \times 10^{+05}$	$1.8760 \times 10^{+06}$	$1.2200 \times 10^{+06}$	$6.9380 \times 10^{+07}$
Human Capital (H)	$1.1930 \times 10^{+00}$	$1.8250 \times 10^{+00}$	$2.4430 \times 10^{+00}$	$2.3420 \times 10^{+00}$	$2.7800 \times 10^{+00}$	$3.4110 \times 10^{+00}$
			— Developed Economies —			
GDP (Y)	$1.0030 \times 10^{+04}$	$1.3560 \times 10^{+05}$	$3.2620 \times 10^{+05}$	$1.2160 \times 10^{+06}$	$1.0450 \times 10^{+06}$	$1.6600 \times 10^{+07}$
Labor (L)	1.8330×10^{-01}	$1.8520 \times 10^{+00}$	$3.9880 \times 10^{+00}$	$1.3380 \times 10^{+01}$	$1.1970 \times 10^{+01}$	$1.4850 \times 10^{+02}$
Capital (K)	$4.0060 \times 10^{+04}$	$4.5960 \times 10^{+05}$	$1.7330 \times 10^{+06}$	$4.8630 \times 10^{+06}$	$3.9450 \times 10^{+06}$	$5.2850 \times 10^{+07}$
Human Capital (H)	$2.4270 \times 10^{+00}$	$3.1180 \times 10^{+00}$	$3.3640 \times 10^{+00}$	$3.3270 \times 10^{+00}$	$3.5940 \times 10^{+00}$	$3.7340 \times 10^{+00}$

Note: GDP is output-side real GDP at chained PPPs (in million 2011 US$); Labor is number of persons engaged (in millions); Capital is capital stock at current PPPs (in million 2011 US$); Human Capital is one index based on years of schooling and returns to education.

Table 14: Annual Growth Rate of Labor, Capital, Human Capital and GDP over 2004-2014

Period	L	K	H	Y
— All Economies —				
2004–2005	0.0289	0.1354	0.0082	0.1053
2005–2006	0.0296	0.1300	0.0082	0.0744
2006–2007	0.0315	0.1063	0.0082	0.0737
2007–2008	0.0250	0.0987	0.0081	0.0753
2008–2009	0.0112	0.0661	0.0080	-0.0112
2009–2010	0.0181	0.0877	0.0078	0.0944
2010–2011	0.0213	0.1214	0.0091	0.0939
2011–2012	0.0179	0.0535	0.0090	0.0462
2012–2013	0.0179	0.0594	0.0089	0.0360
2013–2014	0.0182	0.0608	0.0089	0.0342
2004–2014	0.2639	1.5658	0.0897	0.8618
— Developed Economies —				
2004–2005	0.0178	0.0975	0.0061	0.0752
2005–2006	0.0252	0.1338	0.0063	0.0303
2006–2007	0.0273	0.1069	0.0063	0.0638
2007–2008	0.0175	0.0816	0.0063	0.0352
2008–2009	-0.0180	0.0328	0.0063	-0.0400
2009–2010	-0.0014	0.0458	0.0062	0.0498
2010–2011	0.0111	0.0600	0.0062	0.0416
2011–2012	0.0055	0.0502	0.0063	0.0283
2012–2013	0.0049	0.0429	0.0063	0.0232
2013–2014	0.0112	0.0319	0.0064	0.0278
2004–2014	0.1154	0.9737	0.0656	0.4145
— Developing Economies —				
2004–2005	0.0327	0.1485	0.0090	0.1157
2005–2006	0.0312	0.1287	0.0089	0.0896
2006–2007	0.0330	0.1061	0.0088	0.0771
2007–2008	0.0276	0.1046	0.0088	0.0892
2008–2009	0.0213	0.0776	0.0086	-0.0013
2009–2010	0.0248	0.1022	0.0084	0.1098
2010–2011	0.0248	0.1427	0.0101	0.1120
2011–2012	0.0222	0.0547	0.0099	0.0524
2012–2013	0.0225	0.0652	0.0099	0.0404
2013–2014	0.0206	0.0707	0.0098	0.0364
2004–2014	0.3153	1.7706	0.0980	1.0165

Note: GDP is output-side real GDP at chained PPPs (in million 2011 US$); Labor is number of persons engaged (in millions); Capital is capital stock at current PPPs (in million 2011 US$); Human Capital is one index based on years of schooling and returns to education. The growth rate over 2004–2014 is the accumulated annual growth rate.

Table 15: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Type (with Dimension Reduction, $p = q = 1$)

			— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	n_1	n_2	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004	37	107	-8.1521	3.58×10^{-16}	-11.4682	1.91×10^{-30}	-11.7834	4.75×10^{-32}
2006	37	107	-7.3594	1.85×10^{-13}	-8.2827	1.20×10^{-16}	-8.8034	1.33×10^{-18}
2008	37	107	-4.3511	1.35×10^{-05}	-8.0633	7.43×10^{-16}	-4.9163	8.82×10^{-07}
2010	37	107	-4.5157	6.31×10^{-06}	-7.1762	7.17×10^{-13}	-4.3040	1.68×10^{-05}
2012	37	107	-3.7759	1.59×10^{-04}	-5.0038	5.62×10^{-07}	-4.0607	4.89×10^{-05}
2014	37	107	-8.1837	2.75×10^{-16}	-10.3579	3.85×10^{-25}	-7.5720	3.67×10^{-14}

NOTE: The number of developed economies is n_1, while the number of developing economies is n_2. The numerator of statistics is the difference of estimated mean efficiency of developing economies minus developed economies.

Table 16: Tests of Differences in Means for Productivity Estimates with Respect to Type (with Dimension Reduction, $p = q = 1$)

Year	n_1	n_2	Mean1	Mean2	Statistic	p-value
2004	37	107	2.9420	4.3993	5.1180	3.09×10^{-07}
2006	37	107	2.5810	3.9895	6.3903	1.66×10^{-10}
2008	37	107	2.3264	3.6565	7.8422	4.43×10^{-15}
2010	37	107	2.1908	3.3944	7.5544	4.21×10^{-14}
2012	37	107	2.2020	3.4269	6.8688	6.48×10^{-12}
2014	37	107	2.2701	3.4598	6.3503	2.15×10^{-10}

NOTE: The number of developed economies is n_1, while the number of developing economies is n_2. The numerator of statistics is the difference of estimated mean productivity of developing economies minus developed economies.