KOLMOGOROV-SMIRNOV TYPE TEST FOR GENERATED VARIABLES

TAISUKE OTSU AND GO TANIGUCHI

Abstract. Distribution homogeneity testing, particularly based on the Kolmogorov-Smirnov statistic, has been applied in various empirical studies. In empirical economic analysis, it is often the case that economic variables of interest are obtained as estimated values or residuals of preliminary model fits, called generated variables. In this paper, we extend the Kolmogorov-Smirnov type homogeneity test to accommodate such generated variables, and propose an asymptotically valid bootstrap inference procedure. A small simulation study illustrates that it is crucial for reliable inference to account for estimation errors in the generated variables. The proposed method is applied to compare the total factor productivities across different countries.

1. Introduction

In various areas of empirical studies, researchers are often interested in testing homogeneity of distributions across different samples. The most popular approach for distribution homogeneity testing is based on the Kolmogorov-Smirnov statistic, which is obtained as the largest discrepancy of the empirical distribution functions by these samples, and statistical theory of the Kolmogorov-Smirnov test is well known (e.g., Lehmann and Romano, 2005). An obvious premise behind this statistical theory is that the samples of interest are all observable.

In empirical economic analysis, however, it is often the case that researchers are interested in the distributions of latent or theoretical variables, which are unobservable but may be estimated by observable data. Such variables, called generated variables, are often obtained as fitted values or residuals of preliminary regression fitting. Common examples include expected values of prices or sales, total factor productivity, relative quality of firms, among others. A major econometric issue of use of generated variables is that its statistical inference requires to account for estimation errors contained in the generated variables. Pagan (1984) investigated how to modify standard errors for regression analysis using generated regressors. Hahn and Ridder (2013) studied standard error formulae for semiparametric estimators that involve generated variables. Matsushita and Otsu (2019) proposed a likelihood-based inference method for semiparametric models with generated variables.

In this paper, we consider distribution homogeneity testing of generated variables. The test statistic is the Kolmogorov-Smirnov statistic using the generated variables. However, the null distribution is different from the conventional one due to the estimation errors to construct the generated variables, and takes a somewhat complicated form. Therefore, we propose a bootstrap procedure to compute the critical value for the Kolmogorov-Smirnov statistic. A key ingredient of our bootstrap procedure is to recenter for the bootstrap statistic to impose the null hypothesis (cf. Whang, 2001). By adapting the asymptotic theory developed in Linton, Maasoumi and Whang (2005), the asymptotic validity of our bootstrap procedure is established.

The proposed homogeneity test is illustrated by Monte Carlo simulation and an empirical application. Our simulation study illustrates that it is crucial for reliable inference to account for estimation errors in the generated variables. Especially the conventional Kolmogorov-Smirnov test may exhibit severe size distortions. Also our empirical application on comparisons of the total factor productivities across different countries illustrate usefulness of the proposed test.

2. Distribution homogeneity test

Consider scalar latent variables X_{ki} for $k = 1, \dots, K$, which are not directly observable. Suppose that the latent variables are specified as $X_{ki} = g(Z_{ki}, \theta_{k0})$, where g is a known function up to the unknown parameters θ_{k0} and Z_{ki} is a vector of observables. Let F_k be the cumulative distribution function of X_k for $k = 1, \dots, K$. We wish to conduct hypothesis testing on distribution homogeneity of the latent variables

$$H_0 : F_1(x) = \cdots = F_K(x) \quad \text{for all } x,$$

against $H_1 : H_0$ is false. If $\{X_{ki}\}_{i=1}^n$ is directly observable, then various homogeneity tests are available in the literature. In this paper we focus on the situation where some estimates $\hat{\theta}_k$ for the parameters θ_{k0} are available to the researcher so that the generated variables $\hat{X}_{ki} = g(Z_{ki}, \hat{\theta}_k)$ can be constructed for $i = 1, \dots, n$.

A typical example of generated variables is the error term e_{ki} in a regression model $Z_{ki}^{(1)} = m(Z_{ki}^{(2)}, \theta_{k0}) + e_{ki}$ for some regression function m. In this case, we set $X_{ki} = e_{ki}$ and $g(Z_{ki}, \theta_{k0}) = Z_{ki}^{(1)} - m(Z_{ki}^{(2)}, \theta_{k0})$, and $\hat{\theta}_k$ may be given by the (nonlinear) least squares or GMM estimator. Another example would be expected values of some economic variables (e.g., expected inflation or demand). In this case, X_{ki} may be written by some conditional expectation $X_{ki} = E[Z_{ki}^{(1)} | Z_{ki}^{(2)}]$ using observables and its proxy $\hat{X}_{ki}$ is typically obtained as a fitted value for a regression model on $E[Z_{ki}^{(1)} | Z_{ki}^{(2)}]$. Furthermore, the proxy $\hat{X}_{ki}$ for the latent variable X_{ki} may emerge from more involved structural economic models. For example, Berry, Levinsohn and Pakes (1995) proposed a method to estimate unobservable qualities of goods, which are estimated by the simulated method of moments. In this case, estimated qualities may be considered as generated variables. The estimation method for production functions developed by Olley and Pakes (1996) is another example, where the estimated unobservable productivities may be considered as generated variables.

We now present our testing approach. Let $\hat{F}_k(x) = n^{-1} \sum_{i=1}^n \mathbb{I}\{\hat{X}_{ki} \leq x\}$ be the empirical distribution function based on the generated variables $\{\hat{X}_{ki}\}_{i=1}^n$, where $\mathbb{I}\{\cdot\}$ is the indicator function. $\hat{F}_k(x)$ is a consistent estimator of $F_k(x)$ under mild regularity conditions as far as $\hat{\theta}_k$ is consistent for θ_{k0}. In order to test the distribution homogeneity hypothesis H_0, we employ the Kolmogorov-Smirnov type statistics

$$\begin{aligned} KS_1 &= \max_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{n} |\hat{F}_k(x) - \hat{F}_l(x)|, \qquad (1) \\ KS_2 &= \max_{k} \sup_{x \in \mathcal{X}} \sqrt{n} \left| \hat{F}_k(x) - \frac{1}{K} \sum_{l=1}^{K} \hat{F}_l(x) \right|, \end{aligned}$$

where $\mathcal{X}$ is a given set. The first statistic KS_1 is the maximal pairwise sup-norm distance of the empirical distributions $\{\hat{F}_k(\cdot)\}_{k=1}^K$. The second statistic KS_2 is obtained by the maximal deviation from the average $K^{-1}\sum_{l=1}^K \hat{F}_l(x)$, which is equivalent to the empirical distribution by the pooled sample $\{\hat{X}_{ki}\}_{i=1}^n$ over $k = 1, \ldots, K$. When the number of different samples K is large, KS_2 is computationally more attractive.

The limiting distributions of the Kolmogorov-Smirnov statistics under the null hypothesis H_0 are obtained as follows.

Proposition 1. For each $k = 1, \ldots, K$, suppose

(i): $\{Z_{ki}\}_{i=1}^n$ is a strictly stationary and α-mixing sequence with the mixing coefficient $\alpha(m) = O(m^{-a})$ for some $a > \max\{(q-1)(q+1), 1 + 2/r\}$, where q is an even integer satisfying $q > 3(\max\{\dim \theta_1, \ldots, \dim \theta_K\} + 1)/2$ and r appears in (ii) below. For a neighborhood $\mathcal{N}_k$ around θ_{k0}, $g(Z_k, \theta_k)$ is continuously differentiable on $\theta_k \in \mathcal{N}_k$ with probability one, $E[\sup_{\theta_k \in \mathcal{N}_k} |\partial g(Z_k, \theta_k)/\partial \theta_k|^{2+\delta}] < \infty$ for some $\delta > 0$;

(ii): the estimator $\hat{\theta}_k$ satisfies $\sqrt{n}(\hat{\theta}_k - \theta_{k0}) = n^{-1/2} \sum_{i=1}^n \psi_k(Z_{ki}, \theta_{k0}) + o_p(1)$ with $E[\psi_k(Z_k, \theta_{k0})] = 0$ and $E[|\psi_k(Z_k, \theta_{k0})|^{2+r}] < \infty$ for some $r > 0$.

Then under H_0,

$$\begin{aligned} KS_1 &\xrightarrow{d} \max_{k \neq l} \sup_{x \in \mathcal{X}} \left| \nu_{kl}(x) + \frac{\partial F_k(x, \theta_{k0})}{\partial \theta_k'} \xi_k - \frac{\partial F_l(x, \theta_{l0})}{\partial \theta_l'} \xi_l \right|, \qquad (2) \\ KS_2 &\xrightarrow{d} \max_{k} \sup_{x \in \mathcal{X}} \left| \frac{1}{K} \sum_{l=1}^{K} \left\{ \nu_{kl}(x) + \frac{\partial F_k(x, \theta_{k0})}{\partial \theta_k'} \xi_k - \frac{\partial F_l(x, \theta_{l0})}{\partial \theta_l'} \xi_l \right\} \right|, \end{aligned}$$

where $(\nu_{kl}(\cdot), \xi_k', \xi_l')$ is a mean zero Gaussian process with the covariance kernel $C_{kl}(x_1, x_2) = \lim_{n\to\infty} E[S_{kl,n}(x_1) S_{kl,n}(x_2)']$,

$$\begin{aligned} S_{kl,n}(x) &= \left(\hat{\nu}_k(x, \theta_{k0}) - \hat{\nu}_l(x, \theta_{l0}), \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_k(Z_{ki}, \theta_{k0})', \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_l(Z_{li}, \theta_{l0})' \right)', \\ \hat{\nu}_k(x, \theta_k) &= n^{-1/2} \sum_{i=1}^n \{ \mathbb{I}\{g(Z_{ki}, \theta_k) \leq x\} - F_k(x, \theta_k) \}. \end{aligned}$$

□

If X_{ki}'s are directly observable, then the limiting null distributions of KS_1 and KS_2 reduce to $\max_{k \neq l} \sup_{x \in \mathcal{X}} |\nu_{kl}(x)|$ and $\max_k \sup_{x \in \mathcal{X}} |K^{-1} \sum_{l=1}^K \nu_{kl}(x)|$, respectively. Thus, the terms in (2) containing ξ_k's are considered as correction terms to account for the estimation errors of the parameter estimators, $\hat{\theta}_k$'s.

The assumptions for this proposition are adapted from Linton, Maasoumi and Whang (2005) to our setup. Assumption (i) is on the data and the shape of the function g to specify the latent variable. Assumption (ii) is a high level condition for the estimators $\{\hat{\theta}_k\}_{k=1}^K$. ψ_k is called the influence function for $\hat{\theta}_k$, and this assumption is typically satisfied for various estimators (e.g., the OLS, GMM, and maximum likelihood) under certain regularity conditions. For example, if $\hat{\theta}_k$ is a GMM estimator that minimizes $\{n^{-1}\sum_{i=1}^n m(Z_{ki},\theta_k)\}' W \{n^{-1}\sum_{i=1}^n m(Z_{ki},\theta_k)\}$ for some moment function m and weight W, then certain regularity conditions (see, e.g., Theorem 3.4 of Newey and McFadden, 1994) guarantee Assumption (ii) with $\psi_k(Z_{ki},\theta_{k0}) = -(M'WM)^{-1}M'Wm(Z_{ki},\theta_{k0})$ and $M = E[\partial m(Z_k,\theta_{k0})/\partial\theta_k']$.

Also, the proof is obtained by modifying that of Linton, Maasoumi and Whang (2005, Theorem 1). Here we sketch the proof for the case of KS_1. The proof for KS_2 is similar. Under H_0 (i.e., $F_k(x,\theta_{k0}) - F_l(x,\theta_{l0}) = 0$), KS_1 can be written as

$$KS_1 = \max_{k\neq l}\sup_{x\in\mathcal{X}} |\hat{\nu}_k(x,\hat{\theta}_k) - \hat{\nu}_l(x,\hat{\theta}_l) + \sqrt{n}\{F_k(x,\hat{\theta}_k) - F_l(x,\hat{\theta}_l)\} - \sqrt{n}\{F_k(x,\theta_{k0}) - F_l(x,\theta_{l0})\}|.$$

By Linton, Maasoumi and Whang (2005, Lemmas 2 and 3),

$$\sup_{x\in\mathcal{X}} |\hat{\nu}_k(x,\hat{\theta}_k) - \hat{\nu}_k(x,\theta_{k0})| \xrightarrow{p} 0,$$

$$\sup_{x\in\mathcal{X}} \sqrt{n}\left| F_k(x,\hat{\theta}_k) - F_k(x,\theta_{k0}) - \frac{\partial F_k(x,\theta_{k0})}{\partial\theta_k'}\frac{1}{n}\sum_{i=1}^n \psi_k(Z_{ki},\theta_{k0})\right| \xrightarrow{p} 0,$$

for $k = 1,\ldots,K$. Therefore, we obtain

$$KS_1 = \max_{k\neq l}\sup_{x\in\mathcal{X}} \left| \begin{array}{l} \hat{\nu}_k(x,\theta_{k0}) - \hat{\nu}_l(x,\theta_{l0}) \\ +\frac{\partial F_k(x,\theta_{k0})}{\partial\theta_k'}\frac{1}{n}\sum_{i=1}^n \psi_k(Z_{ki},\theta_{k0}) - \frac{\partial F_l(x,\theta_{l0})}{\partial\theta_l'}\frac{1}{n}\sum_{i=1}^n \psi_l(Z_{li},\theta_{l0}) \end{array}\right| + o_p(1), \quad (3)$$

and the conclusion in (2) follows by the weak convergence of the empirical process $\hat{\nu}_k(x,\theta_{k0})$ and continuous mapping theorem.

We can also show that the Kolmogorov-Smirnov tests based on KS_1 and KS_2 are consistent under fixed alternative hypotheses and have non-trivial power under the local alternative hypotheses at the rate of $n^{-1/2}$.

Due to the correction terms associated with ξ_k's, the limiting null distributions of the Kolmogorov-Smirnov type statistics are not asymptotically pivotal and somewhat complicated. Here we focus on the case where $\{Z_{1i},\ldots,Z_{Ki}\}_{i=1}^n$ is iid, and present the following bootstrap procedure to approximate the critical values. This procedure can be extended to the case of dependent data by modifying the resampling method, such as block bootstrap (see, Hall and Horowitz, 1996).

Bootstrap procedure for critical values:

(1) Draw the bootstrap resample $\{Z_{1i}^*,\ldots,Z_{Ki}^*\}_{i=1}^n$ from the joint empirical distribution of $\{Z_{1i},\ldots,Z_{Ki}\}_{i=1}^n$, and compute the estimator $\{\hat{\theta}_1^*,\ldots,\hat{\theta}_K^*\}$ by using $\{Z_{1i}^*,\ldots,Z_{Ki}^*\}_{i=1}^n$.

(2) Then compute the recentered bootstrap statistics

$$KS_1^* = \max_{k \neq l} \sup_x \sqrt{n} |\hat{F}_k^*(x) - \hat{F}_l^*(x) - \{\hat{F}_k(x) - \hat{F}_l(x)\}|,$$

$$KS_2^* = \max_k \sup_x \sqrt{n} \left| \hat{F}_k^*(x) - \frac{1}{K} \sum_{l=1}^{K} \hat{F}_l^*(x) - \left\{ \hat{F}_k(x) - \frac{1}{K} \sum_{l=1}^{K} \hat{F}_l(x) \right\} \right|,$$

where $\hat{F}_k^*(x) = n^{-1} \sum_{i=1}^n \mathbb{I}\{g(Z_{ki}^*, \hat{\theta}_k^*) \leq x\}$.

(3) Repeat (2) B times to obtain $\{KS_{1,b}^*\}_{b=1}^B$ (or $\{KS_{2,b}^*\}_{b=1}^B$). Their $(1-\alpha)$-th quantiles $q_{1,\alpha}^*$ (or $q_{2,\alpha}^*$) provide the bootstrap critical values of KS_1 (or KS_2) to test H_0.

The asymptotic properties of this bootstrap test are presented as follows. Let $E^*[\cdot]$ be the expectation relative to the distribution of the bootstrap sample conditional on the original sample.

Proposition 2. Suppose that Assumptions (i) and (ii) in Proposition 1 hold true for iid data $\{Z_{1i}, \ldots, Z_{Ki}\}_{i=1}^n$. Furthermore, for each $k = 1, \ldots, K$, assume that $\hat{\theta}_k^*$ satisfies $\sqrt{n}(\hat{\theta}_k^* - \hat{\theta}_k) = n^{-1/2} \sum_{i=1}^n \psi_k(Z_{ki}^*, \hat{\theta}_k) + o_p(1)$ conditional on $\{Z_{1i}, \ldots, Z_{Ki}\}_{i=1}^n$ with probability one for $E^*[\psi_k(Z_k^*, \hat{\theta}_k)] = 0$. Then for $l = 1, 2$, $\lim_{n\to\infty} \Pr\{KS_l \geq q_{l,\alpha}^*\} = \alpha$ under H_0 and $\lim_{n\to\infty} \Pr\{KS_l \geq q_{l,\alpha}^*\} = 1$ under H_1. $\square$

This proposition presents the asymptotic validity and consistency of our bootstrap tests. The proof follows by similar arguments in Whang (2001, Theorems 2 and 4). The additional assumption on $\hat{\theta}_k^*$ is a bootstrap analog of Assumption (ii) in Proposition 1.

Note that the recentering in KS_1^* or KS_2^* is crucial to impose the null hypothesis H_0. The idea of recentering was suggested in the literature by Hall and Horowitz (1996), Whang (2001), and Linton, Maasoumi and Whang (2005), for example. To obtain an intuition, consider the bootstrap counterpart KS_1^*. We can show that KS_1^* satisfies

$$KS_1^* = \max_{k \neq l} \sup_{x \in \mathcal{X}} \left| \begin{array}{l} \hat{\nu}_k^*(x, \hat{\theta}_k) - \hat{\nu}_l^*(x, \hat{\theta}_l) \\ + \frac{\partial F_k(x,\theta_{k0})}{\partial \theta_k'} \frac{1}{n} \sum_{i=1}^n \psi_k(Z_{ki}^*, \hat{\theta}_k) - \frac{\partial F_l(x,\theta_{l0})}{\partial \theta_l'} \frac{1}{n} \sum_{i=1}^n \psi_l(Z_{li}^*, \hat{\theta}_l) \end{array} \right| + o_p(1),$$

conditional on $\{Z_{1i}, \ldots, Z_{Ki}\}_{i=1}^n$ with probability one, where $\hat{\nu}_k^*(x, \theta_k) = n^{-1/2} \sum_{i=1}^n [\mathbb{I}\{g(Z_{ki}^*, \theta_k) \leq x\} - \mathbb{I}\{g(Z_{ki}, \theta_k) \leq x\}]$. This expression is analogous to (3) and guarantees the asymptotic validity of our bootstrap procedure. Without recentering in KS_1^*, we would have additional terms in the above expansion of KS_1^* that may diverge.

3. Simulation

To illustrate the finite sample performance of our Kolmogorov-Smirnov type tests, we conduct a small simulation study. We consider two regression models

$$Y_{1i} = 1 + (1,1)X_{1i} + e_{1i}, \qquad Y_{2i} = 1 + (1,1)X_{2i} + e_{2i}, \tag{4}$$

for $i = 1, \ldots, n$, where X_{1i} and X_{2i} are bivariate regressors generated from

$$X_{1i} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}\right), \qquad X_{2i} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.7 \\ 0.7 & 1 \end{pmatrix}\right).$$

In this setup, we consider homogeneity testing of the distribution functions of the error terms e_1 and e_2, i.e.,

$$H_0 : F_{e_1}(t) = F_{e_2}(t) \quad \text{for all } t,$$

based on the OLS residuals $\hat{e}_{1i}$ and $\hat{e}_{2i}$ for the regressions from Y_1 on X_1 and Y_2 on X_2, respectively.

For the test statistic KS_1 in (1), we compare the proposed bootstrap procedure with the conventional Kolmogorov-Smirnov critical value that does not take into account for the estimation errors $\hat{e}_{1i} - e_{1i}$ and $\hat{e}_{2i} - e_{2i}$. Note that the conventional critical value is asymptotically invalid and is used to illustrate importance of accommodating the estimation errors for the generated variables. Furthermore, to evaluate the effect of estimation errors, $\hat{e}_{1i} - e_{1i}$ and $\hat{e}_{2i} - e_{2i}$, we also report infeasible versions of the bootstrap and Kolmogorov tests for H_0 based on e_{1i} and e_{2i} using knowledge of the true parameter values. Note that the conventional Kolmogorov test based on e_{1i} and e_{2i} is asymptotically valid since it does not involve generated variables.

To assess the size properties, we consider three distributions of e_1 and e_2: $N(0, 1)$, standardized $t(3)$, standardized $\chi^2(3)$, and standardized $LN(0, 1)$ (log-normal generated by $\exp(Z)$ for $Z \sim N(0, 1)$). For the sample size, we consider n =100, 200, and 500. The number of bootstrap replications is 99 and the number of Monte Carlo replications is 1000. The nominal size is 0.05. Table 1 presents the rejection frequencies under the null hypotheses. Our bootstrap procedure works well for all cases although it shows under-coverage for most cases. Interestingly, its infeasible version based on e_1 and e_2 does not exhibit such under-coverage. Thus we conjecture the under-coverage for our bootstrap procedure is due to estimation errors for generated variables. Although it is beyond the scope of this paper, it is interesting to investigate whether this under-coverage is a general phenomenon for our bootstrap procedure.

On the other hand, the conventional Kolmogorov-Smirnov critical value clearly fails to control the size for the case of $LN(0, 1)$. This point is confirmed by comparing with its infeasible version based on e_1 and e_2, which is asymptotically valid. This indicates that it is crucial to take into account for the estimation errors of the generated variables to conduct homogeneity testing.

We next evaluate power properties of the proposed test. Now the error terms in (4) are generate by

$$\begin{aligned} e_{1i} &= \sqrt{\rho}\epsilon_{0i} + \sqrt{1-\rho}\epsilon_{1i}, \\ e_{2i} &= \sqrt{\rho}\epsilon_{0i} + \sqrt{1-\rho}\epsilon_{2i}, \end{aligned}$$

for $i = 1, \ldots, n$, where $\rho \in \{0.0, 0.2, 0.4\}$, $\epsilon_{0i} \sim N(0, 1)$, $\epsilon_{1i} \sim N(0, 1)$, and ϵ_{2i} follows the standardized $t(3)$, standardized $\chi^2(3)$, and standardized $LN(0, 1)$. The distributions of e_{1i} and e_{2i} become similar as ρ increases.

TABLE 1. Rejection Frequencies (Size)

$F_{e_1} = F_{e_2}$	Method	n 100	200	500
$N(0,1)$	Bootstrap	.001	.003	.006
	Conventional KS	.001	.001	.002
	Bootstrap (based on e_1 and e_2)	.046	.035	.046
	Conventional KS (based on e_1 and e_2)	.058	.067	.082
$t(3)$	Bootstrap	.009	.007	.013
	Conventional KS	.052	.064	.078
	Bootstrap (based on e_1 and e_2)	.064	.057	.047
	Conventional KS (based on e_1 and e_2)	.079	.088	.095
$\chi^2(3)$	Bootstrap	.009	.017	.018
	Conventional KS	.005	.072	.069
	Bootstrap (based on e_1 and e_2)	.052	.048	.045
	Conventional KS (based on e_1 and e_2)	.085	.076	.092
$LN(0,1)$	Bootstrap	.043	.048	.053
	Conventional KS	.390	.441	.506
	Bootstrap (based on e_1 and e_2)	.060	.040	.040
	Conventional KS (based on e_1 and e_2)	.081	.064	.078

Table 2 presents the rejection frequencies under the alternative hypotheses based on 1000 Monte Carlo replications. The proposed bootstrap test shows reasonable power properties when the sample size is large enough. The power of the conventional Kolmogorov-Smirnov for $LN(0,1)$ is spurious because of severe over-rejection under the null hypothesis. Also, as ρ increases, the power of the conventional Kolmogorov-Smirnov deteriorates faster than the bootstrap test. Finally, although the results are not reported, we find that the bootstrap statistic without recentering has zero power for all cases.

TABLE 2. Rejection Frequencies (Power)

		$n = 100$			$n = 200$			$n = 500$		
		ρ			ρ			ρ		
F_{ϵ_2}	Method	0.0	0.2	0.4	0.0	0.2	0.4	0.0	0.2	0.4
$t(3)$	Bootstrap	.087	.024	.007	.378	.108	.026	.946	.579	.191
	Conventional KS	.128	.027	.007	.495	.107	.022	.987	.539	.084
$\chi^2(3)$	Bootstrap	.114	.045	.016	.605	.243	.061	1	.883	.416
	Conventional KS	.255	.063	.012	.779	.239	.049	1	.841	.255
$LN(0,1)$	Bootstrap	.312	.195	.038	.780	.719	.237	.997	.997	.901
	Conventional KS	.979	.384	.067	1	.884	.260	1	1	.855

4. EMPIRICAL ILLUSTRATION: TOTAL FACTOR PRODUCTIVITY

We apply our Kolmogorov-Smirnov type test to compare the total factor productivities among different countries. Following Solow's (1957) classical approach, we specify the production function as $Y_t = A_t K_t^{\alpha} L_t^{\beta}$, where Y_t is total output, K_t and L_t are capital and labor inputs, respectively, and A_t is total factor productivity. Solow (1957) specified the production function as $Y_t = A_t F(K_t, L_t)$, where Y_t is total output, K_t and L_t are capital and labor inputs, respectively, and A_t is total factor productivity.[1] If we specify F by the Cobb-Douglas function $F(K_t, L_t) = K_t^\alpha L_t^\beta$, we can derive

$$\Delta \log(A_t) = \Delta \log(Y_t) - \alpha \Delta \log(K_t) - \beta \Delta \log(L_t). \tag{5}$$

We use the datasets offered by the Federal Reserve Bank of St. Louis (FRED), OECD, United Nations Statistics Division, and World Bank (see Table 3).

TABLE 3. List of variables

Variable	Measurement
Y	Gross domestic products in the country (World Bank)
I	Gross fixed capital formation in the country (FRED)
L	Population (OECD) $\times$ Percentage of working age population (OECD) $\times$ Average annual hours worked by persons engaged for the country (FRED)

Using these data, we calculate the annual growth rate of total factor productivity from 1971 to 2014 for 15 countries (Australia, Austria, Belgium, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, Norway, South Korea, Sweden, United Kingdom, and United States). In particular, $\Delta \log(A_t)$ is estimated by evaluating α and β in (5) with the OLS estimator.

We apply the Kolmogorov-Smirnov type test for generated variables to test homogeneity of pairs of distributions of $\Delta \log(A_t)$ from 15 countries. As in the simulation study, we compare our bootstrap method with the conventional Kolmogorov-Smirnov critical value that does not take into account for the estimation errors. The results are presented in Table 4. We can see that the conclusions of the tests are different for several cases (indicated by bold letters). Also those conclusions can be different in either ways. For reliable inference, it is critical to incorporate estimation errors for generated variables as in our bootstrap method.

[1]The capital input K_t is computed by the permanent inventory method, i.e., $K_t = I_t + (1-\delta)K_{t-1}$, where I_t and δ are gross investment and depreciation rate, respectively. We use the data on gross fixed capital formulation for I_t, and set the depreciation rate as $\delta = 0.05$. The initial capital stock K_0 is calculated by $K_0 = I_0/\delta$.

TABLE 4. Test results for (i) Bootstrap (ii) Conventional KS (R=reject, N=not reject)

	Austria		Belg.		Canada		Denm.		France		Germ.		Italy		Japan		Korea		Neth.		Norway		Sweden		UK		US	
	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
Australia	R	R	R	R	N	N	**R**	**N**	R	R	R	R	N	N	N	N	N	N	**N**	**R**	N	N	R	R	R	R	**R**	**N**
Austria			N	N	N	N	N	N	N	N	N	N	N	N	R	R	R	R	N	N	**R**	**N**	N	N	N	N	N	N
Belgium					N	N	N	N	N	N	N	N	N	N	R	R	N	N	N	N	N	N	N	N	N	N	N	N
Canada							N	N	**R**	**N**	R	R	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Denmark									N	N	N	N	N	N	**R**	**N**	R	R	N	N	**R**	**N**	N	N	N	N	N	N
France											N	N	N	N	R	R	R	R	N	N	R	R	N	N	N	N	N	N
Germany													**R**	**N**	R	R	R	R	N	N	R	R	N	N	N	N	**R**	**N**
Italy															N	N	**R**	**N**	N	N	N	N	N	N	N	N	N	N
Japan																	N	N	R	R	N	N	R	R	N	N	R	R
Korea																			R	R	**R**	**N**	R	R	R	R	R	R
Netherlands																					R	R	N	N	N	N	N	N
Norway																							R	R	N	N	N	N
Sweden																									N	N	N	N
UK																											N	N

References

[1] Berry, S., Levinsohn, J. and A. Pakes (1995) Automobile prices in market equilibrium, *Econometrica*, 63, 841-890.

[2] Hahn, J. and G. Ridder (2013) Asymptotic variance of semiparametric estimators with generated regressors, *Econometrica*, 81, 315-340.

[3] Hall, P. and J. L. Horowitz (1996) Bootstrap critical values for tests based on generalized-method-of-moments estimators, *Econometrica*, 64, 891-916.

[4] Lehmann, E. L. and J. P. Romano (2005) *Testing Statistical Hypotheses*, 3rd ed., Springer.

[5] Linton, O., Maasoumi, E. and Y.-J. Whang (2005) Consistent testing for stochastic dominance under general sampling schemes, *Review of Economic Studies*, 72, 735-765.

[6] Matsushita, Y. and T. Otsu (2019) Likelihood inference on semiparametric models with generated regressors, forthcoming in *Econometric Theory*.

[7] Newey, W. K. and D. L. McFadden (1994) Large sample estimation and hypothesis testing, in Engle, R. F. and McFadden, D. L. (eds.), *Handbook of Econometrics*, Ch. 36, vol. 4, Elsevier.

[8] Olley, S. and A. Pakes (1996) The dynamics of productivity in the telecommunications equipment industry, *Econometrica*, 64, 1263-1297.

[9] Pagan, A. (1984) Econometric issues in the analysis of regressions with generated regressors, *International Economic Review*, 25, 221-247.

[10] Solow, R. M. (1957) Technical change and the aggregate production function, *Review of Economics and Statistics*, 39, 312-320.

[11] Whang, Y.-J. (2001) Consistent specification testing for conditional moment restrictions, *Economics Letters*, 71, 299-306.

Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, UK.

Email address: t.otsu@lse.ac.uk

Research Center for Policy and Economy, Mitsubishi Research Institute, Inc., 10-3, Nagatacho 2-chome, Chiyoda-ku, Tokyo 100-8141, Japan.

Email address: go_taniguchi@mri.co.jp