12 Infrastructure Productivity with a Long Persistent Effect

Tsukai, Makoto[1] and Kobayashi, Kiyoshi[2]

[1]Graduate School of Engineering, Hiroshima University
[2]Graduate School of Management, Kyoto University

12.1 Background and purpose

12.1.1 Production function approach

Appropriate investment and management of infrastructure is an important issue in national or regional planning. Aschauer (1989) reported two important findings: that infrastructure productivity was significantly higher than expected, and that fewer investments in the U.S after 1970 would result in less growth of national productivity. Since then, measurement of infrastructure productivity has become a focal issue in national or regional management policy, and a large number of empirical studies have accumulated; see Strurm (1998). An aggregate production function approach for time-series applied in Aschauer's study, however, was criticized by economists, or econometricians. Among the assumptions inherent in the production function approach, "constant return to scale" and "competitive input factor market" are often viewed with suspicion from the viewpoint of the endogenous growth theory by Romer (1986) and Lucas (1988) or from an uncompetitive structure of infrastructure market, respectively. Basu and Fernald (1997) empirically tested these assumptions with the production function approach, based on data from 34 sectors of U.S. industries. They concluded that in some segmented sectors of industry, the constant return to scale assumption was violated but that it held true in most sectors.

Holtz-Eakin and Lovely (1996) analyzed the effect of infrastructure on economic activities by using a general equilibrium system explicitly considering scale economies. They showed that infrastructure decreases the input factor cost, increases the variety of industries, and increases the number of newly founded companies. Haughwout (2002) applied the spatial equilibrium theory considering regional monopolistic power over input factor market, and measured infrastructure productivity. This study showed that infrastructure provided significant marginal benefit on regional economic activities. Most studies have reported that there is a positive production /cost reduction effect caused by infrastructure, but unfortunately, some studies have paid little attention to the data generation process that significantly influences findings through model specifications. Recent American, European, and Japanese studies about the production function approach have been widely reviewed by Ejiri et al. (2000).

From an econometric standpoint, model misspecification, the possibility of reverse causation from the product to infrastructure, and spurious correlation are often focused upon as nuisance problems in statistical regression analysis. All these faults in modeling result in inconsistent, inefficient, or biased estimates. A typical model misspecification is to dismiss technological improvement or scale economies. Duggal et al. (1999) formulated the time series production function model which explicitly considered technological growth as a function of infrastructure, and nonlinear labor productivity (allowing increasing and decreasing return to scale), and then applied the model to U.S. data. In this approach, it is necessary to estimate factor demand function, due to a violation of the constant return to scale assumption; hence the two stage least square method was adopted. The result obtained in this study supported that of Aschauer. Everaert and Heylen (2001), used the simultaneous equations method based on the error correction mechanism, which uses a differenced data series with first order, in order to test spurious correlation, reverse causation, and endogenous bias problems in the production function approach. In the application to Belgian data, the reverse causation from output product to infrastructure was rejected. To address the spatial correlation problem, many studies using the spatial econometrics approach have been conducted. Tsukai et al. (2002) estimated regional production function with spatial spillover effect, and in fact found a significant spatial spillover effect. The serial autocorrelation problem, however, remains unresolved in their study, since the production function is estimated by cross-sectional data.

Unfortunately, our review found no studies focusing on the persistent (lagged) production effect of infrastructure. An analysis of the persistent production effect of infrastructure requires an appropriate model specification for both spatial expansion and temporal persistency. Another requirement for the time-series production function approach is an appropriate consideration of the technological innovation effect. The technological innovation effect is modeled as residuals that cannot be explained in structural terms of the production function, but which might show particular spatio-temporal structure. In other words, the characteristics of spatial / cross-sectional structure should be explicitly modeled. In this study, the lagged effects of infrastructure productivity and technological innovation are specified as a multiple time series model with a long persistent effect.

12.1.2 Long persistent productivity of infrastructure

If the production effect of infrastructure lasts over years after the investment, multiple time series used for regional production function would have the long persistent property, which in the time series data is a kind of statistical property of the data generation process (DGP), mathematically defined in Sect. 12.2.1. Various empirical studies about time series data such as those involving temperature in climatology, voting behavior in political science, currency exchange rate in macro-economics, and merger and acquisition behavior in management economics, have reportedly found the long persistent property in these time-series data (Box-Steffensmeier and Tomlinson 2000; Barkoulas et al. 2001).

The studied series obtained by aggregation over some regions or several durations would have the long persistent property, even if each series does not have long persistency. Davidson and Sibbertsen (2005) gave simulation results to demonstrate what appears to be the long persistent property in aggregated data series from non-persistent disaggregated series. Suppose a process exhibits short memory fluctuations (i.e. uncorrelated standard disturbance such as with white noise) around a local average which randomly switches such that the durations of the regimes follow a power low. The authors showed that the aggregated process composed of a large number of independent copies of such process results in fractional Brownian motion, known to show long persistent behavior. They stated that an aggregated time series actually does not correspond to the single representative firm, but encompasses an extensive geographical area including many heterogeneous firms or regions that are interdependent. In terms of long persistency in economical data, Michelacci (2004) refers to the delay of innovation to replace capital in the private sector. Even though new private capital embedding the latest technology can make more profit, an immediate catch-up strategy to replace present capital with the latest one may not be optimal for each individual firm. An option to postpone the catch-up may rather be preferred under the rapid growth of technology, or in the case where present capital has been recently replaced. In a situation where some firms immediately adopt new technology but others do not, technological innovation in aggregated series would show long persistency.

Michelacci's consideration about the "delay of innovation" phenomenon may also be applicable to infrastructure productivity measured through private sector output. Suppose there is a change in land use, such that factories are not immediately, but gradually concentrated along newly constructed infrastructure. Then the production effect by new infrastructure would be observed with some delays. The other notable point is that infrastructure data is estimated by aggregation of individual stock with different vintage. If infrastructure productivity depends on its vintage, because of, for e.g., gradual change in land use, it is not appropriate to assume an instantaneous productivity effect of infrastructure. Therefore, the long persistent effect of infrastructure on production output seems worth explicitly considering in the econometric measurement of infrastructure production effect.

12.1.3 Time-series model with a long persistent effect

In the Box-Jenkins approach known as a standard time series analysis, a non-stationary time series should be differenced in order to convert it into stationary. Then the converted (stationary) series is used for modeling the stochastic process with AR and MA parameters, called the ARIMA model. While differentiation to remove the non-stationary property of time series data is a popular data handling technique proposed in the standard Box-Jenkins method, Munnel (1992) argued that first differencing destroys long-term relationships in the data, and therefore does not make economic sense. Strum and deHaan (1995) estimated production function, and obtained insignificant estimates of private capital and labor by using first differenced data series (they assumed original data series are integrated).

The loss of information by first differencing was originally pointed out by Granger(1980) and Hosking(1981). Instead of first differencing, they proposed fractional integration order between no integration (level) and integration. Fractionally integrated time series exhibits long persistent behavior but still holds the stationary property if fractional integration order is less than 0.5. The ARFIMA (Auto-Regressive, Fractionally Integrated, Moving Averaged) model that combines fractional order of time series integration with the conventional autoregressive / moving average method has been widely applied to prove its empirical applicability, such as for stock price series or international exchange rate (Smith 1997; Henry and

Olekalns 2002; Robinson and Hidalgo 2003). The ARFIMA model can include exogenous variables for deterministic trend, called the ARFIMAX model (ARFIMA model with exogenous variables). In order to measure infrastructure productivity, regional output series should be modeled as stochastic process, which deterministic term is explained by the production function including labor, private capital, and infrastructure term.

The purpose of this study is to clarify the long persistent effect of infrastructure on regional productivity. The regional production function with fractional integration is formulated and is applied to Japanese data. This Chap. is organized as follows: Section 12.2 shows the general formulation of the ARFIMA model with exogenous variables, and specifies the production function to be estimated. Data and the results of estimation with some discussion are presented in Sect. 12.3. Finally, the conclusion and a summary of further issues for study appear in Sect. 12.4.

12.2 Long persistent model

12.2.1 ARFIMA model with exogenous variables (ARFIMAX model)

The formulation of an ARFIMA model without exogenous variables is similar to the ARIMA model proposed by Box and Jenkins, but an ARFIMA model permits fractional differencing / integration order. Suppose that t is discrete time indices ($t = 0,...,T$), Y_t is an endogenous variable driven by normally distributed disturbances as $\varepsilon_t \sim N(0,\sigma^2)$, L is a lag operator defined by $LY_t = Y_{t-1}$, and d is fractional differencing / integration order. The ARFIMA model is formulated as follows in Eq. (12.1).

$$(1-L)^{d_Y}\phi(L)Y_t = \varphi(L)\varepsilon_t \tag{12.1}$$

$$(1-L)^{d_Y} = \sum_{j=0}^{\infty}\frac{\Gamma(d_Y+1)}{\Gamma(j+1)\Gamma(d_Y-j+1)}\cdot(-1)^j\cdot L^j \tag{12.2}$$

$$\phi(L) = 1-\phi_1 L-\phi_2 L^2 - \cdots \tag{12.3}$$

$$\varphi(L) = 1-\varphi_1 L-\varphi_2 L^2 - \cdots \tag{12.4}$$

In Eq. (12.1), $d_Y > 0$ indicates fractional differencing, $d_Y < 0$ indicates fractional integration, and $d_Y = 0$ indicates no differencing, no integration. Note that Eq. (12.2) is obtained as Maclaurin expansion of $(1-L)^{d_Y}$ which has a structure mathematically identical with $f(x) = (1-x)^d$ as follows:

$$f^{(j)}(x) = d \cdot (d-1) \cdots (d-j+1) \cdot (1-x)^{d-j} (-1)^j = \frac{d!(1-x)^{d-j}(-1)^j}{(d-j)!}$$

Therefore

$$(1-x)^d = \sum_{j=0}^{\infty} \frac{x^j}{j!} f^{(j)}(0) = \sum_{j=0}^{\infty} \frac{x^j}{j!} \cdot \left(\frac{d!(-1)^j}{(d-j)!} (1-x)^{d-j} \right)_{x=0} = \sum_{j=0}^{\infty} \frac{d!(-1)^j}{j!(d-j)!} \cdot x^j = \sum_{j=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(j+1)\Gamma(d-j+1)} \cdot (-1)^j \cdot x^j$$

In Eq. (12.2), $\Gamma(a)$ is a gamma function, which is interpreted as expansion of factorial calculation into real numbers, such as $\Gamma(a) = \int_0^{\infty} x^{a-1} \mathrm{e}^{-x}\, dx$, $\Gamma(a) = a\Gamma(a-1)$. If a is an integer, $\Gamma(a) = (a-1)!$. Note that $\Gamma(a)$ can also be defined for negative real numbers except for negative integers. Equation (12.2) indicates that fractional differencing can be expanded as an infinite series of lagged Y_t s. Here, an invertible condition for d_Y that $d_Y < 1/2$ is assumed (Granger, 1980). $\phi(L)$ in Eq. (12.3) and $\varphi(L)$ in Eq. (12.4) are called AR, or MA lag polynomials. Assume invertible condition for $\phi(L)$ and stationary condition for $\varphi(L)$, which is $\phi_1, \cdots, \phi_p < 1$ and $\varphi_1, \cdots, \varphi_q < 1$ (see Box et al.1994). In this case, Eq. (12.1) can be transformed into Eq. (12.1a).

$$Y_t = (1-L)^{-d_Y} \frac{\varphi(L)}{\phi(L)} \varepsilon_t = H(L)\varepsilon_t \qquad (12.1a)$$

$$(1-L)^{-d_Y} = \sum_{j=0}^{\infty} \frac{\Gamma(d_Y + j)}{\Gamma(j+1)\Gamma(d_Y)} \cdot L^j = \sum_{j=0}^{\infty} \pi_j L^j \qquad (12.2a)$$

Equation (12.2a) is again obtained, as Maclaurin expansion of $(1-L)^{-d_Y}$ indicates that the series of π_j contains infinite terms (see details in Minotani 1995). Since the inverse of AR polynomial $\phi^{-1}(L)$ can be expanded into an infinite series of lagged parameters, Y_t is driven by an infinite series of disturbance terms. When $\phi^{-1}(L)$ and $(1-L)^{-d_Y}$ satisfy stationality, they eventually reach 0. However, two parameter series differ in diminishing speed with a lag increase. Suppose that h_j is the composite parameter of Eq. (12.1). Gourierouex and Monfort (1997) showed the following approximation for h_j if $j \to \infty$:

$$h_j \approx \frac{\varphi_1}{\phi_1 \Gamma(d)} j^{d-1} \tag{12.5}$$

In Eq. (12.5), $j^{d-1}/\Gamma(d)$ stems from π_j converges to 0 under the stational condition $d < 1/2$. While AR and MA parameters also exponentially converge to 0 such that k^j $(|k| < 1)$, these parameters converge faster than $j^{d-1}/\Gamma(d)$, so that they do not appear in Eq. (12.5). Lower convergence property with lag in fractionally differenced series is called long persistency, or long memory.

The ARFIMA model with exogenous variable, called the ARFIMAX model, is formulated in Eq. (12.6).

$$(1-L)^{d_Y} \phi_Y(L) Y_t - (1-L)^{d_X} \phi_X(L) f(X_t) = \varphi(L)\varepsilon_t \tag{12.6}$$

where X_t are exogenous explanatory variables, $f(X_t)$ is a linear function of X_t, d_X and d_Y are fractional integration parameters (i.e. $d_X > 0$ indicates fractional integration), respectively, and $\phi_Y(L), \phi_X(L)$ are AR lag polynomials. Stationality and invertibility are assumed to be parameters in Eq. (12.6), and the transformation of Eq. (12.6) is shown as Eq. (12.6a).

$$Y_t = (1-L)^{d_X - d_Y} \frac{\phi_X(L)}{\phi_Y(L)} f(X_t) + (1-L)^{-d_Y} \frac{\varphi(L)}{\phi_Y(L)} \varepsilon_t \tag{12.6a}$$

Equation (12.6a) indicates that the long persistent property appears for both deterministic and stochastic processes as $f(X_t)$ and ε_t, hence they have a parameter series with infinite lags. Note that if $d_Y = 0$, the stochastic process is identical to the ARMA process in the latter case, hence long persistency does not appear in error term, and if $d_X = d_Y$, long persistency does not appear in infrastructure term.

In the ARFIMA model and the ARFIMAX model, differencing parameters are not given but estimated from observed data series. Difficulties in parameter estimation can be summarized into three types. First, because of slow and infinite persistency of fractionally differenced / or integrated process, a relatively larger number of samples is necessary to obtain good (i.e., consistent, unbiased, and efficient) parameters (Bhardwai and Swanson 2005). Second, a step-wise estimation procedure that initially estimates $\hat{d}$, then estimates the other AR and MA parameters using $\hat{d}$ to filter the series, such as the GPH method, tends to be biased under small samples, because d and the other parameters are dependent (Beran 1992; Igresias et al. 2005). Therefore, differencing parameters and other parameters should be simultaneously estimated. Robinson and Hidalgo proposed the maximum likelihood estimator obtained by frequency domain data series, which has desirable asymptotic properties (2003), and Tanaka showed that the ML estimator by time domain data also has such properties (1999). Finally, when long persistency is observed, it is difficult to distinguish whether the long memory process is truly appropriate as DGP. If regime switch (average drift) occurred in a time series, a misspecified long persistent model should yield a significant fractional differencing parameter (Davidson and Sibbertsen 2005; Dufrenot et al. 2005). Such difficulties should be carefully considered in empirical application.

Using composite parameters functioning as $H_X(L), H_\varepsilon(L)$, Eq. (12.6a) can be written into Eq. (12.6b).

$$Y_t = H_X(L) f(X_t) + H_\varepsilon(L)\varepsilon_t \tag{12.6b}$$

$H_X(L)$ and $H_\varepsilon(L)$ determine the shape of lag distribution of deterministic term and stochastic term. The lagged parameters for deterministic terms are already known as distributed lag models (Almon 1965; Siller 1973), in which average lag delay, l_x^e, can be calculated by using estimated lag parameters. Suppose that h_{t-i}^x s are a series of lag parameters ($h_t^x = 1$), a^T is a data aggregation span, and m is lag truncation length. l_x^e shown in Eq.(12.7).

$$l_x^e = \frac{\sum_{j=0}^{m} (j+1) h_{t-j}}{\sum_{j=0}^{m} h_{t-j}} \times a^T \tag{12.7}$$

Note that l_x^e of ARFIMAX model does not converge for infinite m, since l_x^e includes the term $j \cdot j^{d-1} = j^d$ if $d > 0$, see and compare Eq. (12.5) and Eq. (12.7). When finite m is set at empirical parameter estimation, l_x^e can be calculated by Eq. (12.7), but the result could be interpreted as a kind of approximated index. Since further discussion about this point is beyond the scope of the present paper, we adopt Eq. (12.7) to calculate average lag delay.

12.2.2 Specification of production function with a long persistent effect

In this paper, we estimate a simple Cobb-Douglas type production function. Unfortunately, the number of available samples is insufficient among annual statistics in Japan, so that cross-sectional and longitudinal data are pooled in order to obtain stable parameter estimates. The data we use is provided by Doi (2003) on his website, available for 46 prefectures from 1955 to 1998 (44 years), details of which are provided in Sect. 12.3.1. Suppose that i is regions ($i = 1,...,46; N = 46$), t is discrete time indices ($t = 1,...,44-m; T = 44-m$, from 1955 $+m$ to 1998, m is truncated lag length), Y_{it} is gross regional product, N_{it} is labor input, K_{it} is private capital stock, and G_{it} is infrastructure stock. The log-transformed production function with a constraint of constant returns to scale, and the deterministic progress of technology is shown in Eq. (12.8).

$$\log Y_{it} = \alpha_0 + \alpha_1 \log N_{it} + (1-\alpha_1) \log K_{it} + \alpha_2 \log G_{it} + \alpha_3 T_t \tag{12.8}$$

Among input factors, long persistency would appear in G_{it} but not in N_{it} or K_{it} because labor and private capital can be easily adjusted to maximize productivity, so no long persistent effect is expected. Equation (12.8) is a linear additive function in parameters, hence the long persistent parameters can be set for infrastructure as d_G. This setting makes it possible to test the hypothesis that the long persistent effect would be longer for infrastructure than for private capital. These lagged variables are shown in Eq. (12.9).

$$\overline{\log G_{it}} = (1-L)^{-d_G} \log G_{it} = \sum_{j=0}^{\infty} \binom{d_G + j - 1}{j} L^j \log G_{it} = \sum_{j=0}^{\infty} \frac{\Gamma(d_G + j)}{\Gamma(j+1)\Gamma(d_G)} \log G_{i,t-j} \tag{12.9}$$

In Eq. (12.9), non-lagged variable is included due to $\Gamma(d_G)/\Gamma(1)\Gamma(d_G) = 1$ if $j = 0$. Note that positive d_G indicates differencing and negative d_G indicates integration. Figure 12.1 plots lag coefficient values given by $\frac{\Gamma(d+j)}{\Gamma(j+1)\Gamma(d)}$ for lags ($j \geq 1$) at $d_G = 0.2$ and $d_G = -0.2$. In the case of $d_G = -0.2$, all lagged coefficients are positive.

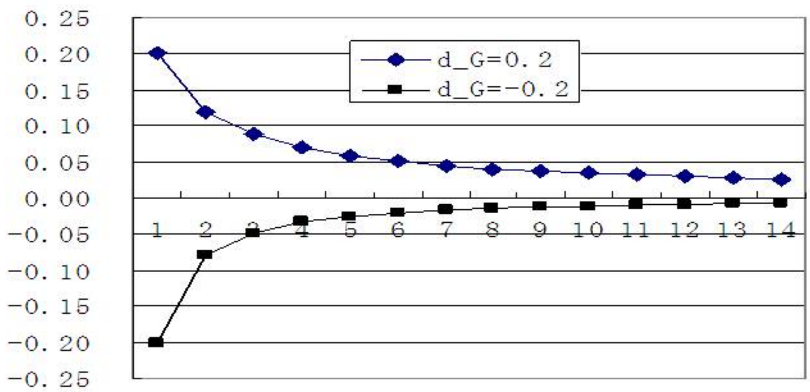


Fig.12.1 Decays in coefficients for lagged variable with long persistency

Therefore d_G is expected to be positive if past infrastructure has a positive production effect on present output.

In addition to lagged variables, also considered is the spatial spill over effect brought by infrastructure in other regions $G_{-i,t}$. $G_{-i,t}$ are weighted by inverse of distance, then summed up for all regions except for the own region.

$$\overline{\log G_{it}} = \sum_{j}^{j \neq i} w_{ij} \log G_{jt} \tag{12.10}$$

Equation (12.9) and Eq. (12.10) are added to Eq. (12.8). The regional production function with stochastic disturbance u_{it} is in Eq. (12.11). In this model, it is important to set the structure of the stochastic process appropriately.

$$\log Y_{it} = \alpha_0 + \alpha_1 \log N_{it} + (1-\alpha_1)\log K_{it} + \phi\overline{\log G_{it}} + \alpha_2 \overline{\overline{\log G_{it}}} + \alpha_3 T_t + u_{it} \tag{12.11}$$

In Eq. (12.6), the long persistent property will also appear in the stochastic error term in the case of $d_Y > 0$. In our model, long persistency and first order MA polynomial structure are assumed for the stochastic process, in Eq. (12.12).

$$u_{it} = \rho\varepsilon_{i,t-1} + (1-L)^{-d_\varepsilon}\varepsilon_{it}, \; or \;\; u_{it} = (1-\varphi L)(1-L)^{-d_\varepsilon}\varepsilon_{it} \tag{12.12}$$

This specification corresponds to the case of $\phi_X(L) = \phi_Y(L) = 1$, and $\varphi_\varepsilon(L) = 1 - \varphi L$ in Eq. (12.6). Parameters to be estimated are $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \varphi$ in Eq. (12.11), d_G in Eq. (12.9), d_ε and MA parameter φ in Eq. (12.12). Here, let us introduce some matrix algebra in order to clarify the difference in differencing order for each term.

$$\mathbf{Y} = \{\ldots, \log Y_{it}, \ldots\}', \;\; \mathbf{X} = \left\{\begin{matrix} \vdots \\ 1, \log N_{it}, \log K_{it}, \sum_{j \neq i} w_{ij} \log G_{jt}, \log T_t \\ \vdots \end{matrix}\right\}, \tag{12.13}$$

$$\mathbf{g} = \{\ldots, \log G_{it}, \ldots\}', \; \boldsymbol{\varepsilon} = \{\ldots, \varepsilon_{it}, \ldots\}', \boldsymbol{\alpha} = \{\alpha_0, \alpha_1, (1-\alpha_1), \phi, \alpha_3\}'$$

Using the above algebra, Eq. (12.14) or Eq. (12.15) is obtained (as to no differencing parameter for error term).

$$\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha} - \alpha_2(1-L)^{-d_G}\mathbf{g} = (1-\varphi L)(1-L)^{-d_\varepsilon}\boldsymbol{\varepsilon} \tag{12.14}$$

$$(1-\varphi L)^{-1}\left[(1-L)^{d_\varepsilon}\left(\mathbf{Y}-\mathbf{X}\boldsymbol{\alpha}\right)-\alpha_2(1-L)^{d_\varepsilon-d_G}\,\mathbf{g}\right]=\boldsymbol{\varepsilon} \tag{12.15}$$

For the error term, $\boldsymbol{\varepsilon} \sim N(0,\sigma^2\mathbf{I})$ is assumed. Parameters are estimated by pseudo maximum likelihood estimation for time domain data series proposed by Chung and Baille (1993) and Tanaka (1999), called the CSS estimator. In this study, the data set is composed of 46 time series for each prefecture. The log-likelihood function is shown in Eq. (12.16).

$$\log \mathcal{L} = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}\sum_i\sum_t \varepsilon_{it}^2 \tag{12.16}$$

By using estimated parameters, instantaneous marginal productivity and the sum of persistent marginal productivity of infrastructure are calculated in Eq. (12.17), and Eq. (12.18).

$$\frac{\partial Y_t}{\partial G_{it}} = \frac{\partial Y_{it}}{\partial G_{it}} + \sum_j^{j\neq i}\frac{\partial Y_{jt}}{\partial G_{it}} = \alpha_2\frac{Y_{it}}{G_{it}} + \phi\sum_j^{j\neq i} w_{ij}\frac{Y_{jt}}{G_{it}} \tag{12.17}$$

$$\sum_{j=1}^{m}\frac{\partial Y_{i,t+j}}{\partial G_{it}} = \alpha_2\sum_{j=1}^{m}\frac{\Gamma\left(d_G+j\right)}{\Gamma\left(j+1\right)\Gamma\left(d_G\right)}\frac{Y_{i,t+j}}{G_{it}} \tag{12.18}$$

In Eq. (12.17), the first term on the right indicates productivity contribution to own region, while the second term is contribution to other regions. The truncation order m in Eq. (12.18) is exogenously set at parameter estimation.

12.2.3 Diagnostic tests

Since the CSS estimator is asymptotically consistent and asymptotically efficient, a likelihood ratio test based on Large-sample theory can be applied as a diagnostic test for parameter significance (Chung and Baille 1993). Now the parameter vector of the CSS estimator in Eq. (12.15) is denoted by $\ddot{\boldsymbol{\beta}}^\circ = (\ddot{\beta}_1^\circ,\cdots,\ddot{\beta}_i^\circ,\cdots)$. The null and alternative hypothesis of likelihood test is in Eq. (12.19).

$$\begin{cases} H_0 & \ddot{\beta}_i^\circ = 0 \\ H_0 & \ddot{\beta}_i^\circ \neq 0 \end{cases} \tag{12.19}$$

The parameter vector of the CSS estimator estimated under the constraint of objective parameter as $\ddot{\beta}_i^\circ = 0$ is denoted by $\ddot{\bar{\boldsymbol{\beta}}}_i^\circ$. A statistics of likelihood ratio for $\ddot{\beta}_i^\circ$ is ξ_i in Eq. (12.20).

$$\xi_i = -2\left\{\ln\left[S(\ddot{\boldsymbol{\beta}},\ddot{\sigma}^2)\right] - \ln\left[S(\ddot{\bar{\boldsymbol{\beta}}}_i,\ddot{\bar{\sigma}}^2)\right]\right\} \tag{12.20}$$

Since ξ_i asymptotically converges to $\chi^2(1)$ distribution, the critical value to reject null hypothesis H_0 is obtained in Eq. (12.21).

$$\xi_i \geq \chi^2_{(100-\alpha)\%}(1) \tag{12.21}$$

As in Eq. (12.11), under the null hypothesis α_2 (for infrastructure) should be considered simultaneously with estimates of d_G, because these two parameters are dependent. Statistics for simultaneously testing α_2 with d_G are ξ_{G,d_G}, shown in Eq. (12.22),

$$\xi_{G,d_G} = -2\left\{\ln\left[S(\ddot{\boldsymbol{\beta}},\ddot{\sigma}^2)\right] - \ln\left[S(\ddot{\bar{\boldsymbol{\beta}}}_{G,d_G},\ddot{\bar{\sigma}}^2)\right]\right\} \tag{12.22}$$

where $\ddot{\bar{\boldsymbol{\beta}}}_{G,d_G}$ is the parameter vector of CSS estimator under the constraint with $\alpha_2 = d_G = 0$. ξ_{G,d_G} follows $\chi^2(1)$ distribution.

In terms of longitudinal correlation, Durbin-Watson statistics (DW) are used to detect first order serial autocorrelation. Suppose $\hat{\varepsilon}_{it}$ are the residuals at region i and time t. Overall Durbin-Watson statistics are in Eq. (12.23a), and Regional DW statistics are in Eq. (12.23b). If positive first order autocorrelation remains in residuals then DW is close to 0, and in the case of no first order autocorrelation, DW is close to 2.

$$DW = 2 - 2 \cdot \frac{\sum_{i=1}^{N} \sum_{t=m+1}^{T} \hat{\varepsilon}_{it} \hat{\varepsilon}_{i,t-1}}{\sum_{i=1}^{N} \sum_{t=m+1}^{T} \left(\hat{\varepsilon}_{it}\right)^2} \tag{12.23a}$$

$$DW_i = 2 - 2 \cdot \frac{\sum_{t=m+1}^{T} \hat{\varepsilon}_{it} \hat{\varepsilon}_{i,t-1}}{\sum_{t=m+1}^{T} \left(\hat{\varepsilon}_{it}\right)^2} \tag{12.23b}$$

When the residuals of cross sectional models show unspecified spatial correlation, the spatial correlation left in $\hat{\boldsymbol{\varepsilon}}_t$ can be tested by Moran's *I* statistics (Moran, 1948). Suppose $\hat{\boldsymbol{\varepsilon}}_t$ is a residual vector for each cross-section. Based on $\hat{\boldsymbol{\varepsilon}}_t$, I^{M} is calculated by Eq. (12.24).

$$\mathrm{I}^{\mathrm{M}} = \frac{\hat{\boldsymbol{\varepsilon}}_t' \left(\frac{1}{2} (\mathbf{W}' + \mathbf{W}) \right) \hat{\boldsymbol{\varepsilon}}_t}{\hat{\boldsymbol{\varepsilon}}_t' \hat{\boldsymbol{\varepsilon}}_t} \tag{12.24}$$

where, $\mathbf{W}$ is a spatial weight matrix. We assumed it by inverse distance matrix identical which have w_{ij} s in Eq. (12.10). The standardized Moran's *I* statistics $\overline{\mathrm{I}}^{\mathrm{M}}$ is defined as follows;

$$\overline{\mathrm{I}}^{\mathrm{M}} = \frac{\mathrm{I}^{\mathrm{M}} - E\left[\mathrm{I}^{\mathrm{M}}\right]}{Var\left[\mathrm{I}^{\mathrm{M}}\right]^{1/2}}$$

where, $E\left[\mathrm{I}^{\mathrm{M}}\right] = \frac{\mathrm{tr}\left(\mathbf{U}\right)}{n-k}$, $\quad Var\left[\mathrm{I}^{\mathrm{M}}\right] = \frac{\mathrm{tr}\left(\mathbf{UPW}'\right) + \mathrm{tr}\left(\mathbf{U}^2\right) + \left[\mathrm{tr}\left(\mathbf{U}\right)\right]^2}{(n-k)(n-k+2)} - \left(E\left[\mathrm{I}^{\mathrm{M}}\right]\right)^2$

and n is a number of samples, k is a number of parameters, $\mathbf{P}$ (projection matrix) and $\mathbf{U}$ are defined as $\mathbf{P} = \mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)\mathbf{X}'$, $\mathbf{U} = \mathbf{PW}$, respectively. It is known that $\overline{\mathrm{I}}^{\mathrm{M}}$ asymptotically follows the standardized normal distribution.

12.3 Empirical measurement of infrastructure productivity

12. 3.1 Data

The privatization of Japan Railway companies (JRs), Nippon Telephone and Telecommunications (NTTs), and Japan Tobacco (JT) caused a disjoint at 1987 in the longitudinal statistics of infrastructure. The corrections in the data set of infrastructure and private capital are provided by Doi in order to remove the inconsistency before and after 1987 (2000). Both data were depreciated with the identical rule to infrastructure data. Gross Regional Product (GRP) and labor force are also estimated based on the national census. Input and output data are deflated at the 1995 price. These data sets were provided for each prefecture unit, from 1955 through 1999. In order to ameliorate the small sample problem, we pooled all the available data. Okinawa prefecture was excluded because of its distant location and the lack of data availability before 1972.

12.3.2 Results of estimation

In order to estimate Eq. (12.11), lag truncation order m needs to be set. After several trials, m is set at 10. Therefore, 34 cross-sections (from 1965 through 1998) and 46 prefectures data are pooled up to be used for parameter estimation. Then totally 1568 observations are used. Previous studies in the ARFIMA model reported that setting time-trend parameters affect d_G, d_ε and other structural parameters. Considering the economical situation of the data period, we set three durations with different time-trend parameters as follows: 1) 1965 to 1973 as a rapid economic growth term up to the oil shock in 1973; 2) 1974 to 1990 as a post oil shock term up to the end of the bubble economy in 1991; and 3) 1992 to 1998 as the post bubble economy. In the regional production function, time-trend parameter indicates an average growth of total factor products, or of technological innovation rate.

Table 12.1 shows the estimation results; the first two columns are the estimates of production function with long persistency and three deterministic trends, referred to as long memory model-1. The estimated value of infrastructure parameter is considerably small compared with that of conventional studies (0.2 to 0.3). Because the long persistent model measures persistent productivity and spatial spillover effect of infrastructure, the productivity index would separately appear in different terms. As alternate model specifications, two models are estimated. The middle two columns are the estimates of the model without long persistent terms for both as $d_{\varepsilon} = d_G = 0$, referred to as short memory model. The last two columns are the estimates of the model without long persistent term of infrastructure as $d_G = 0$, referred to as long persistent model-2. First, we compare the long memory model-1 with the short memory model. The adjusted R-squared coefficient is slightly high in the long persistent model. The critical values of Durbin-Watson statistics (1564 samples, 1 percent significant level, 11 parameters) are 1.868 to 1.896 for $\varphi > 0$ and 2.104 to 2.132 for $\varphi < 0$, so that $\varphi > 0$ is accepted if $DW < 1.868$, $\varphi = 0$ is accepted if $1.896 < DW < 2.104$, and $\varphi < 0$ is accepted if $DW > 2.132$, respectively. The statistic values calculated by Eq.(12.23a) indicate that $\varphi = 0$ is accepted for long persistent model-1, while $\varphi > 0$ is accepted for the MA model. The prefecture's Durbin-Watson statistics calculated by (12.23b) ranges between 1.576 and 2.189. The critical values of Durbin-Watson statistics (34 samples, 1 percent significant level, 11parameters) are 0.610 to 2.160 for $\varphi > 0$ and 1.840 to 3.390 for $\varphi < 0$. Table 12.2 shows that only two prefectures can reject $\varphi < 0$, but the others cannot test the autocorrelation. Figure 12.2 shows the standardized Moran's I for each cross-section. The rejecting range of no spatial correlation is $\left|\bar{I}^{M}\right| > 1.96$; therefore there is no significant spatial correlation for all cross-sections.

Table 12.1 Parameter estimates of regional production function

	Long persistent model-1		MA Model		Long persistent model-2	
Variables	estimates	Chi-squared $^{\$\$}$	estimates	Chi-squared$^{\$\$}$	estimates	Chi-squared $^{\$\$}$
Labor$^{\$}$	0.544**	24.9	0.488**	35.69	0.459**	62.69
Private Capital$^{\$}$	0.456**	99.94	0.512**	138.45	0.541**	105.73
Infrastructure$^{\$\$\$}$	0.028**	13.36	0.149**	14.73	0.072**	10.73
Spatial Spillover of infrastructure	0.030**	6.63	0.013**	1.13	0.022**	8.04
Long persistency of infrastructure	0.310**	13.36	-	-	-	-
Long persistency of stochastic error	0.438**	57.24	-	-	0.487**	150.52
MA(1) in stochastic error	0.753**	194.19	0.972**	370.25	0.734**	187.12
Deterministic trend (1965-1973)	0.077**	35.78	0.008**	0.06	0.078**	48.64

Table 12.1. (contd.)

Deterministic trend (1974-1991)	0.025**	9.57	(-0.021)**	1.47	0.027**	15.3
Deterministic trend (1992-1998)	(-0.016)**	23.9	(-0.051)**	78.86	(-0.014)**	38.1
Constant	(-2.612)**	13.73	(-0.320)**	0.03	(-2.516)**	48.3
Variance	0.032		0.034		0.033	-
No.of Samples	1564		1564		1564	
Adjusted R^2	0.999		0.998		0.993	
REG statistics	1.716		3.935**		2.024*	
Durbin-Watson statistics	1.93		1.718**		1.924	

$: Parameters of private capital and labor are constrained by constant returns to scale

$$: ** indicates null hypothesis of zero slope is rejected with 1 percent significance, and * indicates the null is rejected with 5 percent significance. Critical values of Chi-squared test are 3.84 (1 d.f., 5 percent significance), 6.63(1 d.f., 1 percent significance).

$$$: In Eq.(9), infrastructure parameter and long persistency parameter of infrastructure are jointly tested because they are nonlinearly dependent. Critical values of Chi-squared test are 5.99 (2 d.f., 5 percent significance), 9.21 (2 d.f., 1 percent significance)

Table 12.2 Results of Durbin-Watson Test

φ> 0	φ<0	Number .of preferences
neither reject nor accept	Rejected	17
(0.610 <DW <= 1.840)		
neither reject nor accept	neither reject nor accept	27
(1.840 < DW <= 2.160)		
Rejected	neither reject nor accept	2
(2.160 < DW <= 3.390)		

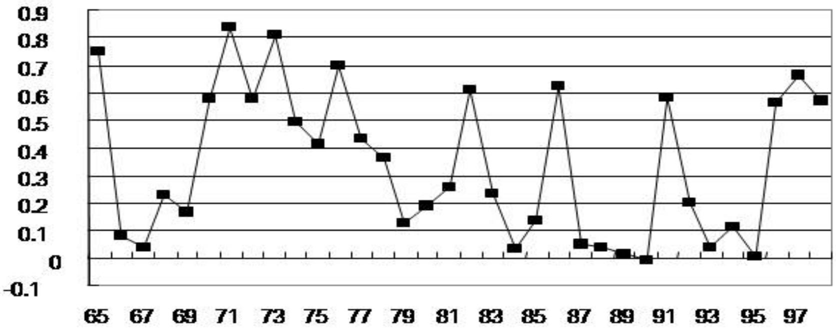


Fig. 12.2. Moran's Index

Parameters of labor, private capital and infrastructure are significant in long persistent model-1. Compared with the MA model, the labor parameter is slightly large, and private capital parameter is slightly small due to the constant returns to scale constraint. The estimated infrastructure parameter is considerably small compared with that of conventional studies (0.2 to 0.3). Because the long persistent model measures persistent productivity and the spatial spillover effect of infrastructure, productivity index would separately appear in different terms. Three time-trend parameters in the long persistent model are significant, but in the MA model, only one parameter for 1991 to 1998 is significant. The spatial spillover parameter is significant in long persistent model-1, but not for the MA model. The MA parameter is significant for both models. Note that the MA parameter is smaller in long persistent model-1. Both long persistent parameters in long persistent model-1 are positive, and fulfill stationary and invertible conditions as $|d_\varepsilon|, |d_G| < 0.5$. Since positive d_G is estimated, GRP for each region enjoys a positive production effect from past infrastructure as shown in Fig. 12.1, which shows the persistent production effect of infrastructure. And positive d_ε indicates that production activity is stationary but that exogenous shock will last for a while.

12.3.3 Statistical tests for long persistent specification

The appropriateness of long persistent specification for production function is tested to confirm two hypotheses simultaneously: that long persistent property remains in the model without long persistent specification and that long persistent property does not appear in the model with long persistent specification. In this Sect., we test these hypotheses by applying the REG test for residual series estimated by three specifications of regional production function. The REG test is proposed by Agiakloglou and Newbold (1994), which is an expansion of the Lagrangean Multiplier (LM) test for ARIMA model parameters by Godfrey (1979). The residuals of long persistent model-1, the MA model, and long persistent model-2 are denoted as $\hat{\boldsymbol{\varepsilon}}_t^1, \hat{\boldsymbol{\varepsilon}}_t^2$, and $\hat{\boldsymbol{\varepsilon}}_t^3$, respectively. The null hypothesis is that the residual series follows the white noise process as shown in Eq. (12.26).

$$\hat{\boldsymbol{\varepsilon}}_t^k = \mathbf{W}_k^t \quad (k = 1, 2, 3) \tag{12.26}$$

An alternative model is ARFIMA(0,d,0), as shown in Eq. (12.27).

$$(1 - L)^{d_k} \hat{\boldsymbol{\varepsilon}}_t^k = \mathbf{W}_k^t \quad (k = 1, 2, 3) \tag{12.27}$$

The LM test statistics on long persistent parameter can be seen in Eq. (12.28) where, T and R are number of cross-sections, and number of regions, respectively.

$$U^k = \sum_{t=1}^{T} \left. \frac{R\mathbf{W}_t^k \boldsymbol{\lambda}_t^k \left(\boldsymbol{\lambda}_t^{k\prime} \boldsymbol{\lambda}_t^k \right)^{-1} \boldsymbol{\lambda}_t^{k\prime} \mathbf{W}_t^k}{\mathbf{W}_t^{k\prime} \mathbf{W}_t^k} \right|_{d_k = 0} \tag{12.28}$$

$$\boldsymbol{\lambda}_t^k = \frac{\partial \mathbf{W}_t^k}{\partial d_k} \tag{12.29}$$

Note that in Eq. (12.29), partial differentiated vector $\boldsymbol{\lambda}_t^k$ is evaluated at neighbor of $d_k = 0$. Between Eq. (12.26) and Eq. (12.27), difference of degree of freedom is 1, hence the test statistics U^k follows chi-squared distribution with d.f.1. Agiakloglou and Newbold noted that Eq. (12.28) is proportional to the R-squared coefficient of the regression model that $\boldsymbol{\lambda}_t^k$ is regressed on $\mathbf{W}_t^k$ ($\boldsymbol{\lambda}_t^k$ and $\mathbf{W}_t^k$ are evaluated at $d_k = 0$), then proposed the REG test as the T-statistic test of $\boldsymbol{\lambda}_t^k$'s coefficient by estimating the regression model by OLS.

Note that $\hat{\boldsymbol{\varepsilon}}_t^k = \mathbf{W}_k^t\big|_{d_k=0}$ from Eq. (12.27), $\boldsymbol{\lambda}_t^k$ can be calculated as in Eq. (12.30).

$$\left.\frac{\partial \mathbf{W}_t^k}{\partial d_k}\right|_{d_k=0} = \left(\log(1-L)\right)(1-L)^{d_k}\,\hat{\boldsymbol{\varepsilon}}_t^k\Big|_{d_k=0} \tag{12.30}$$
$$= \log(1-L)\mathbf{W}_t^k\Big|_{d_k=0}$$
$$= \log(1-L)\hat{\boldsymbol{\varepsilon}}_t^k$$
$$= -\sum_{j=1}^{\infty} j^{-1}\hat{\boldsymbol{\varepsilon}}_{t-j}^k \ (=\mathbf{S}_t^k\ )$$

Equation (12.31) shows the auxiliary regression model of the REG test.

$$\hat{\boldsymbol{\varepsilon}}_t^k = \tau^k \mathbf{S}_k^t + \boldsymbol{\zeta}_t^k \quad (k = 1,2,3) \tag{12.31}$$

where, $\boldsymbol{\zeta}_t^k$ is the error term of auxiliary regression model. Therefore, a long persistent property test for the estimated residual is shown in Eq. (12.32).

$$\begin{cases} H_0^k & \hat{\tau}^k = 0 \\ H_1^k & \hat{\tau}^k \neq 0 \end{cases} \tag{12.32}$$

If H_0^k is rejected, the long persistent property remains in the residual series.

The result of the REG test is shown in Table 12.1 in the second row from the bottom. The REG statistic for the MA model is 3.935. Since H_0^2 is rejected with 1 percent significance, the long persistent effect remains in the residual series of the MA model. The REG statistic for the long persistent model-1 is 1.716. Since H_0^1 is not rejected with 5 percent significance, the residual series does not have the long persistent property. The REG statistic for the long persistent model-2 is 2.024. Since H_0^3 is rejected with 5 percent significance, the residual series has the long persistent property. Summarizing the results of the REG test, the residual series of MA model (without d_ε and d_G) and of long persistent model-2 (without d_G) have the long persistent property, while the residual series of long persistent model-

2 do not have the long persistent property. Therefore, we can conclude that the long persistent property of infrastructure productivity can not be rejected.

However, the implications from the above statistical test for the long persistent model have certain limitations. First, the possibility of misspecification for spatial auto-correlation remains in our model because spatial correlation structure in error term was not considered. Second, a problem might arise from an insufficient number of observations. Hosking (1996) asserted that the expected mean of the ARFIMAX process asymptotically distributes with normal distribution, but its parameters converge with $n^{1/2-d_\varepsilon}$, while the parameters of a linear regression model with i.i.d. error converge with $n^{1/2}$. Note that $\hat{d}_\varepsilon$ of our long persistent model-1 is 0.428; therefore the convergence rate of our model is very slow. In other words, a much larger number of observations is required in the estimation of the ARFIMAX model to keep its efficiency as high as that of the corresponding ARIMAX model. Third, structural change of deterministic time-trend might occur. Previous studies reported that misspecification of deterministic time-trend would affect $\hat{d}_\varepsilon$ (Davidson and Sibbertsen 2005; Dfrenot et-al. 2005). In order to specify these points properly, the long persistent property and spatial auto-correlation should be simultaneously modeled by spatially and temporally correlating with the long persistent effect. The possibility of structural change in deterministic time-trend, which is exogenously given in our production function, should be statistically tested. These points are further issues which need to be addressed in regional production function approach.

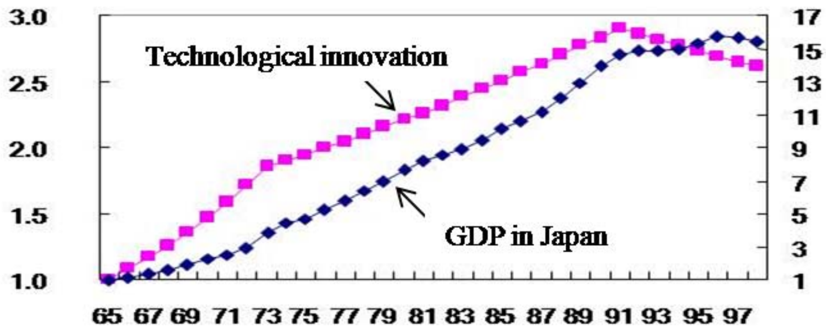
12.3.4 Average growth of technological innovation

From the estimates of deterministic time-trend, average growth of technological innovation is calculated. Table 12.3 shows a comparison of average growth of technological innovation between long persistent model-1 and the MA model. Figure 12.3 shows graphs of the average growth of technological innovation calculated from long persistent model-1 and growth of GDP. In this figure, the growth of technological innovation is prior to the growth of GDP. The technological innovation calculated from the MA model, however, does not fit with GDP growth. Figure 12.3 indicates that the long persistent property should be considered to estimate appropriate deterministic trends, but even with our approach some problems remain. In conventional studies, the least squared model with i.i.d. error structure or the ARIMA model is widely adopted. Our results indicate that the conventional approach, when assessing growth productivity, cannot separate between technological innovation and long persistent infrastructure productivity, and that the conventional approach cannot properly estimate deterministic time-trend parameters.

Table 12.3. Technological Innovation Rate

Term	Long-Persistent Model (percent)	MA Model (percent)
1965-1973	8.00	0.80
1974-1991	2.53	-2.08
1992-1998	-1.59	-4.97

Note: Technological innovation rate is calculated by using time trend parameters as as α_t^k as exp (α_t^k)-1



(Red line/left axis: Technological innovation, blue line: GDP in Japan)

Fig.12.3. Longitudinal plots of technological innovation and GDP growth in Japan

12.3.5 Infrastructure productivity

Figure 12.4 and Fig. 12.5 show marginal productivity of infrastructure in 1970 and in 1990, respectively, calculated from long persistent model-1. It is decomposed into three types: instantaneous productivity for own region (MPG) calculated by the first term of Eq. (12.17); the sum of instantaneous productivity for the other regions (spill-over effect, MPGS) calculated by the second term of Eq. (12.17); and the sum of persistent productivity for own region (long persistent effect, MPGM) calculated by Eq. (12.18). In order to calculate MPGM in 1990, GRP information in 1999 and 2000 is required. In the calculation, we assumed that GRP after 1998 is identical to

1998. In 1970, MPG and MPGM are relatively high in industrialized areas, such as Tokyo, Osaka, Nagoya, Hiroshima, and Fukuoka, while MPGS is high in the surroundings of industrialized areas. In 1990, MPG and MPGM are high in Aichi and Shizuoka, where the car industries (Toyota, Honda and Yamaha) are clustered, and in the prefectures surrounding Tokyo and Osaka. A comparison of 1970's with 1990's results show that prefectures with high MPG are distributed throughout Japan, and that prefectures with high MPGS can commonly be found around high MPG prefectures. MPGM is still high around the three metropolitan regions (Tokyo, Nagoya, and OSAKA), while MPGM in western prefectures between Osaka to Fukuoka is significantly decreased.

Figure 12.6 shows the total marginal productivity of infrastructure (TMPG) in 1970, 1980, and 1990, which is the sum of MPG, MPGS and MPGM for each cross-section. In Fig. 12.6, bi-polar distribution of TMPG around Tokyo and Osaka appears in 1970. In 1990, high TMPG prefectures successively appear between Tokyo and Osaka, including Nagoya. Compared to those of conventional studies, values of TMPG of the long persistent model are slightly lower (Ejiri et al. 2000). The estimated marginal productivity obtained by the long persistent model is lower than in conventional studies. However, the decreasing trend of marginal productivity of infrastructure, which is often pointed out in conventional studies, is not observed, but rather the increasing trend is observed. Since diagnostic tests for residuals show that the long persistent model does not seriously suffer from serial- and spatial- autocorrelation problems, the obtained results seem to be credible. Such a difference in longitudinal and cross-sectional behavior of infrastructure productivity stems from the difference in model specification. Neglecting the persistent effect of past infrastructure in marginal productivity would result in overestimation of present infrastructure productivity. In our result, one third of marginal productivity stems from past infrastructure, hence its contribution should not be neglected.

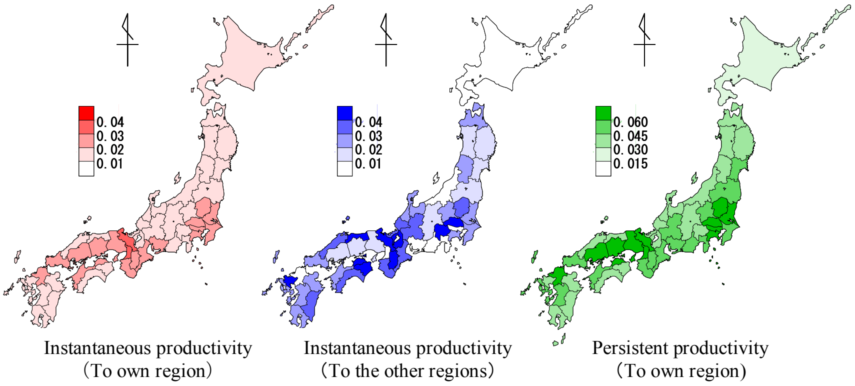


Fig.12.4. Marginal productivity of infrastructure in 1970

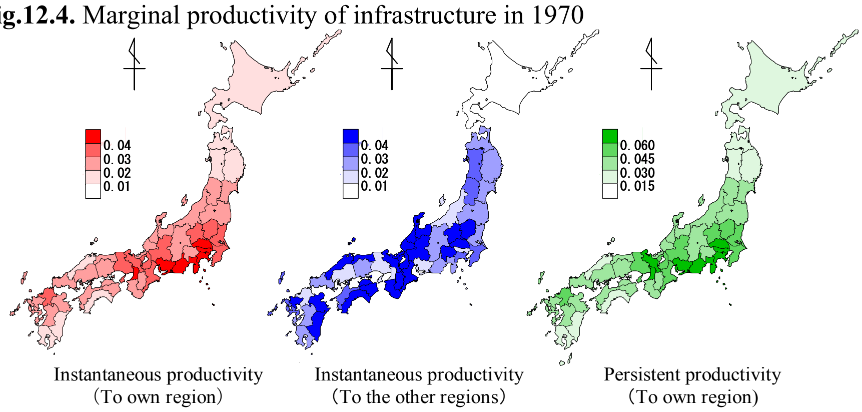


Fig. 12. 5. Marginal productivity of infrastructure in 1990

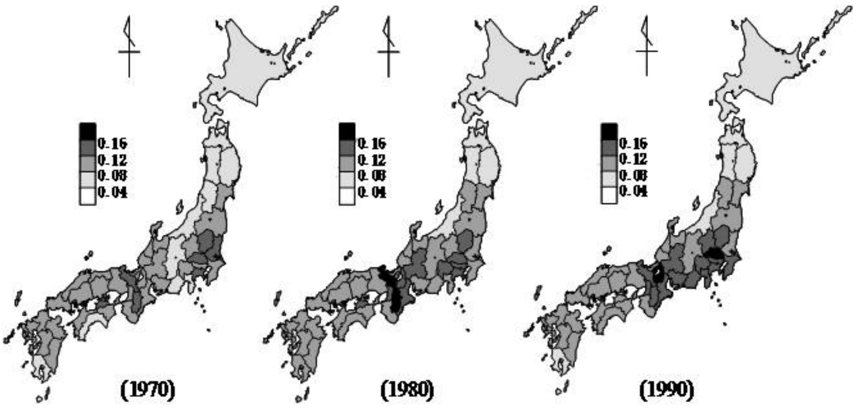


Fig. 12.6. Total marginal productivity of infrastructure

Figure 12.7 shows the share of three types of marginal productivity in total marginal productivity averaged for all regions. The averaged shares are around 0.33 to 0.34 for all cross-sections, hence the three types of marginal productivity are almost at the same level. The share of MPG is highest in 1968, and continuously decreases up to 1998. On the other hand, the shares of MPGM and MPGS continuously increase from 1965 to 1998. Therefore, long persistent productivity of infrastructure is stable and has increased for whole regions. The growth of MPGS would correspond to the globalization of domestic economic activities, and such trends will not change in the future. The average delay of infrastructure productivity can be calculated from Eq. (12.7) .Using the estimated $\hat{\varphi}$ and $\hat{d}_G$ of long persistent model-1, the average delay of infrastructure productivity is about 5.78 years. As discussed in Sect. 12.2.1, infrastructure productivity of the long persistent model under the infinite lags diverges to positive infinite. Note that the calculated 5.78 years implies approximated delay truncated with 10 years. If longer series of infrastructure stock data were available, the average delay would be more largely estimated.

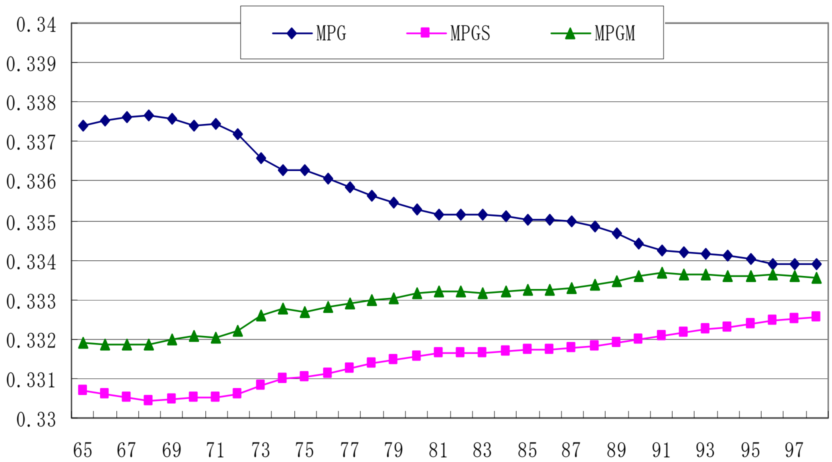


Fig.12.7. Shares of components in total marginal productivity, averaged for all regions

Most conventional studies propose to measure infrastructure productivity for own regions (Ejiri 2000). However during the planning stage, infrastructure is expected to have productivity effect extending over a long term and to global regions. An empirical analysis for productivity measurement neglecting the effect to the future and to neighboring regions would result in underestimation of infrastructure productivity. Of course, it is impossible to conclude immediately the extent of such underestimation. More careful studies about the long persistent effect or spatial spillover effect should be undertaken, followed by intense discussion of the productivity of infrastructure.

12.4 Conclusions

In order to measure infrastructure productivity with lasting effects for the future, we formulated a production function with a long persistent effect, and the proposed model was applied to measure infrastructure productivity in Japan from 1965 to 1998. The estimated model showed that a positive and significant long persistent effect was observed for infrastructure and for stochastic error term. Based on the estimated production function with a long persistent effect, the longitudinal change in marginal productivity of infrastructure was estimated. The estimated series of marginal productivity gave us the novel implication that the marginal productivity of infrastructure is longitudinally increasing, which is different from the implication of conventional approaches measuring infrastructure productivity only within the own region. The share of persistent productivity over the total infrastructure productivity composed of instantaneously to the own, instantaneously to the other regions and lasting for the future to the own region was about one-third throughout the data duration. Because of its stable share, the persistent productivity effect of infrastructure should not be neglected. In other words, considerable underestimation of infrastructure productivity would occur if a policy maker neglected the long persistent productivity effect of infrastructure.

Our approach is significant in that it shows an approach to measure the long persistent productivity effect of infrastructure by using a long memory model, but there remain further issues to be resolved. First, the long persistent property should expand for spatial correlation structure. Our model considered only longitudinal correlation structure, but the spatial or spatial-temporal correlation structure was neglected. The development of a model specification and estimation procedure is one important issue. Second, points of trend shift in deterministic time-trend were exogenously set in our approach based on retrospective assessment in the economic situation. The estimated time trend parameters were the appropriate signs due to the ad-hoc shift setting. However, in long persistency in empirical analysis, many papers stress the importance of properly specifying regime switching in empirical modeling in order to improve long persistent (fractional integration) parameters (Gonzalo J and Lee T 1998; Cheung Y and Lai K 2001; Patel A and Shoesmith G 2004). Much careful treatment is required for deterministic time-trend. A further issue is to develop a statistical test in regime switching applicable for multiple time-series data. Third, about the spatial spillover effect, our model only considered the infrastructure spillover effect. However, the knowledge of spillover effect embedded in regional output would affect social capital spillover. In applied spatial econometrics, a production function approach explicitly considering spatial heterogeneity is proposed as regional specific fixed or random effect based on panel data analysis (Baltagi et al. 2001; Juhn et al. 2003). Panel data analysis with the long persistent property is another target issue. Finally, technological innovation is modeled as a deterministic trend in our model. Investment in infrastructure would induce technological innovation. Therefore, investment in infrastructure, long persistent property in stochastic error term, and deterministic trend would have correlations. . Studies on proper estimation procedure about production function based on endogenous growth theory are required.

All the above issues are related with an appropriate specification of spatial-temporal process in the long persistent model, so that a theoretical approach for the spatial-temporal stochastic process is important.

References

Agiakloglou C, Newbold P (1994) Lagrange multiplier tests for fractional difference. Journal of Time Series Analysis 15, pp 253-262

Almon S (1965) The distributed lag between capital appropriation and expenditures. Econometrica 33, pp 178-196

Aschauer D (1989) Is public expenditure productive? Journal of Monetary Economics 23, pp 177-200

Baltagi B, Song S, Jung B (2001),The unbalanced nested error component regression model. Journal of Econometrics 101, pp 357-381

Barkoulas T, Baum F, Chakraborty A (2001) Waves and persistence in merger and acquisition activity. Econometrics letters 70, pp 237-243

Basu S, Fernald J (1997) Returns to scale in U.S. production: Estimates and implications. Journal of Political Economy 105, pp 249-283

Beran J (1992) Statistical methods for data with long-range dependence. Statistical science 7, pp 404-427

Bhardwai G , Swanson N (2006) An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series., Journal of Econometrics 121, pp 539-578

Box,P, Jenkins M, Reinsel C (1994) Time series analysis/forecasting and control. 3rd edn, Prentice hall

Box-Steffensmeier M, Tomlinson R (2000) Fractional integration methods in political science. Electoral studies 19, pp 63-76
Cheung Y, Lai K (2001) Long memory and nonlinear mean reversion in Japanese yen-based real exchange rate. Journal of International Money and Finance , 20, pp 115-132
Chung F, Baillie R (1993) Small sample biases in conditional sum-of-squares estimators of fractionally integrated ARMA process. Empirical Economics 18, pp 791-806
Davidson J, Sibbertsen P (2005) Generating schemes for long memory processes: regimes, aggregation and nonlinearlity. Journal of Econometrics 128, pp 253-282
Dfrenot G, Guegan D, Peguin-Feissolle A (2005) Long-memory dynamics in a SETAR mode – applications to stock markets. Journal of Financial Markets, institutions & money 15, pp 391-406
Doi T (2003) http://www.econ.keio.ac.jp/staff/tdoi/index-J.html (in Japanese)
Duggal V, Saltzman C, Klein L (1999) Infrastructure and productivity: a nonlinear approach. Journal of Econometrics 92, pp 47-74
Ejiri R, Okumura M, Kobayashi K (2001), Productivity of social capital and economic growth: state of the art, Journal of infrastructure planning and management, 688/IV-53, pp 75-87 (in Japanese)
Everaert G, Heylen F (2001) Public capital and productivity growth: evidence for Belgium, 1953-1996. Economic Modeling 18, pp 97-116
Godfrey L (1979) Testing the adequacy of a time series model, Biometrica, 66, pp 62-72
Gonzalo J, Lee T (1998) Pitfalls in testing for long run relationships. Journal of Econometrics 86, pp129-154
Gourierouex C, Monfort A (1997) Time series and dynamic models. Cambridge University Press
Granger CWJ, Joyeux R (1980) An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis 1, pp 15-29
Haughwout A (2002) Public infrastructure investments, productivity and welfare in fixed geographic areas. Journal of Public Economics 83, pp 405-428
Henry T, Olekalns N (2002) Does the Australian dollar real exchange rate display mean reversion? Journal of International Money and Finance, 21, pp 651-666
Holtz-Eakin D, Lovely M (1996) Scale economies, returns to variety, and productivity of public infrastructure. Regional Science & Urban Economics 26, pp 105-123
Hosking J (1981) Fractional differencing. Biometrika 68, pp 165-176
Hosking J (1996) Asymptotic distributions of the sample mean, auto-covariances, and autocorrelations of long-memory time series. Journal of Econometrics 73, pp 261-284
Igresias P, Jorquera H, Palma W (2005) Data analysis using regression models with missing observations and long-memory: an application study. Computational Statistics & Data Analysis, (article in press)
Jhun M, Song S, Jung B (2003) BLUP in the nested regression model with serially correlated errors. Computational Statistics & Data Analysis 44, pp 77-88

Lucus R (1988) On the mechanics of economic development. Journal of Monetary Economics 22, pp 3-42

Michelacci C (2004) Cross-sectional heterogeneity and persistence of aggregate fluctuations. Journal of Monetary Economics 51, pp 3121-1352

Minotani C (1995) Differencing and Integral for economic analysis.Taga-Shuppan

Moran P (1948) The interpretation of statistical maps, Journal of the Royal Statistical Society B 10, pp 243-251

Munnel H (1992) Policy Watch: infrastructure investment and economic growth Journal of Economic Perspectives 6, pp 189-198

Patel A, Shoesmith G (2004)Term structure linkages surrounding the Plaza and Louvre accords:evidence from Euro-rates and long-memory components. Journal of Banking & Finance 28, pp 2051-2075

Robinson P, Hidalgo F (2003) Time-series regression with long-range dependence. In: Time series with long memory, Oxford, pp 305-333

Romer M (1986) Increasing returns and long-run growth, Journal of Political Economy 94, pp 1002-1037

Shiller J (1973) A distributed lag estimator derived from smoothness priors. Econometrica 41, pp 775-778

Smith J, Taylor N, Yadav S (1997) Comparing the bias and misspecification in ARFIMA models. Journal of Time Series Analysis 18, pp 507-527

Strum E, de Haan J (1995) Technological change and aggregate production function. The Review of Economics and Statistics 39, pp 312-320

Sturm E (1998) Public capital expenditure in OECD countries. In: The causes and impact of the decline in public capital spending, Edward Elger

Tanaka K (1999) The non-stationary fractional unit root, Econometric Theory 15, pp 549-582

Tsukai M, Ejiri R, Okumur M, Kobayashi K (2002) Productivity of Infrastructure and spillover effects, Journal of Infrastructure Planning and Management 716/IV-57, pp 53-67 (in Japanese)