Conditional time-dependent nonparametric estimators with an application to healthcare production function

Kourtzidis, Stavros; Tzeremes, Panayiotis; Tzeremes, Nickolaos

Published in:
Journal of Applied Statistics

DOI:
10.1080/02664763.2019.1588234

Publication date:
2019

Document Version
Peer reviewed version

Author Accepted Manuscript of Kourtzidis, S, Tzeremes, P & Tzeremes, N 2019, 'Conditional time-dependent nonparametric estimators with an application to healthcare production function', Journal of Applied Statistics.10.1080/02664763.2019.1588234

Conditional time-dependent nonparametric estimators with an application to healthcare production function

Stavros Kourtzidis[a], Panayiotis Tzeremes[b], Nickolaos G. Tzeremes[c*]

[a]*University of Dundee, School of Business, Perth Road 1-3, DD1 4HN, Dundee, United Kingdom.*

[b,c]*Laboratory of Economic Policy and Strategic Planning, Department of Economics, University of Thessaly, 28th October street, 78, 38333, Volos, Greece.*

**Corresponding author's email: bus9nt@econ.uth.gr*

Abstract

By using the probabilistic framework of production efficiency, the paper develops time-dependent conditional efficiency estimators performing a non-parametric frontier analysis. Specifically, by applying both full and quantile (robust) time-dependent conditional estimators, it models the dynamic effect of health expenditure on countries' technological change and technological catch-up levels. The results from the application reveal that the effect of per capita health expenditure on countries' technological change and technological catch-up is nonlinear and is subject to countries' specific income levels.

Keywords: Conditional efficiency measures; Health expenditure; Non-parametric analysis; Probabilistic approach.

1. Introduction

Public and private investment affect the mechanism of social cohesion, which interrelates with countries' path of economic development [13]. Such investment policies are favoured by countries since they are the principal transmission channel through which countries' fiscal and financial development policies affect their productivity levels [7, 17]. Specifically, the relative literature asserts that healthcare expenditure influences countries' technological change levels [19, 20] . Similarly, Baldacci et al. [5] provide evidence that health expenditure has a positive and significant impact on economic growth through the enhancement of health capital. Bloom and Canning [6] demonstrate the positive effect of health capital on aggregated output. From the other hand Aisa and Pueyo [1] provide a model capturing the non-monotonic effect of government health spending on economic growth. Whereas, from the early years Azariadis and Drazen [2] highlight the existence of such threshold phenomena in growth paths which are more pronounced in the early stages of growth.

This paper in a fully nonparametric framework contributes to the relative literature by providing an innovative application of the probabilistic framework of efficiency measurement [8,10], alongside with its latest methodological developments [3,18] on modelling the dynamic effect of countries' healthcare expenditure on their estimated production efficiency levels. According to Daraio et al. [11] the adoption of such probabilistic efficiency measures gives the researcher the ability to investigate the influence of the examined factors not only on the shape and the level of the boundary of the attainable set, but also on the distribution of the estimated inefficiencies. This attribute suggests that the adopted models do not assume that the "separability" assumption holds. If the "separability" assumption is not verified or wrongly (ad hoc) imposed, then the it is assumed that time and health expenditure influences only countries' distribution of the estimated inefficiencies [11].

The efficiency estimators developed are applied on a sample of 158 countries over the period of 1995-2014. Specifically, we investigate the effect of health expenditure separately on countries' technological change (shift of the frontier) and technological catch –up (distribution of efficiency) levels, allowing for the first time such distinct effects on countries' growth paths to be revealed. As a result our applied approach enables us also to uncover nonlinearities (threshold phenomena) which may be present when examining the health expenditure impact. The reminder of the paper is structured as follows: Section 2 presents the methodological framework, whereas, Section 3 presents our empirical findings. Finally, the last Section concludes the paper.

2. Methodological framework

Based on the probabilistic approach of nonparametric frontier analysis [3,18] we let countries' aggregate production to be characterized by the input vector $X \in \mathbb{R}_{+}^{p}$ and the output vector $Y \in \mathbb{R}_{+}^{q}$, whereas let $V \in \mathbb{R}_{+}^{d}$ denote the per capita health expenditure. Then we can define the attainable set as:

$$\tau = \{(x, y) | H_{X,Y}(x, y) > 0\}, \text{ where } H_{X,Y}(x, y) = \text{Prob}(X \leq x, Y \geq y). \quad (1)$$

Then country's production process at (x, y) can be measured on Farrell-Debreu's efficiency measure as:

$$\beta(x, y) = \sup\{\beta | (x, \beta y) \in \tau\} = \sup\{\beta | \mathcal{S}_{Y|X}(\beta y | x) > 0\}, \quad (2)$$

where $\mathcal{S}_{Y|X}(y|x) = \text{Prob}(Y \geq y | X \leq x)$. Then at time s the attainable set $\tau_s^V \subset \mathbb{R}_{+}^{p+q}$ can be characterized as the support of the following conditional probability:

$$H_{X,Y|V}^{s}(x, y|v) = \text{Prob}(X \leq x, Y \geq y | V = v, S = s). \quad (3)$$

As a result under the effect both of time s and healthcare expenditure v we can define country's production efficiency levels $(x, y) \in \tau_s^v$ as:

$$\beta_s(x,y|v) = \sup\{\beta|(x,\beta y) \in \tau_s^v\} = \sup\{\beta|\mathcal{S}_{Y|X,V}^s(\beta y|x,v) > 0\}, \tag{4}$$

where $\mathcal{S}_{Y|X,V}^s(y|x,v) = \text{Prob}(Y \geq y|X \leq x, V = v, S = s)$.

Let our panel data $x_{i,s}, y_{i,s}, v_{i,s}$ for $i = 1, \dots, n$ and $s = 1, \dots, w$, then we can estimate the unconditional and conditional attainable sets using data envelopment analysis (DEA) measures as:

$$\hat{\tau}_{DEA} = \{(x,y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q | y \leq \sum_{m=(i,s)} \omega_m y_m; x \geq \sum_{m=(i,s)} \omega_m x_m; \omega \geq 0, \text{s.t.} \sum_{m=(i,s)} \omega_m = 1\}, \tag{5}$$

$$\hat{\tau}_{s,DEA}^v = \{(x,y) \in \mathbb{R}_+^p \times \mathbb{R}_+^q | y \leq \sum_{m\in\Phi(v,s)} \omega_m y_m; x \geq \sum_{m\in\Phi(v,s)} \omega_m x_m; \omega \geq 0, \text{s.t.} \sum_{m\in\Phi(v,s)} \omega_m = 1\}, \tag{6}$$

where $\Phi(v,s) = \{m = (i,\eta)|v - h_v < v_{i,\eta} < v + h_v; s - h_s < \eta < s + h_s\}$, describes the localizing procedure; and h_v, h_s represent the bandwidths using data-driven methods [4]. Moreover, following Hall et al. [14] the estimation of the conditional distribution presented previously can be estimated as:

$$\hat{\mathcal{S}}_{Y|X,V}^s(y|x,v) = \frac{\sum_{m=(i,\eta)} \mathbb{I}(x_m \leq x, y_m \geq y) K_{h_v}(v_m - v) K_{h_s}(\eta - s)}{\sum_{m=(i,\eta)} \mathbb{I}(x_m \leq x) K_{h_v}(v_m - v) K_{h_s}(\eta - s)}. \tag{7}$$

In equation (7) $\mathbb{I}(\cdot)$ is the indicator function and $K(\cdot)$ are kernels with compact support (in our case Epanechnikov). The optimal bandwidths have been selected using the least squares cross-validation (LSCV) criterion as has been described in the procedure by Bădin et al. [4]. Another issue that needs to be addressed is the i.i.d. structure of the data in the smoothing techniques applied in (7). Based on Mastromarco and Simar [18] since the time variable takes many different values continues kernels can be applied. In addition as has been stressed by Hart [15] since in our smoothing we are applying Epanechnikov kernels with a support [-1,1] the independence of the observations can be assumed. This is evident since the dependency is decayed within the "window" and therefore makes the data within that "window" independent from the rest of the data. This process is also known as the "whitening by windowing" principle [15].

Additionally to the full frontiers presented previously we follow Daouia and Simar [9] and we apply *Order-α* quantile (robust) efficiency measures[1] which can be presented for the unconditional and conditional case as:

$$\beta_\alpha(x, y) = \sup\{\beta | \mathcal{S}_{Y|X}(\beta y|x) > 1 - \alpha\},$$

$$\beta_{s,\alpha}(x, y|v) = \sup\{\delta | \mathcal{S}^t_{Y|X,V}(\beta y|x, v) > 1 - \alpha\}. \quad (8)$$

Then as a robustness check we further apply the *Order-m* frontiers first introduced by Cazals et al. [8]. For a given level of x let $m\ i.i.d.$ random variables $Y_i, i = 1, \dots, m$ which is generated by the conditional q-variate distribution function $\Phi_{Y|X}(y|x) = \text{Prob}(Y \leq y|X \leq x)$. Then we can define the set:

$$\tau_m(x) = \{(\acute{x}, y) \in R_+^{p+q} | \acute{x} \leq x, Y_i \leq y, i = 1, \dots, m\}. \quad (9)$$

We can define the *Order-m* efficiency score as:

$$\tilde{\beta}_m(x, y) = \sup\{\beta | (x, \beta y) \in \tau_m(x)\} = \underset{i=1,\dots,m}{max}\left\{\underset{j=1,\dots,q}{min}\left(\frac{Y_i^j}{y^j}\right)\right\}. \quad (10)$$

Following Daraio and Simar [10] we can adapt the following Monte Carlo algorithm in order to obtain the estimated efficiency score:

[Step 1] Draw a sample of size *m* (with replacement) among Y_i with $X_i \leq x$ denoted as $(Y_{1,b,\dots,}Y_{m,b})$.

[Step 2] Compute $\tilde{\beta}_m^b(x, y) = \underset{i=1,\dots,m}{max}\left\{\underset{j=1,\dots,q}{min}\left(\frac{Y_{i,b}^j}{y^j}\right)\right\}$.

[Step 3] Compute again Step 1 and Step 2 for $b = 1, \dots, B$, where B is large.

[Step 4] Then $\hat{\beta}_m(x, y) \approx \frac{1}{B}\sum_{\beta=1}^{B} \tilde{\beta}_m^b(x, y)$

[1]In these partial frontiers α∈(0,1), in our case we have used as α a median value (α=0.5) in order to evaluate later the effect on countries' technological catch-up levels [3]. Jeong et al. [16] provide the asymptotic properties of the conditional estimators applied.

In a similar manner following the smoothing procedures described previously we can estimate the conditional *Order-m* frontiers as:

$$\beta_{s,m}(x,y|v,s) = E_{Y|X,V,S}\left(\tilde{\beta}_m^{v,s}(x,y)\middle|X \leq x, V = v, S = s\right) \quad (11)$$

As suggested by Bădin et al. [3], in order to explore the effect of per capita healthcare expenditure and time on countries' production efficiency levels we construct the following ratios:

$$\hat{Q}(x,y|v,s) = \frac{\hat{\beta}_s(x,y|v)}{\hat{\beta}(x,y)}, \hat{Q}_\alpha(x,y|v,s) = \frac{\hat{\beta}_{s,\alpha}(x,y|v)}{\hat{\beta}_\alpha(x,y)}, \hat{Q}_m(x,y|v,s) = \frac{\hat{\beta}_{s,m}(x,y|v)}{\hat{\beta}_m(x,y)}. \quad (12)$$

Then we apply $\hat{Q}(x,y|v,s)$, $\hat{Q}_\alpha(x,y|v,s)$ and $\hat{Q}_m(x,y|v,s)$ as a function of S and V using a local linear nonparametric regression examine the effect of time and healthcare on countries' technological change $[\hat{Q}(x,y|v,s)]$ and technological catch-up $[\hat{Q}_\alpha(x,y|v,s)]$ levels. Finally, in order to check the robustness of our findings, $\hat{Q}_m(x,y|v,s)$ reveal us a robust picture of the effect of time and healthcare on countries' technological change. Therefore in its general multivariate form the nonparametric regression will take the form:

$$Q_i = f(\Delta_i) + \epsilon_i, i = 1, \dots, \mathrm{n} \quad (13)$$

where Q_i are the ratios in equation (12) and $\{\Delta_i = (\Delta_{i1}, \dots, \Delta_{il})^T\}_{i=1}^n$ are i.i.d. random vectors and representing v and s. Then the local linear estimator is $\hat{\alpha}$ representing the conditional mean function $f(\cdot)$, at $\delta = (\delta_1, \dots, \delta_l)^T$ and can be obtained by solving the following problem:

$$\min_{\alpha,\beta} \sum_{i=1}^n \{Q_i - a - \beta^T(\Delta_i - \delta)\}^2 \prod_{j=1}^l K\left(\frac{\Delta_{ij} - \delta_j}{b_j h}\right). \quad (14)$$

In equation (14) $K(\cdot)$ represents the Epanechnikov kernel and h the optimal bandwidths using the criterion [14]. As has been analyzed by Bădin et al. [3] an increasing nonparametric regression line will signify a positive effect of v and s on countries' production performance levels, whereas, a decreasing line will suggest a negative effect.

3. Empirical findings

This paper uses a sample of 158 countries[2] over the period 1995-2014 extracted from Penn World Table v9.0 [12]. For the purpose of our analysis we consider capital stock and labor force as inputs and GDP as output. Finally, the $V \in \mathbb{R}_+^d$ variable is represented by the per capita government expenditure on health, which have been extracted from World Health Organization database.

Figure 1 presents diachronically the unconditional Farrell-type production efficiency measures (subfigures 1a and 1b)[3]. Specifically, subfigure 1a presents diachronically countries' production efficiency measures under the assumption of variable returns to scale, based on their income classification. As it was expected low income countries achieve the worst performance, whereas, the high income countries achieve the highest. However, it is also evident that upper middle income countries after the initiation of Global Financial Crisis-GFC (2007-2008 period) appear to catch-up more with high income countries' productive efficiency levels. Similarly when looking the estimated production efficiency levels of the largest economies (subfigure 1b), it is evident that Russian Federation, China and the United States have similar efficiency levels especially after 2011. Nevertheless, the results signify that the United States is the most efficient country in terms of production efficiency throughout the entire examined period.

Figure 1 about here

[2]Specifically our sample contains: 49 high income, 43 upper middle income, 43 lower middle income and 23 low income countries. All data are available upon request.
[3]The efficiency scores equal to 1 indicate efficient countries, whereas, values >1 indicate inefficiency. The analytical per country results are available upon request.

Based on Bădin et al. [3], Figure 2 presents the three-dimensional pictures of the main effect (based on the entire sample) of time and per capita health expenditure on countries' technological change (subfigure 2a) and technological catch-up (subfigure 2b) levels. The results reveal that the effect of health expenditure both on countries' technological change and technological catch-up is nonlinear and exhibits a "U" shape form. This indicates that for the largest part of the per capita health expenditure the effect is positive. However, after a certain threshold level a deteriorating effect on countries' technological change and technological catch-up is revealed. Finally, it is observed that over the examined period the effect of time on technological change is negative, whereas, on technological catch-up is positive. Our results confirm (at least for the largest part of countries' per capita health expenditure levels) the results by several other studies [5, 6, 20] suggesting a positive effect of health on countries' economic growth. Similar to Azariadis and Drazen [2], threshold effects are also reported, confirming a nonlinear relationship between health expenditure and economic growth [1].

As a robustness check we perform our analysis based on countries' income classification. The overall finding of an inverted "U" shape effect is also confirmed for the cases of technological change of high, upper-middle and low income countries (subfigure 2c, 2e, 2i). However, a 'U'-shape effect is evident for the low middle income countries (subfigure 2g). In addition the inverted "U" shape effect for countries' technological catch-up levels is confirmed only for the cases of upper middle and low income countries (subfigure 2f and 2j). However, a 'U'-shape effect on countries' technological catch-up levels is evident for the cases of high and low middle income countries (subfigures 2d and 2h). Even though the results confirm the nonlinear relationship of health expenditure and countries' technological change and technological catch-up levels, the shape of the effect is determined by the countries' different income levels.

Finally, as a further robustness check of our findings, Figure 3 presents the dynamic effects of health expenditure on countries' technological change levels using the *Order-m* frontiers presented in equations (10) and (11). The adopted frontiers are less sensitive to extreme values and outliers and they don't suffer from the curse of dimensionality. In order to compute the unconditional and conditional estimates we have adopted the Monte Carlo algorithm [10] presented previously and we have set $m = 50$[4]. Our findings verify the results presented previously regarding the effect of time and per capita health expenditure on countries' technological change. Finally, it is evident that under the application of *Order-m* frontiers our overall findings are verifying the existence of nonlinear effects. This is also evident both for the entire sample (subfigure 3a) but also when accounting for countries' income classification.

Figure 2 & 3 about here

4. Conclusions

By employing time-dependent conditional efficiency estimators the paper provides an innovative application examining in a nonparametric framework the effect of time and per capita healthcare expenditure on countries' technological change and technological catch-up. The nonparametric frontier analysis is applied on a sample of 158 countries over the period 1995-2014. The overall results present for the first time evidence of a nonlinear effect of healthcare expenditure on countries' technological change and technological catch-up, indicating an initial inverted "U"-shape relationship. However, when the analysis was carried forward taking into account

[4] Different values of *m* have been applied and the results are available upon request. Note also that when the *Order-m* estimators are $\sqrt{n}$ consistent and $\hat{\beta}_{m,n}(x,y) \rightarrow \hat{\beta}_n(x,y)$ when $m \rightarrow \infty$.

countries' income classifications the effect in many cases has been changed to a 'U'-shape form revealing that the influence of both time and healthcare expenditure is determined by the countries' different income levels. Finally the results revealed that for high income countries the per capita health expenditure impacts differently on countries' technological change and technological catch-up levels.

Acknowledgements

We would like to thank the Associate Editor and the reviewers of the journal for the comments raised to our manuscript. Any remaining errors are solely the authors' responsibility.

References

[1] R. Aisa and F. Pueyo, *Government health spending and growth in a model of endogenous longevity*, Econ. Lett. 90 (2006), pp. 249-253.

[2] C. Azariadis and A. Drazen, *Threshold externalities in economic development*, Q. J. Econ. 105 (1990), pp. 501-526.

[3] L. Bădin, , C. Daraio and L. Simar, *How to measure the impact of environmental factors in a nonparametric production model*, Eur. J. Oper. Res. 223 (2012), pp. 818-833.

[4] L. Bădin, , C. Daraio and L. Simar, *Optimal bandwidth selection for conditional efficiency measures: a data-driven approach*, Eur. J. Oper. Res. 201 (2010), pp. 633-640.

[5] E. Baldacci, B. Clements, S. Gupta, and Q. Cui, *Social spending, human capital, and growth in developing countries*, World Dev. 36 (2008), pp. 1317-1341.

[6] D. Bloom and D. Canning, *The health and poverty of nations: from theory to practice,* J. Hum. Dev. 4 (2003), pp. 47-71.

[7] I. Bournakis and S. Mallick, *TFP estimation at firm level: The fiscal aspect of productivity convergence in the UK*, Econ. Model. 70 (2018), pp. 579-590.

[8] C. Cazals, J.P. Florens and L. Simar, *Nonparametric frontier estimation: a robust approach*, J. Econometrics 106 (2002), 1-25.

[9] A. Daouia and L. Simar, *Nonparametric efficiency analysis: a multivariate conditional quantile approach*, J. Econometrics 140 (2007), pp. 375-400.

[10] C. Daraio and L. Simar, *Introducing environmental variables in nonparametric frontier models: a probabilistic approach*, J. Prod. Anal. 24 (2005), 93-121.

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

[12] R. C. Feenstra, R. Inklaar and M. P. Timmer, *The next generation of the Penn World Table*, Am. Econ. Rev. 105 (2015), pp. 3150-82.

[13] P. Gupta, S. Mallick and T. Mishra, *Does social identity matter in individual alienation? Household-level evidence in post-reform India*, World Dev. 104 (2018), pp. 154-172.

[14] P. Hall, J. Racine and Q. Li, *Cross-validation and the estimation of conditional probability densities*, J. Am. Stat. Assoc. 99 (2004), pp. 1015-1026.

[15] J.D. Hart, *Some automated methods of smoothing time-dependent data*, J. Nonparametr. Stat. 6 (1996), pp. 115-142.

[16] S.O., Jeong, B.U. Park and L. Simar, *Nonparametric conditional efficiency measures: asymptotic properties*. Ann. Oper. Res. 173 (2010), pp. 105-122.

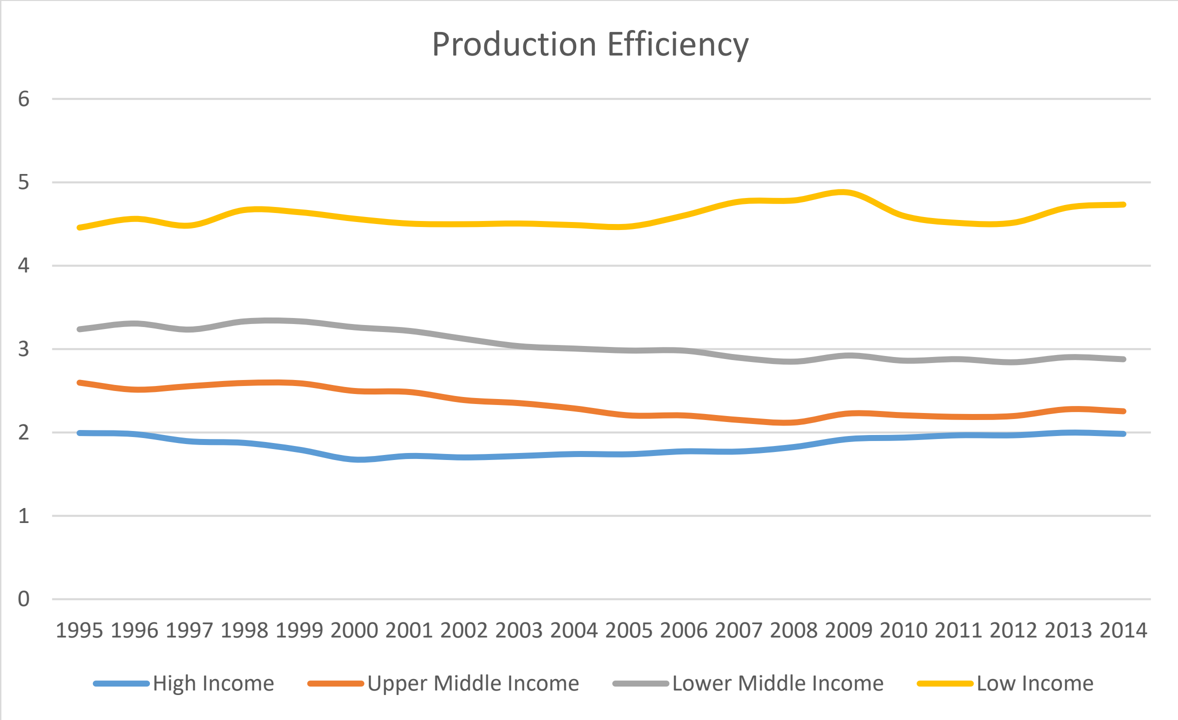
[17] S. Mallick, R. Matousek and N.G. Tzeremes, *Financial development and productive inefficiency: A robust conditional directional distance function approach*, Econ. Lett. 145 (2016), pp. 196-201.

[18] C. Mastromarco and L. Simar, *Effect of FDI and time on catching up: New insights from a conditional nonparametric frontier analysis*, J. Appl. Econometrics 30 (2015), pp. 826-847.

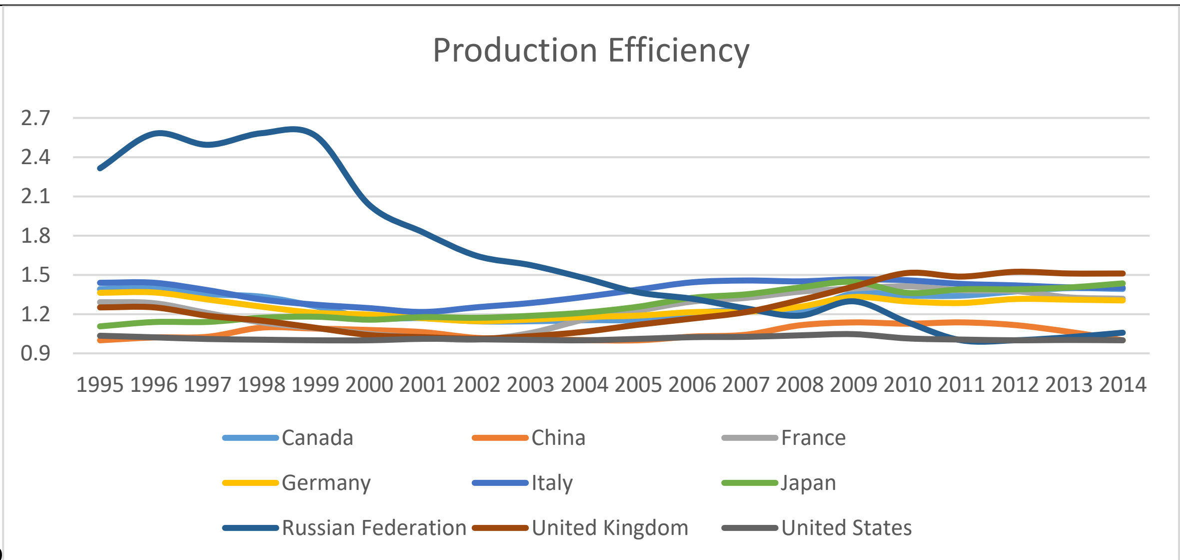
[19] J.P. Newhouse, *Medical care costs: how much welfare loss*? J. Econ. Perspect. 6 (1992), pp. 3-21.

[20] B. Rivera and L. Currais, *Public health capital and productivity in the Spanish regions: a dynamic panel data model*, World Dev. 32 (2004), pp. 871-885.

Figure 1: Diachronic representation of estimated production efficiency levels

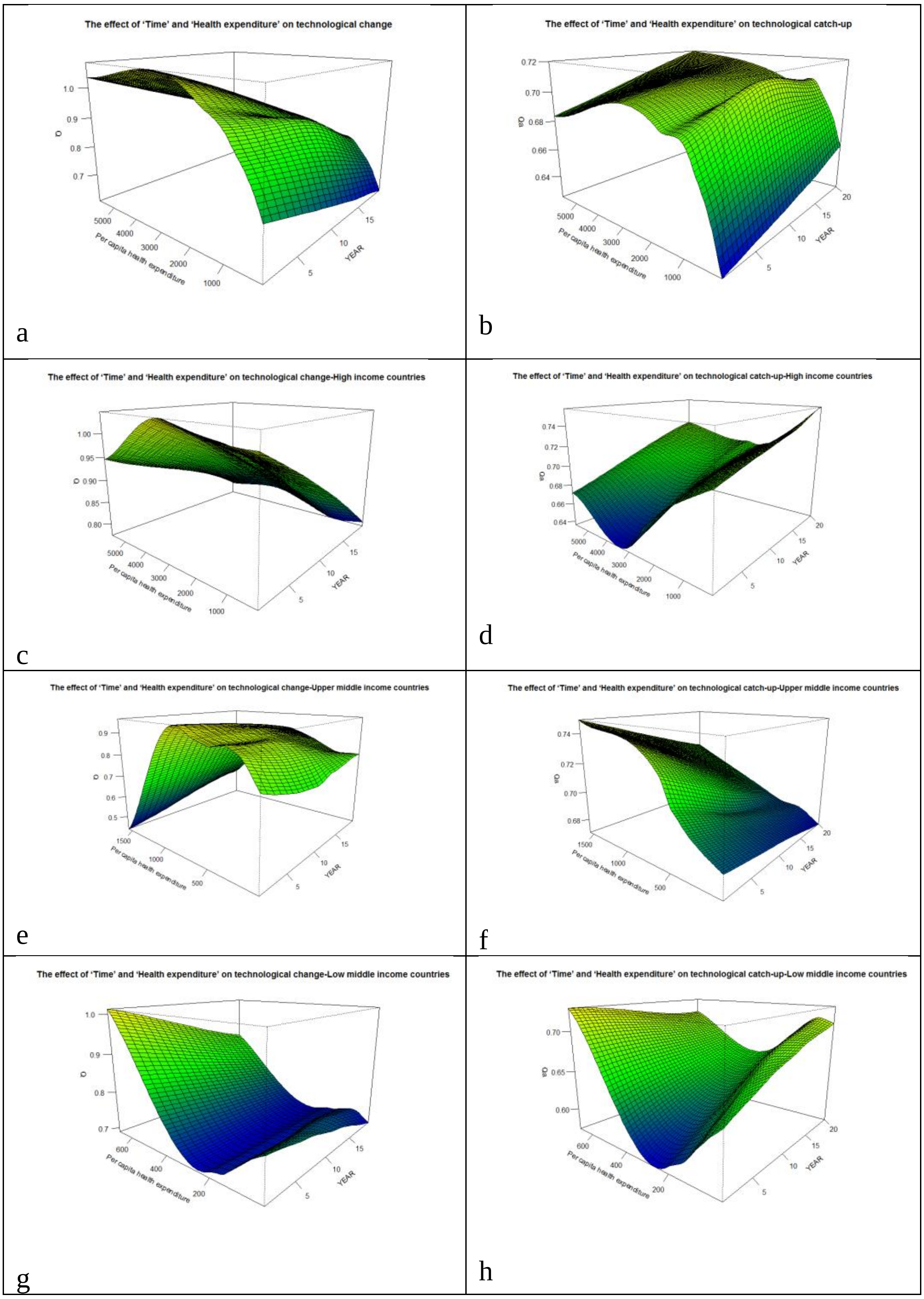


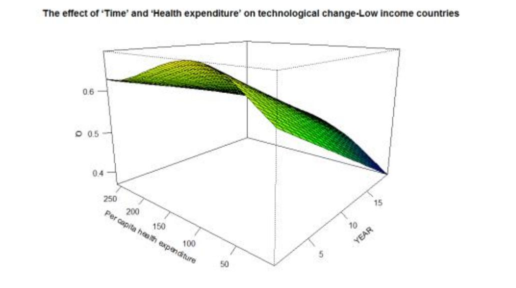
a



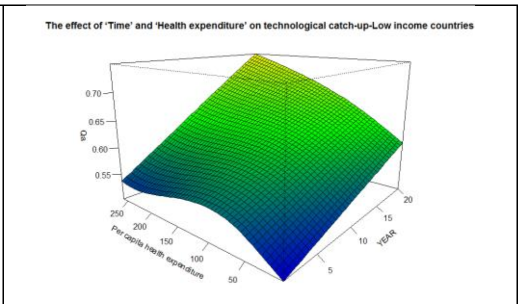
b

Figure 2: The effect of time and per capita health expenditure on countries' technological change and catch-up levels




The effect of 'Time' and 'Health expenditure' on technological change-Low income countries
0.6
0.5
0.4
250
200
150
100
50
Per capita health expenditure
5
10
15
YEAR

i


The effect of 'Time' and 'Health expenditure' on technological catch-up-Low income countries
0.70
0.65
0.60
0.55
250
200
150
100
50
Per capita health expenditure
5
10
15
20
YEAR

j

Figure 3: The effect of time and per capita health expenditure on countries' technological change using Order-m efficiency measures