An empirical examination of the effect of infrastructure on economic development: A large and heterogeneous panel data analysis *

Iziga I. Jude and Shingo TAKAGI

Hokkaido University

July 26, 2022

Abstract

Infrastructure drives economic development. This study investigates to what extent infrastructure and skilled labour affect aggregate output, by analysing large heterogeneous panel data of 130 countries over two decades. We implement an autoregressive distributed lag (ARDL) model to extract the long-run production technology relationship among economic growth, infrastructure, and skilled labour. The complementarity of skilled labour and infrastructure is conducive to skill-biased economic growth. Skill differences account for disparities among workers' wages worldwide, thereby widening inequalities in income and consequently, living standards. Contrary to previous studies, such as Calderón et al. [2015], that have used frameworks assuming production function homogeneity across countries, we propose a methodology to identify latent country groups based on the long-run production technology embedded in the ARDL model, using the estimation procedure of Liu et al. [2020]. We select the optimal number of groups by implementing a new information criterion under multiple nuisance parameters and estimate the coefficients of the production functions for each country group. Based on the complementarity estimates of country groups and the estimated country classifications, we find that the effects of infrastructure generated grouped-heterogeneity of growth span across countries in estimated production relationships.

JEL classification: C23, D24, I30

Keywords: identifying latent groups, heterogeneous large panel data, complementarity, infrastructure

*The authors acknowledge financial support from JST SPRING, Grant Number JPMJSP2119; JSPS KAKENHI, Grants-in-Aid for Scientific Research (C), Grant Number JP18K01603, JP21K01436; and Hokkaido University President's Fellowship.

1 Introduction

Infrastructure is one of the foundations of both social and economic life, as it contributes to economic growth and improvements in quality of life. From the road and water supply networks of Ancient Rome to the recent information and communications technology networks, infrastructure has been crucial to the maintenance and improvement of the productivity of commerce, agriculture, and industry. Therefore, many scholars and policymakers have studied the effects of infrastructure on economic growth and welfare, as well as its underlying mechanisms (Jimenez, 1995).The primary purpose of this study is to investigate to what extent infrastructure and other related inputs affect aggregate output, using large heterogeneous panel data of 130 countries over a period of two decades.

In previous theoretical and empirical studies, most macroeconometric models assume that physical capital (possibly including infrastructure) is the key input of a production function that generates aggregate output. This implies that public and private capital are perfect substitutes for each other. However, empirical evidence does exist against the 'perfect substitutability' theory, in that public and private capital can be complementary inputs of production (An et al., 2019). Therefore, infrastructure must be separated from other forms of physical capital while measuring returns.

Another important factor to consider regarding the specification of production function is the productive effect of capital-labour complementarity, which has received overwhelming attention in the literature (Krusell et al., 2000; Maliar et al., 2020; Na et al., 2020). Although some studies have found empirical evidence against capital-skill complementarity (Duffy et al., 2004), most evidence continues to favour the hypothesis of capital-skill complementarity (Correa et al., 2019; Tyers and Yang, 2000). Previous studies have also examined the theoretical and empirical implications of capital-skill complementarity on economic inequality, the wage-gap between skilled and nonskilled labour, and the productivity gaps among economic sectors (Krusell et al., 2000; Maliar et al., 2020).

Following the same reasoning of capital-skill complementarity, infrastructure (which is deemed as a part of physical capital) should be considered as a complementary or substitute input to skilled labour in macroeconomic production technology. Further, since building infrastructure implies vast expenditure, policymakers are often concerned, not only with its direct effects (such as those that are growth enhancing), but also its indirect effects (including income redistribution). The precise assessment and understanding of the contributions of infrastructure to the global economy are important issues in economics.

A topic that is yet to be fully explored in the assessment of the economic effects of infrastructure is the treatment of heterogeneity across units (Calderón and Servén, 2014). The effect of infrastructure on output may vary across units and over sample points due to various reasons, such as the heterogeneous features of infrastructure or production technology. Although several previous empirical studies have attempted to use panel data to control unobserved heterogeneity (represented by fixed effects), as in Eberhardt et al. [2013], the assumption of poolability (except the constant term across cross sectional units) is generally restrictive and may yield biased estimates. Resultantly, estimating the effect of infrastructure on the output of each country using time-series data may suffer from efficiency loss. To balance the possibility of bias with that of efficiency loss while considering heterogeneity, various econometric methods have been developed to endogenously find and classify latent groups in large panel data settings (Su et al., 2016; Liu et al., 2020). In this study, we classify 130 countries in our panel data into finite groups in terms of the features of the estimated production function. Subsequently, we present the group-wise production functions. The classification algorithm (which is detailed in Section 3) is based on the work of Liu et al. [2020].

Our main contribution to the literature is relating heterogeneity in parameters of production function with the classification of countries in terms of the marginal effect of production inputs, especially between infrastructure and skilled labour. This has important implications for domestic income inequalities across countries. Given the increase in infrastructure investment, income inequality increases in countries with infrastructure-skill complementarity in terms of wage increases for skilled labour. However, income inequality decreases in countries with substitutable inputs between infrastructure and skilled labour arising from the wage reduction for skilled labour. We examine the relationship between complementarity and income inequality based on our estimation results in Section 4.

This paper is organized as follows. Section 2 describes the dataset used in the empirical analysis. Section 3 presents the basic model to be estimated and the estimation procedures. Section 4 reports the estimation results and discusses their empirical implications.

2 Data

As previously mentioned, this study investigates to what extent infrastructure and skilled labour affect aggregate output, based on the analysis of large heterogeneous panel data of 130 countries over two decades (the sample period being 1990-2015). As summarized in Table 1, the data generated are defined as follows: output-side real gross domestic product (GDP) at chained purchasing power parities (PPPs) (in million 2011 US $), a measure of aggregate output shown as Y, and capital stock at constant 2011 national prices (in million 2011 US $) shown as K. The data on Y and K are generated from the Penn World Table (PWT) 9.1 (Feenstra et al., 2015). Data on average years of secondary schooling were collected from the Barro and Lee [2013] database. The average years of secondary schooling of the population is depicted as s, which represents skilled labour (labour effectiveness), that is defined by $S = \exp\{s\}$.

<<Table 1 is around here >>

Additional data include the total length of the road network (TROADS; in kilometres), obtained from the World Road Statistics; power generation capacity (EGC; in megawatts), collected from the United Nations Energy Statistics; and the total number of main telephone lines (MLINES) and labour force (LWDI), both collected from the World Development Indicator 2019. The variables enter production technologies as per person (divided by total labour force), except for the secondary variable, S. Infrastructure variables are each defined as per person before construction of the synthesis of infrastructure index.

The MLINES, EGC, and TROADS are used as to construct the infrastructure index. Following Calderón et al. [2015], a principal component method is applied to the three series to construct an index for infrastructure. The first main component of these three infrastructure availability services captures the synthesis of the infrastructure index[1].

3 Estimation and model selection

Our primary goal was to estimate the long-run relationship between the output and input variables, allowing for heterogeneity in coefficients. However, unrestricted heterogeneity in coefficients can cause under-identification or efficiency loss. When using the framework of the autoregressive distributed lag (ARDL) model to capture the dynamics of the variables in it, ARDL(P, Q), the model is

$$\sum_{p=0}^{P} \lambda_{i,p} y_{i,t-p} = \sum_{q=0}^{Q} \mathbf{f}'_{i,t-q} \boldsymbol{\beta}_{i,q} + \mu_i + u_{i,t},$$

where all coefficients are assumed to be heterogeneous across countries. A country-specific intercept term (i.e. a fixed effect term), μ_i, is included, and the error term $u_{i,t}$ is assumed to be idiosyncratic with a constant variance σ^2. The sample size of the cross-sectional dimension is shown as N and that of the time-series dimension as T. An

[1]The synthesized infrastructure index is given as $0.3654331 \log\left(\frac{\text{TROADS}}{\text{LWDI}}\right) + 0.3719091 \log\left(\frac{\text{EGC}}{\text{LWDI}}\right) + 0.2626578 \log\left(\frac{\text{MLINES}}{\text{LWDI}}\right)$, which explains approximately 84 %of the focussed dimensions' variation.

error-correction model (ECM) representation of the ARDL(P, Q)[2] [3] model is given as follows:

$$\Delta y_{it} = \phi_i \left(y_{i,t-1} - \mathbf{f}'_{i,t-1}\boldsymbol{\theta}_i \right) + \sum_{p=1}^{P-1} \lambda_{i,p} \Delta y_{i,t-p} + \sum_{q=0}^{Q-1} \underbrace{\Delta \mathbf{f}'_{i,t-q}}_{1\times k} \boldsymbol{\delta}_{i,q} + \mu_i + \epsilon_{it} \tag{1}$$

$$= \phi_i \xi_{i,t-1}(\boldsymbol{\theta}_i) + \begin{pmatrix} \Delta y_{i,t-1} & \cdots & \Delta y_{i,t-p+1} & \Delta \mathbf{f}'_{i,t} & \cdots & \Delta \mathbf{f}'_{i,t-q+1} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda}_i \\ \boldsymbol{\delta}_i \\ \mu_i \end{pmatrix} + \epsilon_{it},\ \xi_{i,t-1}(\boldsymbol{\theta}_i) \equiv y_{i,t-1} - \mathbf{f}'_{i,t-1}\boldsymbol{\theta}_i$$

$$= \phi_i \xi_{i,t-1}(\boldsymbol{\theta}_i) + \mathbf{W}_i \boldsymbol{\eta}_i + \epsilon_{it}. \tag{2}$$

The error correction term, $\xi_{i,t-1}(\boldsymbol{\theta}_i)$, captures the stable, long-run relationships between relevant variables. In this case, it is interpreted as the production function ($y_{i,t}$: output, $\mathbf{f}_{i,t} = (k_{i,t}, z_{i,t}, s_{i,t}, z_{i,t}s_{i,t})'$: physical capital, infrastructure, skilled labour, and the cross-term of the last two inputs. Further details are described in Section 2).

$$y_{i,t-1} - \mathbf{f}'_{i,t-1}\boldsymbol{\theta}_i = y_{i,t-1} - \beta_{k,i}k_{i,t-1} - \beta_{z,i}z_{i,t} - \beta_{s,i}s_{i,t} - \beta_{zs,i}z_{i,t}s_{i,t} \tag{3}$$

Although Calderón et al. [2015] assumes the homogeneity of the coefficients on $\mathbf{f}_{i,t}$, that is, $\boldsymbol{\theta}_i = \boldsymbol{\theta} = (\beta_k, \beta_z, \beta_s, \beta_{zs})$ for any i, this is a somewhat restrictive assumption. All countries would have similar production technologies and the differences would only be attributed to those in input quantity. Allowing for country-wise coefficients would result in a serious efficiency loss during estimation. Therefore, we allow for group-wise coefficients, but the membership is unrestricted and estimated from the data. In other words, we assume the presence of group-wise, long-run relationships for the country groups: $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_G$. Each country belongs to one of G groups,

$$\begin{aligned}
\xi_{i,t-1}(\boldsymbol{\theta}^{(1)}) &= y_{i,t-1} - \mathbf{f}'_{i,t-1}\boldsymbol{\theta}^{(1)} = y_{i,t-1} - \beta_k^{(1)}k_{i,t-1} - \beta_z^{(1)}z_{i,t} - \beta_s^{(1)}s_{i,t} - \beta_{zs}^{(1)}z_{i,t}s_{i,t}, \text{ if } i \in \mathcal{G}_1 \\
\xi_{i,t-1}(\boldsymbol{\theta}^{(2)}) &= y_{i,t-1} - \mathbf{f}'_{i,t-1}\boldsymbol{\theta}^{(2)} = y_{i,t-1} - \beta_k^{(2)}k_{i,t-1} - \beta_z^{(2)}z_{i,t} - \beta_s^{(2)}s_{i,t} - \beta_{zs}^{(2)}z_{i,t}s_{i,t}, \text{ if } i \in \mathcal{G}_2 \\
&\vdots \\
\xi_{i,t-1}(\boldsymbol{\theta}^{(G)}) &= y_{i,t-1} - \mathbf{f}'_{i,t-1}\boldsymbol{\theta}^{(G)} = y_{i,t-1} - \beta_k^{(G)}k_{i,t-1} - \beta_z^{(G)}z_{i,t} - \beta_s^{(G)}s_{i,t} - \beta_{zs}^{(G)}z_{i,t}s_{i,t}, \text{ if } i \in \mathcal{G}_G.
\end{aligned}$$

[2]The selection of P and Q is based on the method by Calderón et al., 2015, where the Akaike information criterion (AIC) is used to determine the lag lengths of the ARDL model. Both lag lengths are selected by country, using the AIC. The lengths are confined to 2. Due to the lag-length selection strategy, the length of the in-sample time period is 24, even though the dataset included 26 years' information.

[3]In addition, Calderón et al., 2015 apply the filtration to the original variables to remove the aggregate effects and time effects in the sample; they subtract the cross section means of the variables from the original variables for this purpose. We also adopt their filtration.

We also assume that the short-run dynamics, driven by coefficient parameters, ϕ_i and $\boldsymbol{\eta}_i$, are heterogeneous across countries. Country-specific, unrestricted coefficients are used in the model. Under the assumption of homogeneous long-run coefficients, Pesaran et al. [1999] propose the 'Pooled Mean Group' estimation method. In this study, we extend this model with homogeneous long-run coefficients among all countries to that with heterogeneous coefficients across a finite number of country groups. There are a small number of groups that show similar patterns of production function. Under this set up, there are $kG + N[(p-1) + kq + 3]$ parameters in the model $(\{\boldsymbol{\theta}^{(g)}\}_{g=1}^{G}, \{\boldsymbol{\lambda}_i, \boldsymbol{\delta}_i, \mu_i, \phi_i, \sigma_i^2\}_{i=1}^{N})$.

3.1 Estimation of grouped coefficients

Here, we introduce the group membership variable g_i, which takes a value of $\{1, 2, \ldots, G\}$ according to the group to which the country i belongs. The concentrated log-likelihood function of model (2), after concentrating the parameters $\{\boldsymbol{\lambda}_i, \boldsymbol{\delta}_i, \mu_i\}_{i=1}^{N}$, is given as follows: by using $\mathbf{Q}_{W,i} \equiv \mathbf{I}_T - \mathbf{W}_i \left(\mathbf{W}_i'\mathbf{W}_i\right)^{-1} \mathbf{W}_i'$: define as $\boldsymbol{\theta}_i \equiv \boldsymbol{\theta}^{(g_i)}$,

$$
\begin{aligned}
\ln L\left(\{\boldsymbol{\theta}^{(g)}\}_{g=1}^{G}, \{\phi_i, \sigma_i^2\}_{i=1}^{N}\right) &= \sum_{i=1}^{N}\sum_{t=1}^{T} \ell_{it}(\boldsymbol{\theta}_i, \phi_i, \sigma_i^2) \\
&= \sum_{i=1}^{N}\left\{\sum_{t=1}^{T} \frac{-1}{2}\left(\log(2\pi) + \log\sigma_i^2 + \frac{\left(\Delta\mathbf{y}_i - \phi_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}_i)\right)' \mathbf{Q}_{W,i}\left(\Delta\mathbf{y}_i - \phi_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}_i)\right)}{\sigma_i^2}\right)\right\} \\
&= -\frac{NT}{2}\log(2\pi) + \sum_{i=1}^{N}\left(-\frac{T}{2}\log\sigma_i^2 - \frac{\left(\Delta\mathbf{y}_i - \phi_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}_i)\right)' \mathbf{Q}_{W,i}\left(\Delta\mathbf{y}_i - \phi_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}_i)\right)}{2\sigma_i^2}\right).
\end{aligned}
$$

The estimation algorithm is given as follows (see Liu et al., 2020):

1. Given the number of groups G and an initial value of the long-run parameter[4] $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(G)})$ and $\boldsymbol{\theta}_i \equiv \boldsymbol{\theta}^{(g_i)}$, estimate the parameters of the short-run dynamics, $\left\{\hat{\phi}_i, \hat{\sigma}_i^2\right\}_{i=1}^{N}$, and the error correction term $\boldsymbol{\xi}_i(\boldsymbol{\theta}_i)$ as follows:

$$
\begin{aligned}
\hat{\phi}_i &= \left(\boldsymbol{\xi}_i(\boldsymbol{\theta}_i)'\mathbf{Q}_{W,i}\boldsymbol{\xi}_i(\boldsymbol{\theta}_i)\right)^{-1} \boldsymbol{\xi}_i(\boldsymbol{\theta}_i)'\mathbf{Q}_{W,i}\Delta\mathbf{y}_i \\
\hat{\sigma}_i^2 &= T^{-1}\left(\Delta\mathbf{y}_i - \phi_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}_i)\right)' \mathbf{Q}_{W,i}\left(\Delta\mathbf{y}_i - \phi_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}_i)\right) \\
\boldsymbol{\xi}_i(\boldsymbol{\theta}_i) &= \mathbf{y}_{i,-1} - \mathbf{F}_{i,-1}\boldsymbol{\theta}_i, \ \boldsymbol{\theta}_i \equiv \boldsymbol{\theta}(g_i)
\end{aligned}
$$

for each i, $1 \leq i \leq N$.

[4] The criterion function has multiple optima; the optimal value is sensitive to initial values. After attempting several initial parameters, we select the one reaching the maximum.

2. Select the optimal group for the i-th country as

$$g_i^* = \arg \min_{1 \le g \le G} \left\{ -\frac{T}{2} \log \hat{\sigma}_i^2 - \frac{\left(\Delta \mathbf{y}_i - \hat{\phi}_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}^{(g)})\right)' \mathbf{Q}_{W,i} \left(\Delta \mathbf{y}_i - \hat{\phi}_i \cdot \boldsymbol{\xi}_i(\boldsymbol{\theta}^{(g)})\right)}{2\hat{\sigma}_i^2} \right\}$$

for each i, $1 \le i \le N$. Then, we obtain the optimal membership as $\mathcal{G}^* = (g_1^*, g_2^*, \cdots, g_N^*)$ at this step.

3. Update the long-run parameter given the membership $\mathcal{G}^* = (g_1^*, g_2^*, \cdots, g_N^*)$ as ($\mathcal{G}_g^* \equiv \{\ i \mid g_i^* = g,\ 1 \le i \le N\,\}$),

$$\hat{\boldsymbol{\theta}}_1 = \left\{ \sum_{i \in \mathcal{G}_1^*} \frac{(\phi_i)^2}{\sigma_i^2} \cdot (\mathbf{F}_{i,-1})' \mathbf{Q}_{W,i} (\mathbf{F}_{i,-1}) \right\}^{-1} \sum_{i \in \mathcal{G}_1^*} \frac{\phi_i}{\sigma_i^2} \cdot (\mathbf{F}_{i,-1})' \mathbf{Q}_{W,i} (\Delta \mathbf{y}_i - \phi_i \cdot \mathbf{y}_{i,-1})$$
$$\hat{\boldsymbol{\theta}}_2 = \left\{ \sum_{i \in \mathcal{G}_2^*} \frac{(\phi_i)^2}{\sigma_i^2} \cdot (\mathbf{F}_{i,-1})' \mathbf{Q}_{W,i} (\mathbf{F}_{i,-1}) \right\}^{-1} \sum_{i \in \mathcal{G}_2^*} \frac{\phi_i}{\sigma_i^2} \cdot (\mathbf{F}_{i,-1})' \mathbf{Q}_{W,i} (\Delta \mathbf{y}_i - \phi_i \cdot \mathbf{y}_{i,-1})$$
$$\vdots$$
$$\hat{\boldsymbol{\theta}}_G = \left\{ \sum_{i \in \mathcal{G}_G^*} \frac{(\phi_i)^2}{\sigma_i^2} \cdot (\mathbf{F}_{i,-1})' \mathbf{Q}_{W,i} (\mathbf{F}_{i,-1}) \right\}^{-1} \sum_{i \in \mathcal{G}_G^*} \frac{\phi_i}{\sigma_i^2} \cdot (\mathbf{F}_{i,-1})' \mathbf{Q}_{W,i} (\Delta \mathbf{y}_i - \phi_i \cdot \mathbf{y}_{i,-1})$$

4. Repeat steps 1-3 until convergence

The asymptotic properties of the coefficient estimator and the group membership were established by Liu et al. [2020]. Details of the results are shown in the Appendix.

3.2 Model selection

Information criteria for model selection under the presence of incidental parameters are proposed in Lee and Phillips [2015], who establish the conditions for the consistency of model selection, where the selected model is asymptotically true. We slightly modify their Bayesian-like information criterion, using the modified profile likelihood contribution, $\ell_{it}(\boldsymbol{\theta}_i, \hat{\boldsymbol{\alpha}}_i(\boldsymbol{\theta}_i))$ and the correction term. The information criterion is defined as

$$\mathrm{IC}(g) = -\frac{2}{NT} \sum_{i=1}^{N} \left\{ \sum_{t=1}^{T} \ell_{it}(\boldsymbol{\theta}^{(g_i)}, \hat{\boldsymbol{\alpha}}_i(\boldsymbol{\theta}_i)) - M_i\left(\boldsymbol{\theta}^{(g_i)}\right) \right\} + \frac{h(NT)}{NT} \times gK,\ 1 \le g_i \le g, \tag{4}$$
$$M_i(\boldsymbol{\theta}) = \frac{1}{2} \left\{ -\mathbb{E}_T \left[\frac{\partial^2 \ell_{it}(\boldsymbol{\theta}, \hat{\boldsymbol{\alpha}}_i(\boldsymbol{\theta}))}{\partial \boldsymbol{\alpha}_i \partial \boldsymbol{\alpha}_i'} \right] \right\}^{-1} \left\{ \mathbb{E}_T \left[\frac{\partial \ell_{it}(\boldsymbol{\theta}, \hat{\boldsymbol{\alpha}}_i(\boldsymbol{\theta}))}{\partial \boldsymbol{\alpha}_i} \frac{\partial \ell_{is}(\boldsymbol{\theta}, \hat{\boldsymbol{\alpha}}_i(\boldsymbol{\theta}))}{\partial \boldsymbol{\alpha}_i'} \right] \right\}.$$

The requirement for the consistency is just that $h(NT)$ is a non-decreasing function of NT. After attempting some candidates, we ultimately choose $h(NT) = (NT)^{3/8}$ since $\ln(NT)$ is too loose and $(NT)^{1/2}$ is too severe to pick up moderate group sizes. The results of model selection by the information criterion, defined in (4), and by the criterion proposed by Liu et al. [2020], are shown in Table 2. Both information criteria show that the optimally selected model is the one with five groups ($G = 5$). In the estimation results section, we use those from the model with $G = 4$, $G = 5$ and $G = 6$ to reduce the risk of false selection. Further research examining the validity of our choice for the information criterion is ongoing.

<<**Table 2 is around here** >>

The classification result of the model with 5 groups and the coefficient estimates of the models with 4, 5, and 6 groups are summarized in Table 3 and Table 4. The results and interpretations are discussed in the following section.

<<**Table 3 is around here** >>

<<**Table 4 is around here** >>

4 Results and interpretation

The model we investigate is the production function of the following form,

$$Y = K^{\beta_k} Z^{\beta_z} S^{\beta_s} \exp\{\beta_{zs} \log Z \log S\} \tag{5}$$

Where Y is the output per worker, K is the physical capital stock per worker, Z is the infrastructure service per worker (the geometric average of telecommunication stock, road stock, and electricity-generating stock, defined in the footnote of page 4), and S is the skilled labour (defined as the exponential of the average years of secondary education in the population). This is an extended version of a Cobb-Douglas production function with an interaction term (Na et al. [2020]). In the logarithm form, the estimated model is given as a linear-in-parameters model with an interaction term,

$$\log Y = \beta_k \log K + \beta_z \log Z + \beta_s \log S + \beta_{zs} \log Z \log S.$$

From this setup, the marginal product of infrastructure is

$$\frac{\partial Y}{\partial Z} = (\beta_z + \beta_{zs} \log S)\, \frac{Y}{Z},$$

and the marginal product of the skilled labour is

$$\frac{\partial Y}{\partial S} = (\beta_s + \beta_{zs} \log Z)\, \frac{Y}{S}.$$

The term $(\beta_z + \beta_{zs} \log S)$ represents the total contribution of infrastructure to aggregate output, and $(\beta_s + \beta_{zs} \log Z)$ is similarly defined for skilled labour. That is, they are interpreted as the effect of each input on the output.

The cross derivative of Y with respect to Z and S is given as:

$$\frac{\partial^2 Y}{\partial S \partial Z} = \{\beta_{zs} + (\beta_z + \beta_{zs} \log S)(\beta_s + \beta_{zs} \log Z)\}\, \frac{Y}{ZS}.$$

Therefore, the sign of the cross derivative is related to the concept of complementarity or substitutability in the definition by Milgrom and Roberts [1990], which is determined by $\beta_s + \beta_{zs} \log Z$, $\beta_z + \beta_{zs} \log S$, and β_{zs}. When infrastructure and skilled labour are complementary, as infrastructure investment increases, the marginal product of skilled labour rises. Subsequently, wages paid to skilled labour increase, which will widen the wage gap between skilled and nonskilled workers. This wage increase would manifest in the Gini coefficient of the economy, leading to a rise in the coefficient.

Conversely, when infrastructure and skilled labour are substitutable inputs, increases in infrastructure will reduce the marginal product of skilled labour and the payment to the latter will decrease. This interpretation can help us understand the narrowing mechanism of the income gap. Thus, the substitution between S and Z could lead to decrease in the Gini coefficient of the economy.

4.1 Classification results

In this subsection, we discuss classification results in terms of the signs of the estimated marginal product of skilled labour ($\partial Y/\partial S$), infrastructure ($\partial Y/\partial Z$), and the cross derivative of the production function with respect to both inputs ($\partial^2 Y/\partial S\partial Z$). In Table 5, we show the estimated signs of the groups with the feature of substitution between S and Z ($\partial^2 Y/\partial S\partial Z < 0$). The numbers between the parentheses denote the positive and negative estimates of the parameters for the corresponding model. For example, **Group1** of the model with the number of groups $G = 4$ contains 40 countries. The sample period used for the estimation is 24 years; therefore, there are $960 = 40 \times 24$ observations in this category. The table shows (960,0), (0,960), and (0,960) for $\partial Y/\partial S$, $\partial Y/\partial Z$, and$\partial^2 Y/\partial S\partial Z$ respectively, which reveal that all estimated $\partial Y/\partial S$ in this category are positive. Likewise, all estimated $\partial Y/\partial Z$ and $\partial^2 Y/\partial S\partial Z$ are negative. This sign pattern is common to one of the models with $G = 5$ and $G = 6$. Therefore, **Group1** is characterized by a positive marginal product of skilled labour, a negative marginal product of infrastructure, and a negative cross derivative with respect to both inputs. **Group6** also has the same sign pattern as **Group1**: this is closely related to the grouping process; a large percentage of the countries in **Group6** is separated from **Group1**. **Group4** also has negative cross-derivative estimates. However, the sign pattern of marginal products is the opposite, namely, the one of skilled labour is negative and that of infrastructure is positive. Interpretations of these results and their statistical significance are discussed in the following subsection.

<<**Table 5 is around here** >>

In Table 6, **Group2** and **Group3** show positive cross-derivative estimates, while **Group5** shows a mixed-sign result of cross-derivative estimates.[5] **Group2** is clearly characterized by a negative sign for the marginal product of skilled labour and a positive sign for that of infrastructure. **Group3** and **Group5** have negative marginal products.

<<**Table 6 is around here** >>

From these tables, it is clear that all the groups are classified in terms of the signs of marginal products and cross derivatives, except for the cross-derivative estimates of **Group5**. We emphasize that our grouping method can find and classify the different coefficient patterns in the macro production function.

Accordingly, Canada and Netherlands in **Group2**, Germany and the United Kingdom in **Group3**, and other countries in the respective groups (see Table 3) clearly show skill-infrastructure complementarity. Some countries in **Group5** that Australia and New Zealand fell into (see Table 3 again), similarly had a complementary production technology. These results are consistent with those in Taniguchi and Yamada [2022], Michaels et al. [2014], and Krusell et al. [2000]. Given that the above listed countries are mostly Organisation for Economic Cooperation

[5]The cross-derivative estimates of **Group5** are concentrated on just three points: 50% (42%) of estimates are just approximately zero, 19% (25%) are approximately 0.44, and 31% (33%) are approximately 0.78 for the model with five (six) groups: the feature of this group is that almost half of estimates are approximately zero, and the remaining are positive.

and Development (OECD) countries, our results clearly expand the work of Taniguchi and Yamada [2022]. This suggests that infrastructure-induced technical advancement in the countries raises skilled-labour contributions to output, favouring skill premium. It appears that access to adequate infrastructure enables skilled workers to create new production formula and disseminate the innovative approaches across enterprises in the countries. Additionally, skilled workers diffuse the new production processes in industries that adopt them. These aspects increase the marginal productivity of skilled labour and its remunerations in the form of wages.

Importantly, Burundi, Uganda, and many other African countries fell into **Group1** that is characterized by substitutable skill-infrastructure production technology. Bangladesh and Mongolia in Asia, Romania in Europe, and El Salvador in America also fell into this group. Increasing infrastructure investment in these counties reduces skill premium, implying that it decreases marginal product and wages of the skilled workers. Countries in **Group6**, which are a subset of those in **Group1**, have a similar outcome of technical changes. In Japan, the United States, and some other countries in **Group4**, increasing infrastructure investment and skill acquisitions leave skill premium unchanged in the statistical sense. Overall, infrastructure and skill as complementary production technologies enhance the productivity of skilled labour and increase its wages relative to that of unskilled workers. The magnitude and statistical significance of the production factors are further discussed and additional countries with complementary technologies are enlisted.

4.2 Negative marginal product of infrastructure and positive marginal product of skilled labour

For the first country group (**Group1**), the direct effect of infrastructure on output per worker is negative and significant. The point estimate in Table 4 is -0.296 (s.e.=0.04), -0.336 (s.e.=0.04), and -0.424 (s.e.=0.05) when $G = 4$, $G = 5$, and $G = 6$, respectively. This negative effect can be due to certain data aggregations and the network characteristic of infrastructure. For example, infrastructure (e.g. transport) investment has an output reallocation effect (Melo et al., 2013). By classifying countries into groups, the effect can be negative if infrastructure redistributes output to the winning locations. With network externalities, nonlinearity in infrastructure-output relations implies that a positive effect is feasible when a critical network mass is reached (e.g. universal penetration rate for telephones).

The negative point estimate of infrastructure does not always indicate that infrastructure is irrelevant. Since infrastructure capital is already included in the physical capital stock, this implies that infrastructure has the normal productivity effect of capital as a whole. This explains the very large effect we find for physical capital. We find that the elasticity of output with respect to the physical capital stock amount to approximately 0.80, 0.86, and 0.99, holding constant infrastructure and skilled labour. That is, we find a large effect from increasing the physical capital stock and removing an equal amount of investment in infrastructure capital. This suggests that there is large network externality to physical capital.

The estimated output effects of skilled labour, $\log S$, are 0.172 (s.e.=0.02), 0.163 (s.e.=0.03), and 0.124 (s.e=0.03) when $G = 4$, $G = 5$, and $G = 6$, respectively. This implies that if other forces are held constant, the (average) increases in output per worker resulting from a 1% increase in skilled labour service are roughly 17.2%, 16.3%, and 12.4%, for $G = 4$, $G = 5$, and $G = 6$, respectively. For the coefficient of the interaction term between infrastructure and skilled labour, $\log Z \log S$, the estimated effect is negative and significant, except for the estimate with six groups ($G = 6$). The point estimates of the coefficient on the interaction term are -0.042 (s.e.=0.01), -0.038 (s.e.=0.01), and -0.001 (s.e.=0.02), for $G = 4$, $G = 5$, and $G = 6$, respectively. The estimated total effects (the marginal products) are depicted in Figure 1.

<<**Figure 1 is around here** >>

These estimates reveal the substitutability between skilled labour and infrastructure and could result in a reduction in income inequality. In this case, an increase in infrastructure causes income to flow from skilled to nonskilled labour, indicating that the production sector using the latter requires relatively more intensive infrastructure services. With increases in infrastructure, the wages of skilled labour decline and those of nonskilled labour increase, reducing wage inequality. As evidence of this income gap reduction effect, we refer to Figure 2, which contains Gini coefficient estimates of the last 10 years (2006-2015) and those of the first 10 years (1990-1999) of the sample period with a 45-degree line.[6] Points above (below) 45-degree lines indicate that the Gini coefficients increased (decreased) in the last 10 years of the sample period. Figure 2 shows that the Gini coefficient estimates of **Group2**, **Group3**, and **Group5** (**Group1**, **Group4**, and **Group6**) are relatively increasing (or decreasing, as the case may be) at the end of the sample period.

<<**Figure 2 is around here** >>

The panels in the middle columns of Figure 2 are based on the country classification using the five-groups model,

[6]The Gini coefficient estimates are calculated from Top 10% share, Bottom 50% share, and Top 1% share, taken from the website of the World Inequality Database (https://wid.world/). Except Fiji and Balbados, all countries and almost all sample periods are covered by available Gini coefficient estimates.

and the top panel is the one for **Group1**. The ratio of the points (51.5%) below the 45-degree line is relatively larger than that (48.5%) above the line, which implies that a relatively large number of countries exhibit downward trends in their Gini coefficients over the sample period. Although some large deviations above the line are found for some **Group1** countries with lower Gini coefficients at the beginning of the sample period, a relatively large number of countries are consistent with the above income gap reduction reasoning. Since some countries in **Group1** with the five-groups model are classified into **Group6** with the six-groups model, the income gap reduction effects in **Group1** are mitigated.

The sign pattern of the marginal products of Z and S in **Group6** is similar to that in **Group1**, but the significance and effects of Z are less than those in **Group1** (see Figure 1). The substitution effect between Z and S (the cross derivative estimate) in **Group6** is also negative but close to zero (see the note below Table 5). This similarity in estimates between **Group1** and **Group6** is partly due to the over-specification of the number of groups. Information criteria lead to five being selected as the optimal number of groups. As a result, adding another group (the sixth group) yields estimates like those of the first group, suggesting that the sixth is a subgroup of the first one, as predicted when the consistency of the grouping was established by Liu et al. [2020].

4.3 Positive marginal product of infrastructure, Negative marginal product of skilled labour

Next, we consider estimation results of the second country group (**Group2**). Infrastructure has a positive and significant direct effect on output per worker. From Table 4, the estimated coefficients are 0.578 (s.e.=0.05), 0.617 (s.e.=0.05), and 0.611 (s.e.=0.05), when $G = 4$, $G = 5$, and $G = 6$ respectively. This indicates that infrastructure is an important and robust production factor across models with $G = 4, G = 5$, and $G = 6$. As an example of growth-promoting infrastructure, we can consider road infrastructure, which can link markets and cause an increase in competition. In addition, communication systems can increase the rate of diffusion of technology, increasing output. The increases in output with respect to a 10% increase in infrastructure investments, on average, are 5.8%, 6.2%, and 6.1% for $G = 4$, $G = 5$, and $G = 6$, respectively.

For the coefficient of the interaction term $\log Z \log S$, the point estimates are 0.411 (s.e.=0.04), 0.431 (s.e.=0.04), and 0.434 (s.e.=0.04) for $G = 4, G = 5$, and $G = 6$, respectively. This indicates that, aside from the direct increasing effect on output, an increase in the volume of infrastructure services raises output indirectly by *crowding-in* skilled labour leading to the consequent rise in the marginal products of skilled labour (see the top panels in Figure 3 for the total effect of infrastructure). Infrastructure provision can improve health and education outcomes and enhance skilled labour. Similarly, improved access to electricity may raise educational attainment and reduce the cost of skill acquisition. Generally, infrastructure provision could increase overall economic output and performance.

<<**Figure 3 is around here** >>

Thus, infrastructure raises the marginal product and remuneration of skilled labour. Income flows from non-skilled labour to skilled labour, increasing their wage premium, and consequently, the wage gap between the two. As evidence of the mechanism, the Gini coefficients in this category tend to be upward: the relative frequencies above the 45-degree line (increase in Gini coefficients) in the plot of **Group2** are larger than those below the line (decrease) in Figure 2.

If other forces are held constant, skilled labour earns roughly 4% more than nonskilled labour for every 10% increase in infrastructure provision. The total contribution of infrastructure to growth and development is calculated as its direct marginal product in addition to its indirect marginal growth contribution through the channel of skilled labour. This depends largely on how efficiently skilled labour uses infrastructure in the production process.

Skilled labour is estimated to have a negative and significant relationship with productivity performance. The estimated coefficients are -0.130 (s.e.=0.03), -0.130 (s.e.=0.03), and -0.134 (s.e.=0.03) for $G = 4, G = 5$, and $G = 6$, respectively. Although part of the estimated marginal product of skilled labour is distributed at approximately zero (see the note below Table 5), the negative estimates show that the suggestion to invest in schooling to raise output does not hold in the data. Pritchett [2001] also pointed out the case where education for skill-acquisition is not effective.

The second group (**Group2**) and fourth group (**Group4**) are similar in terms of the signs of the marginal product of infrastructure and skilled labour; the former is positive and significant, and the latter is negative and less significant (see Table 4, the top panels of Figure 3 for **Group2**, and the middle panels of Figure 1 for **Group4**). The extent of the direct effect of both S and Z on the output (measured by β_s and β_z) in **Group2** is larger than that in **Group4**, and the total effect of infrastructure on output is generally positive, whereas the total effect of skilled labour is generally close to zero. The cross derivatives in **Group4** tend to be negative but they are also distributed around zero (see the note below Table 5), except for a few points. Interestingly, these negative cross derivatives in **Group4** might lead a relatively large number of member countries to smaller Gini coefficients in the last period of the sample (see the panels of **Group 4** in Figure 2). The second and fourth groups are similar in terms of the marginal products, but the substitutability and the complementarity between S and Z result in the different patterns of income distribution.

4.3.1 Other cases

Finally, the third group (**Group3**) and fifth group (**Group5**) are similar in the sign patterns of coefficient estimates; both Z and S are negative and significant (see Figure 3) and the coefficient on the interaction term is positive (or negative but almost zero in **Group5**, see the note below Table 6) and significant: the difference between the two groups is in the relative magnitude of the coefficient on the interaction term. It is difficult to explain why both marginal products were negative, but large and positive cross-derivative estimates highlight the strong complementarity of infrastructure and skilled labour and the importance of their joint use as a determinant of output in these categories.

5 Conclusion

In this study, we examined the effect of infrastructure on economic development. We estimated an extended Cobb-Douglas production technology embedded in an autoregressive distributed lag (ARDL) model with grouped coefficients and possible inputs' complementarity. Nuisance parameters were controlled for, and an asymptotically optimal model was selected using the Bayesian-like information criterion, which is based on a modified profile likelihood. We found that the effects of infrastructure generated grouped-heterogeneity of growth across countries in the estimated production relationships.

Another interesting finding is that our method is stable and consistent in classifying countries into groups. While some estimated groups were only subsets of true groups, none were a mixture of elements from multiple true groups. For example, **Group1** exhibits a positive marginal product of skilled labour, a negative marginal product of infrastructure, and a negative cross derivative with respect to both inputs. **Group6** also exhibits the same sign pattern as **Group1**, suggesting that a large part of country members in **Group6** is separated from **Group1**. Since infrastructure capital is somewhat included in the physical capital stock, the negative marginal product of infrastructure suggests that it had the normal productivity effect of capital as a whole, implying a large network externality to physical capital stock. Similarly, the negative cross derivative of inputs suggests that infrastructure and skilled labour are close substitutes in the production process. This has a reduction implication to income inequality of countries in **Group1** and **Group6** in terms of wage redistribution from skilled labour to nonskilled one.

Similarly, **Group2** has a negative marginal product of skilled labour and positive marginal product of infrastructure and cross derivative. The positive cross derivative of inputs indicates that infrastructure crowds in skilled labour in production, leading to a rise in its marginal products. Marginal products, which equal rewards to inputs, imply that infrastructure raises the wages paid to skilled labour. Thus, income flows from nonskilled labour to skilled labour, thus increasing their wage premium and the wage gap between them. **Group2** and **Group4** are similar in terms of the signs of the marginal product of infrastructure and skilled labour, but are different in terms of the complementarity between infrastructure and skilled labour, which lead to different patterns of income distribution.

Likewise, the coefficient estimates of **Group3** and **Group5** exhibit similar sign patterns. One finding that is difficult to explain is why both **Group3** and **Group5** have negative marginal products of the two inputs. However, their large and positive cross-derivative estimates highlight the strong complementarity between infrastructure and skilled labour, as well as the importance of their joint use as a determinant of output. Our model could partially explain these counter-intuitive findings, but it might be too simple to capture the entire features of the macro production function across countries. To obtain estimation results that are more intuitively appealing, more elaborate input variables may be required (Duffy et al., 2004), or more refined specifications of production function, such as nested constant-elasticity-of-substitution production functions (Sato, 1967) or a semiparametric parsimonious flexible functional form (Coppejans, 2003), may be required. Such extensions will comprise our future research.

References

Zidong An, Alvar Kangur, and Chris Papageorgiou. On the Substitution of Private and Public Capital in Production. IMF Working Papers 2019/232, International Monetary Fund, November 2019.

Robert J. Barro and Jong Wha Lee. A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104(C):184–198, 2013.

César Calderón and Luis Servén. Infrastructure, growth, and inequality : an overview. Policy Research Working Paper Series 7034, The World Bank, 2014.

César Calderón, Enrique Moral-Benito, and Luis Servén. Is infrastructure capital productive? a dynamic heterogeneous approach. *Journal of Applied Econometrics*, 30(2):177–198, 2015.

Mark Coppejans. Flexible but parsimonious demand designs: The case of gasoline. *The Review of Economics and Statistics*, 85(3):680–692, 2003.

Juan A. Correa, Miguel Lorca, and Francisco Parro. Capital-skill complementarity: Does capital composition matter? *The Scandinavian Journal of Economics*, 121(1):89–116, 2019.

John Duffy, Chris Papageorgiou, and Fidel Perez-Sebastian. Capital-Skill Complementarity? Evidence from a Panel of Countries. *The Review of Economics and Statistics*, 86(1):327–344, 2004.

Markus Eberhardt, Christian Helmers, and Hubert Strauss. Do spillovers matter when estimaing provate returns to r&d? *The Review of Economics and Statistics*, 95(2):436–448, 2013.

Robert C. Feenstra, Robert Inklaar, and Marcel P. Timmer. The next generation of the penn world table. *American Economic Review*, 105(10):3150–3182, October 2015.

Jinyong Hahn and Guido Kuersteiner. Bias reduction for dynamic nonlinear panel data models with fixed effects. *Econometric Theory*, 27(6):1152–1191, 2011.

Emmanuel Jimenez. Chapter 43 human and physical infrastructure: Public investment and pricing policies in developing countries. volume 3 of *Handbook of Development Economics*, pages 2773–2843. Elsevier, 1995.

Per Krusell, Lee E. Ohanian, José-Victor Rios-Rull, and Giovanni L. Violante. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5):1029–1053, 2000.

Yoonseok Lee and Peter C. B. Phillips. Model selection in the presence of incidental parameters. *Journal of Econometrics*, 188:474–489, 10 2015.

Ruiqi Liu, Zuofeng Shang, Yonghui Zhang, and Qiankun Zhou. Identification and estimation in panel models with overspecified number of groups. *Journal of Econometrics*, 215(2):574–590, 2020.

Lilia Maliar, Serguei Maliar, and Inna Tsener. Capital-Skill Complementarity and Inequality: Twenty Years After. CEPR Discussion Papers 15228, C.E.P.R. Discussion Papers, August 2020.

Patricia C. Melo, Daniel J. Graham, and Ruben Brage-Ardao. The productivity of transport infrastructure investment: A meta-analysis of empirical evidence. *Regional Science and Urban Economics*, 43(5):695–706, 2013.

Guy Michaels, Ashwini Natraj, and John Van Reenen. Has ict polarized skill demand? evidence from eleven countries over twenty-five years. *The Review of Economics and Statistics*, pages 60–77, 2014.

Paul Milgrom and John Roberts. The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review*, 80(3):511–528, 1990.

Kyoung Y. Na, Dong H. Kim, Byoung G. Park, Sang W. Yoon, and Chang-Ho Yoon. Ict and transport infrastructure development: an empirical analysis of complementarity. *Applied Economics*, 52(2):195–211, 2020.

M. Hashem Pesaran, Yongcheol Shin, and Ron P. Smith. Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, 94(446):621–634, 1999.

Lant Pritchett. Where has all the education gone? *The World Bank Economic Review*, 15(3):367–391, 2001.

Kazuo Sato. A two-level constant-elasticity-of-substitution production function. *The Review of Economic Studies*, 34(2):201–218, 1967.

Liangjun Su, Zhentao Shi, and Peter C. B. Phillips. Identifying latent structures in panel data. *Econometrica*, 84 (6):2215–2264, 2016.

Hiroya Taniguchi and Ken Yamada. Ict capital–skill complementarity and wage inequality: Evidence from oecd countries. *Labour Economics*, 76:102151, 2022.

Rod Tyers and Yongzheng Yang. Capital-skill complementarity and wage outcomes following technological change in a global model. *Oxford Review of Economic Policy*, 16(3):23–41, 2000.

Appendix Asymptotic properties of the estimator

We introduce some notations frequently used in large panel data literature, including Hahn and Kuersteiner [2011]. The short-run parameters (nuisance parameters) are shown as $\boldsymbol{\alpha}_i \equiv (\phi_i, \sigma_i^2)'$ and derivatives of the likelihood contribution with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$ as (note that $\mathbb{E}_T[X_t] = T^{-1}\sum_{t=1}^{T} X_t$ and $\bar{\mathbb{E}}_T\left[X_t Y_s\right]$ is an estimate of the long-run covariance matrix between X_t and Y_s):

$$U_{it} = \frac{\partial \ell_{it}(\boldsymbol{\theta}_i, \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\theta}_i},\ V_{it} = \frac{\partial \ell_{it}(\boldsymbol{\theta}_i, \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\alpha}_i},$$

$$U_{\alpha,it} = \frac{\partial^2 \ell_{it}(\boldsymbol{\theta}_i, \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\alpha}_i'}, U_{\alpha\alpha,it} = \frac{\partial^3 \ell_{it}(\boldsymbol{\theta}_i, \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\theta}_i (\partial \boldsymbol{\alpha}_i \otimes \partial \boldsymbol{\alpha}_i)'}, V_{\alpha,it} = \frac{\partial^2 \ell_{it}(\boldsymbol{\theta}_i, \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\alpha}_i \partial \boldsymbol{\alpha}_i'}, V_{\alpha\alpha,it} = \frac{\partial^3 \ell_{it}(\boldsymbol{\theta}_i, \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\alpha}_i (\partial \boldsymbol{\alpha}_i \otimes \partial \boldsymbol{\alpha}_i)'}$$

where $(\partial\boldsymbol{\alpha}_i \otimes \partial\boldsymbol{\alpha}_i)' = \left((\partial\phi_i, \partial\sigma_i^2)' \otimes (\partial\phi_i, \partial\sigma_i^2)'\right)' = \left((\partial\phi_i)^2, \partial\phi_i\partial\sigma_i^2, \partial\sigma_i^2\partial\phi_i, (\partial\sigma_i^2)^2\right)$. and

$$\boldsymbol{\psi}_{it} = \left\{\mathbb{E}_T\left[V_{\alpha,it}\right]\right\}^{-1} V_{it},\ \tilde{U}_{it} = U_{it} - \Xi_i V_{it},\ \tilde{U}_{\alpha,it} = U_{\alpha,it} - \Xi_i V_{\alpha,it},\ \tilde{U}_{\alpha\alpha,it} = U_{\alpha\alpha,it} - \Xi_i V_{\alpha\alpha,it}$$

where $\Xi_i = \mathbb{E}_T\left[U_{\alpha,it}\right]\left\{\mathbb{E}_T\left[V_{\alpha,it}\right]\right\}^{-1}$. Liu et al. [2020] proved that the asymptotic distribution of the estimator is given as follows:

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}^{(g)} - \boldsymbol{\theta}^{(g)}\right) \xrightarrow{D} N\left(\kappa\mathcal{I}_g^{-1}\boldsymbol{d}_g, \frac{1}{\pi_g}\mathcal{I}_g^{-1}\mathcal{D}_g\mathcal{I}_g^{-1}\right),\quad g = 1, 2, \ldots, G\,(\geq G_0),$$

where $\kappa = \lim_{N,T\to\infty}\sqrt{\frac{N}{T}}$, $\pi_g = \lim_{N,T\to\infty}\frac{N_g}{N}$, and N_g is the number of countries in $\mathcal{G}_g$,

$$\begin{aligned}
\mathcal{I}_g &= \frac{1}{N_g}\sum_{i\in\mathcal{G}_g}\left(-\mathbb{E}_T\left[\frac{\partial U_{it}}{\partial\boldsymbol{\theta}_i'}\right] + \mathbb{E}_T\left[\frac{\partial V_{it}'}{\partial\boldsymbol{\theta}_i}\right]\left\{\mathbb{E}_T\left[\frac{\partial V_{it}}{\partial\boldsymbol{\alpha}_i}\right]\right\}^{-1}\cdot\mathbb{E}_T\left[\frac{\partial V_{it}}{\partial\boldsymbol{\theta}_i'}\right]\right),\\
\mathcal{D}_g &= \frac{1}{N_g}\sum_{i\in\mathcal{G}_g}\bar{\mathbb{E}}_T\left[\tilde{U}_{it}\tilde{U}_{is}'\right],\\
\mathbf{d}_g &= \frac{1}{N_g}\sum_{i\in\mathcal{G}_g}\left\{\bar{\mathbb{E}}_T\left[\tilde{U}_{\alpha,it}\boldsymbol{\psi}_{is}\right] + \frac{1}{2}\mathbb{E}_T\left[\tilde{U}_{\alpha\alpha,it}\right]\mathrm{vec}\left(\bar{\mathbb{E}}_T\left[\boldsymbol{\psi}_{it}\boldsymbol{\psi}_{is}'\right]\right)\right\}.
\end{aligned}$$

The bias-corrected estimator is defined as $\tilde{\boldsymbol{\theta}}^{(g)} = \hat{\boldsymbol{\theta}}^{(g)} - T^{-1}\mathcal{I}_g^{-1}\boldsymbol{d}_g$, which we report as the estimation results.

Liu et al. [2020] established not only the consistency and the asymptotic normality of the long-run coefficient parameter, but also the consistency of the group classification. This consistency implies that all estimated groups are surely included in a certain true group if their numbers in the estimated mode (G) are greater than or equal to the true number of groups (G_0): $G \geq G_0$. Some estimated groups are only subsets of true groups if $G \geq G_0$, but the appropriate combination of estimated groups can reproduce the true groups with probability one as the sample size goes to infinity.

The important point is that, asymptotically, none of the estimated groups become a mixture of elements from multiple true groups. Of course, when $G = G_0$, the estimated group memberships are expected to be identical to true group memberships. In this sense, the selection of the number of groups is especially important in our research.

Tables and Figures

Figure 1: Marginal Products of S and Z: Group1, Group4, and Group6

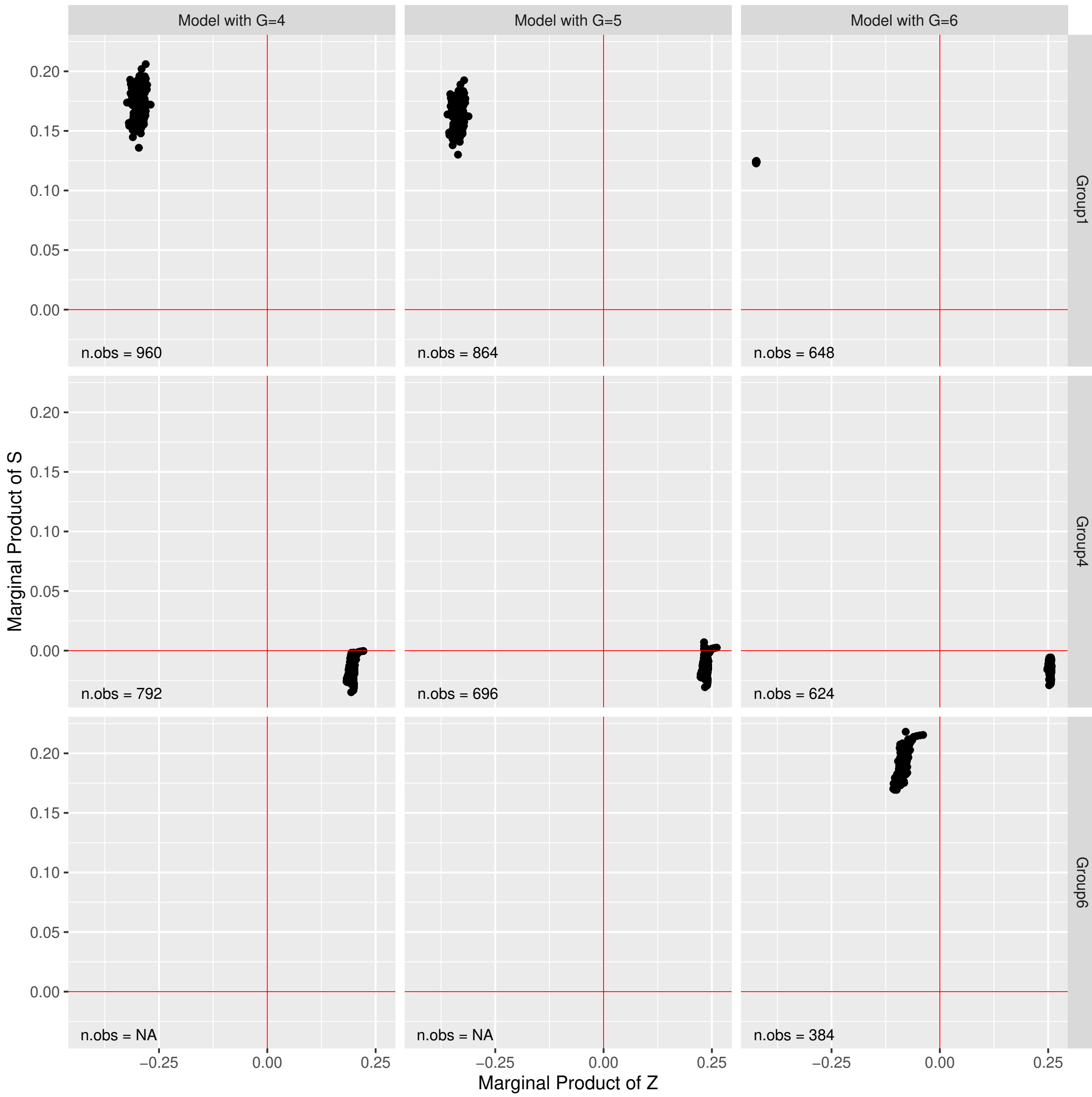
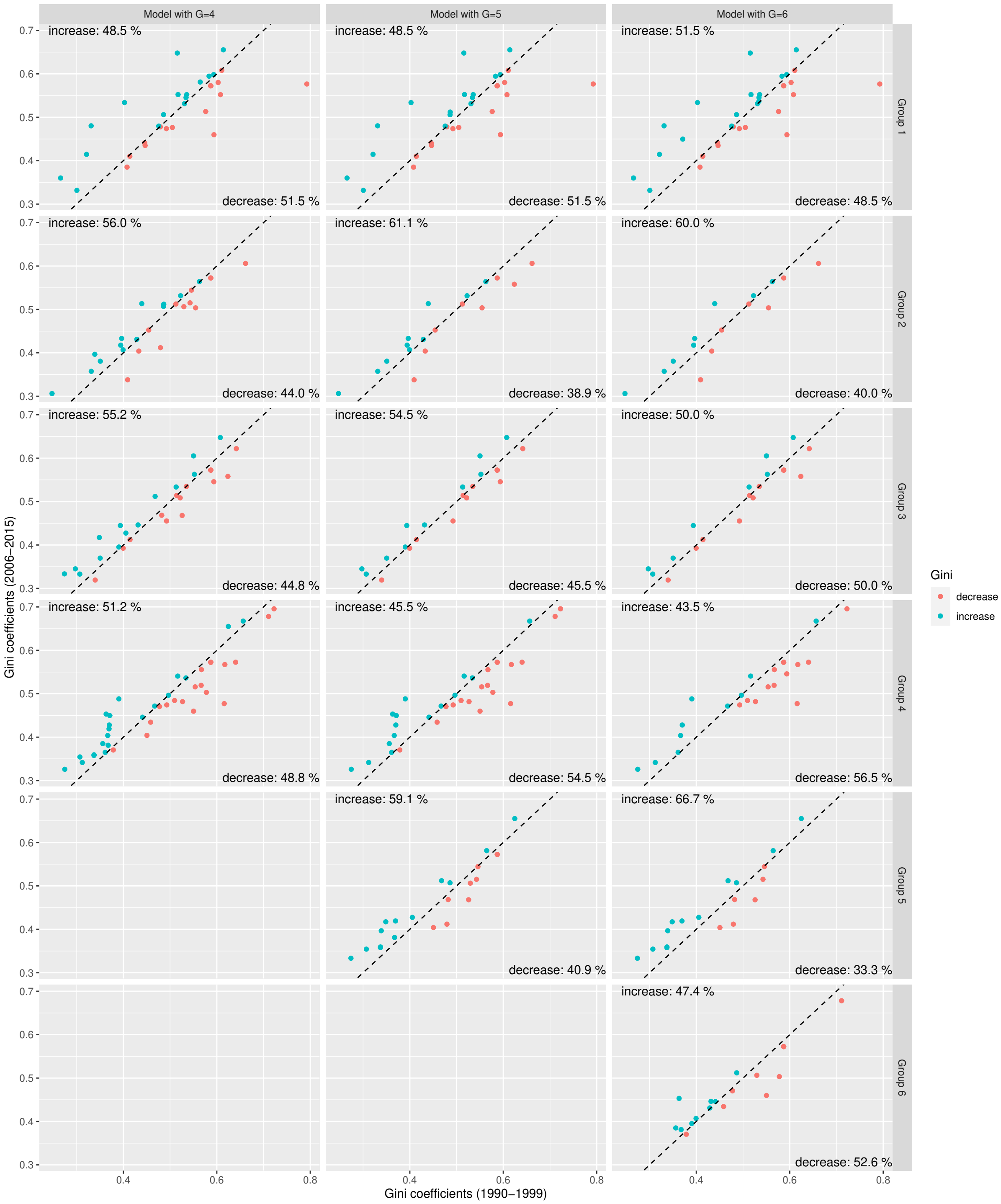


Figure 2: Changes in Gini coefficient estimates from the first 10 years to the last 10 years of the sample period



Note: The graph is reported vertically, starting from $G = 4$ in the first column through to $G = 6$ in the third column. In each of the columns, the uppermost graph is for Group1, followed by Group2 in that order towards the last country-groups.

Figure 3: Marginal Products of S and Z: Group2, Group3, and Group5

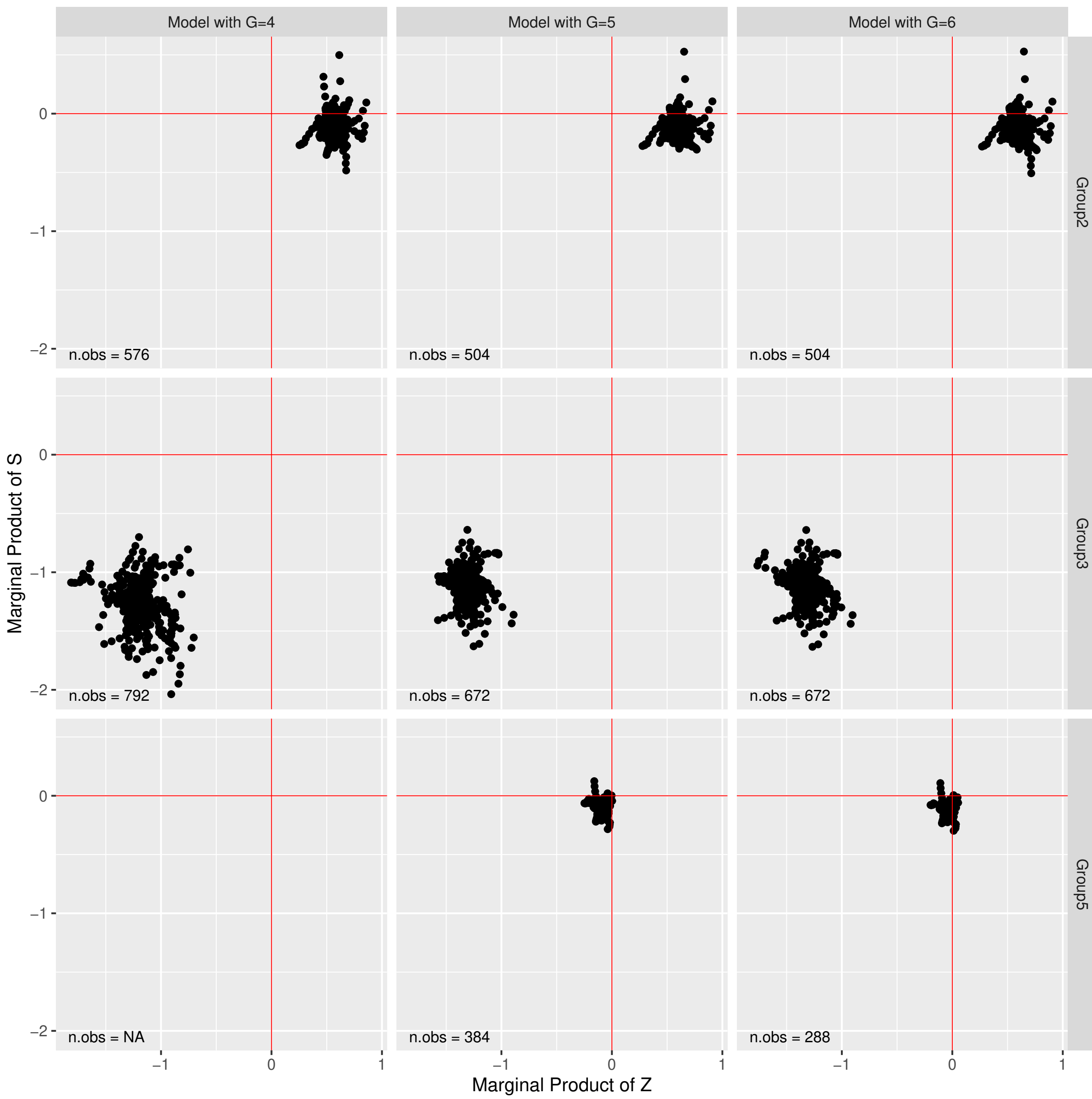


Table 1: Summary statistics: Output and inputs from 1990-2015

	Mean	SD	Min.	Max.	Units
Output-side GDP	0.0309	0.0312	0.0007	0.2248	2011 US dollars
Physical capital	0.1394	0.1453	0.0011	1.1420	2011 US dollars
Secondary education	2.6276	1.5897	0.0600	8.6500	Years
Electricity	0.0019	0.0022	0.0000	0.0143	Gigawatts
Main phone lines	0.4085	0.3860	0.0004	1.4906	Number of lines
Road networks	0.0165	0.0181	0.0004	0.1393	Kilometers

Note: Each variable, except the secondary variable, is transformed in per worker terms. The basic summary statistics were calculated over a sample of 130 countries from 1990-2015.

Table 2: Model selection based on information criteria

	$G=1$	$G=2$	$G=3$	$G=4$	$G=5$	$G=6$	$G=7$	$G=8$	$G=9$	$G=10$
IC in (4)	-3.631	-3.695	-3.721	-3.742	-3.743^{*}	-3.735	-3.720	-3.724	-3.715	-3.696
IC from Liu et al. [2020]	-1.855	-1.889	-1.904	-1.919	-1.922^{*}	-1.921	-1.917	-1.921	-1.919	-1.913

Note: The asterisk ($*$) shows the minimum value among the candidate models. Both information criteria show that the model with five groups is the optimal model.

Table 3: Country Classification List ($G = 5$)

Name	**Code**	**Name**	**Code**	**Name**	**Code**	**Name**	**Code**	**Name**	**Code**
Group1		**Group2**		**Group3**		**Group4**		**Group5**	
36 countries		21 countries		28 countries		29 countries		16 countries	
Bangladesh	BGD	Colombia	COL	Kyrgyz	KGZ	Niger	NER	Tajikistan	TJK
Sudan	SDN	Finland	FIN	Jordan	JOR	Morocco	MAR	Poland	POL
Lithuania	LTU	Cyprus	CYP	Chile	CHL	Switzerland	CHE	Slovak Rep.	SVK
Philippines	PHL	Israel	ISR	Peru	PER	Iceland	ISL	New Zealand	NZL
Latvia	LVA	South Africa	ZAF	France	FRA	Zimbabwe	ZWE	Bulgaria	BGR
Korea	KOR	Croatia	HRV	Greece	GRC	Syria	SYR	Guatemala	GTM
Mozambique	MOZ	Sri Lanka	LKA	Spain	ESP	Paraguay	PRY	Nepal	NPL
Togo	TGO	Portugal	PRT	Slovenia	SVN	Ireland	IRL	Dominica	DOM
Czech Rep.	CZE	Indonesia	IDN	Germany	DEU	Italy	ITA	Australia	AUS
Benin	BEN	Mali	MLI	Cameroon	CMR	Maldives	MDV	Belgium	BEL
Uganda	UGA	Panama	PAN	Austria	AUT	Iraq	IRQ	Lesotho	LSO
Senegal	SEN	Argentina	ARG	Zambia	ZMB	Russia	RUS	Ecuador	ECU
Sweden	SWE	Netherlands	NLD	Ukraine	UKR	Mauritius	MUS	Pakistan	PAK
Egypt	EGY	Kenya	KEN	Tunisia	TUN	Serbia	SRB	Ghana	GHA
Burundi	BDI	Belize	BLZ	Gambia	GMB	Hungary	HUN	Jamaica	JAM
Romania	ROU	Namibia	NAM	Turkey	TUR	Haiti	HTI	Malta	MLT
India	IND	United Arab Emirates	ARE	Costa Rica	CRI	Yemen	YEM		
Estonia	EST	Canada	CAN	Algeria	DZA	Brunei Darussalam	BRN		
Cambodia	KHM	Albania	ALB	Brazil	BRA	United States	USA		
Gabon	GAB	Norway	NOR	Iran	IRN	Trinidad and Tobago	TTO		
Denmark	DNK	Thailand	THA	United Kingdom	GBR	Saudi Arabia	SAU		
Myanmar	MMR			Congo, Rep.	COG	Bolivia	BOL		
Qatar	QAT			Botswana	BWA	Japan	JPN		
Cote d'Ivoire	CIV			Barbados	BRB	Luxembourg	LUX		
Honduras	HND			Kazakhstan	KAZ	Liberia	LBR		
China	CHN			Uruguay	URY	Kuwait	KWT		
Lao PDR	LAO			Mexico	MEX	Vietnam	VNM		
Central Africa	CAF			Nicaragua	NIC	Moldova	MDA		
El Salvador	SLV					Malaysia	MYS		
Sierra Leone	SLE								
Fiji	FJI								
Mongolia	MNG								
Rwanda	RWA								
Malawi	MWI								
Venezuela	VEN								
Mauritania	MRT								

Note: There are a total of 130 countries, and each column reports the names and initials of country-membership of the groups.

Table 4: Coefficient Estimates

	Group1		**Group2**		**Group3**		**Group4**		**Group5**		**Group6**	
$G = 4$												
	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.				
β_k	0.7998	0.0523	0.5853	0.0547	0.3971	0.0414	-0.4348	0.0511				
β_z	-0.2964	0.0409	0.5782	0.0482	-1.1913	0.1041	0.1961	0.0328				
β_s	0.1723	0.0239	-0.1295	0.0311	-1.2690	0.1077	-0.0131	0.0179				
β_{zs}	-0.0422	0.0119	0.4113	0.0422	0.9166	0.0754	-0.0205	0.0125				
$G = 5$												
	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.		
β_k	0.8549	0.0566	0.5410	0.0699	0.4431	0.0555	-0.4369	0.0553	0.3831	0.0175		
β_z	-0.3359	0.0436	0.6167	0.0516	-1.3017	0.1499	0.2364	0.0335	-0.1039	0.0401		
β_s	0.1625	0.0267	-0.1302	0.0347	-1.1192	0.1047	-0.0098	0.0197	-0.1055	0.0290		
β_{zs}	-0.0376	0.0142	0.4305	0.0449	0.7743	0.0734	-0.0196	0.0134	0.2130	0.0303		
$G = 6$												
	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.	Estimate	S. E.
β_k	0.9857	0.0638	0.5466	0.0662	0.4456	0.0575	-0.4476	0.0508	0.3779	0.0176	0.1969	0.0419
β_z	-0.4239	0.0480	0.6111	0.0491	-1.3157	0.1571	0.2540	0.0376	-0.0530	0.0352	-0.0841	0.0480
β_s	0.1238	0.0306	-0.1339	0.0336	-1.1220	0.1040	-0.0138	0.0185	-0.1203	0.0331	0.1928	0.0122
β_{zs}	-0.0013	0.0215	0.4336	0.0446	0.7766	0.0726	-0.0143	0.0143	0.2116	0.0319	-0.0361	0.0103

Note: Information criterion selected five as the optimum groups existing in the panel dataset ($G = 5$). The point estimates from $G = 4$ (underselection) and $G = 6$ (overselection) are presented to clearly show how the methodology used consistently classified countries in the panel dataset.

Table 5: Signs of estimated parameters: substitute technology

	Group1	**Group4**	**Group6**
the number of groups in the estimated model is four ($G = 4$)			
$\frac{\partial Y}{\partial S}$	(960 , 0)	(0 , 792)	
$\frac{\partial Y}{\partial Z}$	(0 , 960)	(792 , 0)	
$\frac{\partial^2 Y}{\partial S \partial Z}$	(0 , 960)	(0 792)	
the number of groups in the estimated model is five ($G = 5$)			
$\frac{\partial Y}{\partial S}$	(864 , 0)	(13 , 683)	
$\frac{\partial Y}{\partial Z}$	(0 , 864)	(696 , 0)	
$\frac{\partial^2 Y}{\partial S \partial Z}$	(0 , 864)	(48 648)	
the number of groups in the estimated model is six ($G = 6$)			
$\frac{\partial Y}{\partial S}$	(648 , 0)	(0 , 624)	(384 , 0)
$\frac{\partial Y}{\partial Z}$	(0 , 648)	(624 , 0)	(0 , 384)
$\frac{\partial^2 Y}{\partial S \partial Z}$	(0 , 648)	(24 , 600)	(24 , 360)

Note: The entry (A,B) shows that the number of positive estimates is A and the number of negative estimates is B. The average cross derivative estimates are -0.0933 ($G = 4$), -0.0922 ($G = 5$), -0.0538 ($G = 6$) for Group1, -0.0231 ($G = 4$), 0.0209 ($G = 5$), 0.0131 ($G = 6$) for Group4, and 0.0271 ($G = 6$) for Group6, where G is the number of groups in the estimated model.

Table 6: Signs of estimated parameters: substitute technology

	Group2	**Group3**	**Group5**
the number of groups in the estimated model is four ($G = 4$)			
$\frac{\partial Y}{\partial S}$	(30 , 546)	(0 , 792)	
$\frac{\partial Y}{\partial Z}$	(576 , 0)	(0 , 792)	
$\frac{\partial^2 Y}{\partial S \partial Z}$	(576 , 0)	(792 0)	
the number of groups in the estimated model is five ($G = 5$)			
$\frac{\partial Y}{\partial S}$	(19 , 485)	(19 , 653)	(6 , 378)
$\frac{\partial Y}{\partial Z}$	(504 , 0)	(0 , 672)	(1 , 383)
$\frac{\partial^2 Y}{\partial S \partial Z}$	(480 , 24)	(672 0)	(193 , 191)
the number of groups in the estimated model is six ($G = 6$)			
$\frac{\partial Y}{\partial S}$	(25 , 479)	(0 , 672)	(4 , 284)
$\frac{\partial Y}{\partial Z}$	(504 , 0)	(0 , 672)	(17 , 271)
$\frac{\partial^2 Y}{\partial S \partial Z}$	(504 , 0)	(672 , 0)	(192 , 96)

Note: The entry (A,B) shows that the number of positive estimates is A and the number of negative estimates is B. The cross derivative estimates are 0.3363 ($G = 4$), 0.3454 ($G = 5$), 0.3679 ($G = 6$) for Group2, 2.4239 ($G = 4$), 2.1891 ($G = 5$), 2.2106 ($G = 6$) for Group3, and 0.3201 ($G = 5$), 0.3685 ($G = 6$) for Group5, where G is the number of groups in the estimated model.