Road infrastructure and TFP in Japan after the rapid growth: A nonstationary panel approach

Koike, Atushi and Sakaguchi, Takuhiro and Seya, Hajime

Kobe University, Kobe University, Kobe University

15 March 2022

Road infrastructure and TFP in Japan after the rapid growth: A nonstationary panel approach

By ATSUSHI KOIKE, TAKUHIRO SAKAGUCHI, AND HAJIME SEYA[1]

Abstract

This study investigates the relationship between road infrastructure stock and total factor productivity (TFP) using R-JIP2017, a database of productivity by industry for each prefecture in Japan, which allows us to estimate TFP with considering the quality of inputs. Specifically, using the growth accounting method, we estimated TFP for each industry in each prefecture from 1972 to 2012, after the period of high economic growth. Afterwards, we conducted a panel data analysis to explain the estimated TFP by road stock. The results of a panel unit root test indicated the existence of unit roots in the road infrastructure stock. Therefore, unlike many previous studies, a panel autoregressive distributed lag (ARDL) model was used as the empirical model, considering the nonstationarity of the variables. The results of the analysis indicated that road stock had a positive and significant relationship with TFP at the 5% level in the majority of industries, even after the period of rapid economic growth. Further, we found that the two-way fixed effects model, which does not consider the non-stationarity of road infrastructure stock, could produce misleading results.

Keywords:
Total factor productivity (TFP); R-JIP; Road infrastructure; ARDL model

JEL: R11; R40; R42

Acknowledgements

This study was supported by JSPS KAKENHI Grant Number 20H02274. The corresponding author states on behalf of all authors that there is no conflict of interest.

[1] Hajime Seya (corresponding author)
Associate Professor, Departments of Civil Engineering, Graduate School of Engineering
Kobe University, Japan
Address: 1-1, Rokkodaicho, Nada, Kobe, Hyogo, 657-8501, Japan
E-mail: hseya@people.kobe-u.ac.jp

1. Introduction

In economic growth theory and development economics, many theoretical studies have been conducted based on the premise that infrastructure is a significant driving force for economic growth. In particular, in the theory of endogenous economic growth, infrastructure plays a major role in the sustainable growth of the economy, with a growing body of literature accumulated in recent decades investigating how government investment in infrastructure affects economic growth and individual welfare.

Furthermore, there have been many empirical studies on the relationship between infrastructure and economic growth (Magazzino and Maltese, 2021). In the U.S., productivity growth has been sluggish since the early 1970s, and many economists have conducted research to clarify the causes of this productivity paradox. Aschauer (1989) argued that a delay in the development of infrastructure could be the answer to the productivity paradox. Since his argument can be linked to expansionary fiscal policies, it attracted the attention of many economists, and a number of empirical studies were conducted using different methods and datasets to validate his conclusions. However, the estimates obtained from these studies varied greatly depending on the methods and datasets used, with no clear conclusion on whether infrastructure can resolve the productivity paradox (however, see the cautionary note by Munnell (1992))[2].

With this background, this study analyzes the relationship between road stock in monetary terms and total factor productivity (TFP) using the 2017 Japan Industrial Productivity (R-JIP 2017) Database (Tokui et al.; 2013, 2019). Specifically, using the growth accounting method introduced by Hulten et al. (2006) and our panel dataset, we estimated TFP for each industry in each prefecture from 1972 to 2012. Then, we conducted panel data analysis to explain the estimated TFP by road stock. In comparison with previous studies, the novelty of this study lies in the following points:

1) We estimated TFP based on the growth accounting frame introduced by Hulten et al. (2006), and thereby derived estimates without specifying the functional form of a production function in a certain degree.
2) We performed an analysis by industry over a long period of time (1972-2012), after the period of high economic growth from 1954 to 1973. The case study of Japan, the world's most aged society, after its rapid economic growth will be of great help to other countries, including developing countries.
3) We explicitly considered the possibility of the existence of unit roots is in the road infrastructure stock and TFP panel data, and thus, conducted our analysis in an attempt to eliminate the problem of spurious correlation.

Regarding point 3), the results of the panel unit root test indicated the existence of unit roots in the road infrastructure stock and TFP of some sectors, and therefore, unlike most previous studies, we used the panel autoregressive distributed lag (ARDL) model, which considers the non-stationarity of variables, as our empirical model.

The results of analysis showed that road stock has a positive and significant relationship with TFP at the 5% level in most industries. Further, we found that the two-way fixed effects (2FE) model, which does not consider the non-stationarity of road infrastructure stock, can produce misleading results.

[2] According to Välilä (2020), Estache and Fay (2007) attributed this inconclusiveness to the presence of possibly bidirectional network effects. All these characteristics render the links between infrastructure and growth complex and difficult to measure.

The remainder of the paper is structured as follows. Section 2 reviews the existing literature on infrastructure and economic growth. Section 3 constructs the empirical model for the study. Section 4 constructs the panel dataset and presents the empirical results. Finally, Section 5 concludes this paper.

2. Literature review

Theoretical studies on infrastructure and economic growth have traditionally focused on the endogenous economic growth theory in macroeconomics. In the 1950s, the neoclassical theory of economic growth originated with the Solow-Swan model developed by Solow (1956) and Swan (1956). The Solow-Swan model is characterized by exogenously determined technological progress and savings rate. The endogenization of the savings rate was solved by Cass (1965) and Koopmans (1965), who advanced the research of Ramsey (1928), and later by Kydland and Prescott (1982), who developed the real business cycle (RBC) theory and Kydland and Prescott (1982) who introduced the dynamic stochastic general equilibrium (DSGE), which is now widely used as a framework for theoretical and empirical analysis. The endogenization of technological progress was developed by Romer (1986) and Lucas (1988) as a theory of endogenous economic growth, which expresses sustained economic growth by modeling the process of accumulation of knowledge, human capital, social infrastructure, and R&D. As the development of the endogenous growth theory was based on the question of whether the disparity in economic growth rates among regions or countries would converge, which is closely related to development economics, many studies have been conducted focusing on the role of infrastructure as an engine of economic development.

Among the researchers who developed the theory of endogenous economic growth, Barro (1990) was a pioneer in explicitly including the public sector in the model. He constructed a model in which the government finances spending with income taxes while being included in the private sector's production function as a public good. Using this model, Barro (1990) showed that the maximization of the economic growth rate coincides with the maximization of the welfare level of a representative individual. Futagami et al. (1993) modified Barro's (1990) model, arguing that the stock of public capital, rather than the flow of capital, should contribute to private production. They showed that, unlike Barro (1990), the tax rate that maximizes the welfare of a representative individual is lower than that maximizes the economic growth rate. Both Barro (1990) and Futagami et al. (1993) explicitly incorporated public sector activities into their models and analyzed the relationship with economic growth. These studies were in line with Aschauer (1989), that is, they focused upon the role played by infrastructure in economic growth. In fact, there could be said to be a mutual relationship, including citations of Aschauer (1989) in Barro (1990).

Aschauer (1989) conducted a pioneering empirical study on infrastructure and economic growth. He examined the relationship between aggregate productivity and stock-flow government spending variables and found that 1) non-military public capital stock is significantly more important in determining productivity than the flow of non-military and military spending, 2) the relationship between military capital and productivity is weak, and 3) "core" infrastructure such as roads, highways, airports, transportation, sewerage, and water supply explains productivity the most. He argued that delayed development of public capital stock caused the slowdown in production growth in the U.S. in the early 1970s. Munnell (1992) provided a defense against the three major criticisms of research on infrastructure and economic growth since Aschauer (1989). These criticisms deserve attention because they are often raised even today in the 2020s; they are as follows:

1) The existence of spurious correlation due to common trends between output and social infrastructure data.
2) Many studies differ in their estimates of the coefficients representing the impact of infrastructure on output.
3) The existence of reverse causality from output to social infrastructure.

With regard to 1), it is pointed out that the first difference method for non-stationary time-series data destroys the long-run equilibrium relationship. Thus, instead of just first-differencing, the variables should be tested for co-integration, adjusted, and estimated accordingly. This study attempts to estimate the long-term relationship while avoiding first differences by making use of non-stationary panel testing and estimation methods that have been developed in recent years.

With regard to 2), although Aschauer (1989) is often regarded as the pioneer in the study of infrastructure and productivity, it should be noted that Hulten and Schwab (1984) conducted a regional study of U.S. manufacturing industries earlier. They were the first to link the prior separate studies on the relationship between infrastructure deterioration, urban environmental degradation, and the economic performance of aging capital stock in the Snow Belt region (East Coast to Midwest) and the slowdown in overall U.S. productivity growth in the 1970s and 1980s. The results showed that the growth rate of TFP was higher in the Snow Belt region (1.80) than in the Sun Belt region (1.61), thus refuting the hypothesis that the slowdown in economic growth in the Snow Belt region was due to a slowdown in productivity growth caused by deteriorating infrastructure. In addition, they argued that the growth rate of labor productivity was almost the same in the Snow Belt and Sun Belt regions, further supporting this result. The results of their study preceded those of Aschauer (1989), who found a positive effect of infrastructure on economic growth, but they represent an important rebuttal. Thus, estimates of the impact of infrastructure on economic output and economic growth in previous studies have varied widely in terms of both sign and magnitude since the early stage of research. Among recent meta-analyses, Melo et al. (2013), who focused on transportation infrastructure, found that the productivity effect of transport was higher in the U.S. than in European countries, while Elburz et al. (2013) found that U.S. studies are more likely to find a negative impact from infrastructure on growth. Thus, we found no consistent conclusions in recent meta-analyses.

Munnell (1992), on the other hand, disputed this typical view that "no consensus has yet been reached." First, she argued that the large discrepancy in the estimated coefficients does not negate the positive impact of infrastructure on production, given that most public capital does not contribute much to production, such as environmental measures or quality of life improvements. In addition, she pointed out that the variation in the estimated coefficients is mainly a result of the fact that the effect of infrastructure becomes smaller as the unit of observation in the comparison studies becomes smaller, from national to state and from state to city, and that a relatively uniform positive effect is observed when the spatial unit is controlled. In other words, it is pointed out that one cannot capture all of the payoffs to infrastructure investment by focusing on a small geographic area. This view was supported by the latest meta-analysis results of Holmgren and Merkel (2017)[3], where the coefficients of the region dummy variable, 1 in the case of regional disaggregation of data, and 0 otherwise, was 0.0808 (p =0.012). Similar results were also confirmed by Melo et al. (2013) and Elburz et al. (2013). The importance of the spatial unit

[3] With heteroscedasticity-robust standard errors.

in analysis of the relationship between infrastructure and economic growth has been pointed out in new economic geography literature (Venables et al., 2014).

Baird (2005) further found that highways have local negative spillover effects that arise from economic activities being drawn to infrastructure-rich locations at the expense of adjacent areas. Recently, there have been extensive studies quantifying this spillover. For example, Deng (2013) pointed out that the variation in effects in previous studies can be explained not only by differences in spatial units, but also by differences in contexts including spatial units (e.g., stage of development of the subject and the time period), differences in measurement units (e.g., industrial divisions), differences in the type and quality of infrastructure, and differences in modeling methods. To clarify these points in detail, it is important to accumulate case study analyses with standard datasets such as R-JIP.

With regard to 3), utilization of instrumental variables or exogenous shocks could be considered (Kawaguchi et al., 2009). It can also be dealt with using other econometric models, such as vector autoregression (VAR) (Kawakami and Doi, 2004), difference generalized method of moments (difference GMM) (Na et al., 2013), System GMM (Barzin et al., 2018), and dynamic ordinary least squares (DOLS) (Okubo, 2008). Nevertheless, using state-level data, Munnell (1992) argued that reverse causality is not necessarily a major problem in the estimation of coefficients, although she admits that it is not a perfect solution to this problem. Indeed, this argument may change with time. For example, preferential investment in less developed regions can be a cause of reverse causality.

While much empirical studies on infrastructure and economic growth has been conducted in the U.S. and Europe, there has been a clear increase in the literature on Asia in recent years (see Magazzino and Maltese, 2021). Here we turn our attention to studies in Japan[4]. In Japan, empirical studies on this topic have been conducted intensively, especially in the 1990s and 2000s[5]. First, there is the pioneering study by Mera (1973). Mera (1973) divided the entire country into 9 regions and social infrastructure into 4 sectors, covering the years 1954 to 1963, and then examined the productivity effects of social infrastructure according to the primary, secondary, and tertiary sectors. The results demonstrated that the productivity effect of social infrastructure is positive in all sectors. As for more recent studies, we look at Miyara and Fukushige (2008), Tsukai and Kobayashi (2009), Nakagishi and Yoshino (2016), and Miyagawa et al. (2013), among others. Miyara and Fukushige (2008) used a Cobb-Douglas production function to examine the productivity of public infrastructure per prefecture from 1976 to 1997. The results suggested that the productivity of public infrastructure differs between prefectures, and that transportation contributes to production in prefectures with many large establishments, whereas congestion reduces the productivity of transportation. Furthermore, water systems and telecommunications might contribute to production in secondary industries. Tsukai and Kobayashi (2009) measured infrastructure productivity with *lasting effects* for the future. They 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. Nakagishi and Yoshino (2016) examined the productivity of public capital using a translog production

[4] For a study on the relationship between infrastructure and economic output and growth using aggregated data at the macro level, see the reviews of Straub (2011) and Välilä (2020). If microdata can be used, various productivity growth rate decomposition methods can be applied (e.g. Petrin and Levinsohn, 2012); however, the bar of obtaining microdata is still high, and this study uses macro level aggregated data (R-JIP). Unlike microdata, which is $N >> T$ (N is the number of units and T is the number of time points), it is important to consider non-stationarity in macrodata (Baltagi, 2005).

[5] Most of them have been published in Japanese domestic journals and are detailed in a comprehensive review by Ejiri et al. (2001).

function for each prefecture grouped by region between 1975 and 2010. They found that in the secondary and tertiary industries, the productivity effect of public capital has been significantly positive throughout the estimated period and has been present in recent years. Miyagawa et al. (2013) conducted a study using R-JIP, and observed the productivity effect of public capital, especially after the collapse of the bubble economy[6]. However, the focus of their study was on regional analysis, and not on detailed sectoral analysis.

In this study, we attempt to add new findings to these previous studies. By estimating TFP based on the growth accounting frame developed by Hulten et al. (2006), we attempt to derive estimates of infrastructure effects without specifying the functional form of a production function in a certain degree. In addition, we analyze the effect of infrastructure by industry from 1973 to 2012, which roughly corresponds to the next period of Japan's rapid economic growth. Such long-run empirical analysis is important because it allows the application of a panel time series approach that can estimate long-term impacts and explicitly account for non-stationarity issues. Note that except for Okubo (2008), studies of Japan did not consider non-stationarity. Since panel data can be regarded as an extension of time-series data, it is necessary to deal with the concern of spurious correlation on the basis of statistical tests as in the case of time-series analysis (Baltagi, 2005). In this study, we use the panel ARDL model as our empirical model, which is generic in the sense that it can be used even when the I(0): integrated variables of order zero and I(1): integrated variables of order one are mixed (Pesaran et al., 2001). We show that misleading results can be obtained using the usual 2FE model that does not consider the non-stationarity of road infrastructure stock.

Studies on the relationship between infrastructure and economic growth using the ARDL model include Calderón et al. (2015), Alam et al. (2020), and Khanna and Sharma (2021). Using balanced panel data comprising annual information on output, physical capital, human capital, and infrastructure capital for 88 industrial and developing countries over the period (1960-2000), Calderón et al. (2015) estimated the long-run elasticity of output with respect to the synthetic infrastructure index. They found that it ranges between 0.07 and 0.10. Alam et al. (2020) also found that transport infrastructure has a long-run positive impact on economic development in Pakistan. Khanna and Sharma (2021) tested the effects of public infrastructure on the TFP of Indian manufacturing industries. The productivity effects of infrastructure were estimated using an ARDL model with cross-sectionally augmented pooled mean group (PMG) estimator. The results of the analysis confirmed the presence of a positive and sizeable effect of infrastructure on manufacturing productivity.

3. The Model

In this study, the method developed by Hulten et al. (2006) for analyzing productivity through growth accounting using industry- and region-specific data is used. This method builds upon the work in Hulten and Schwab (1984), who performed a similar verification using a U.S. dataset. The results of their analysis showed that increases in roads and electricity generation explain about half of productivity growth. A unique feature of their studies was that they explicitly linked infrastructure and productivity growth in the context of growth accounting. In the following, we outline Hulten et al. (2006).

[6]The collapse of the bubble economy is often considered to be the recessionary period lasting from March 1991 to October 1993.

3.1. Infrastructure and Production Functions

Here, we consider the following format for the production function in a given industry:

$$Q_{i,t} = A(B_{i,t}, t)F_i\left(K_{i,t}, L_{i,t}, M(B_{i,t})\right); \tag{1}$$

where i is the region index, t is the time (year) index, $Q_{i,t}$ is the total output, $A(B_{i,t}, t)$ is Hicks-neutral technical change, $B_{i,t}$ represents the infrastructure stock, $K_{i,t}$ represents the private capital stock, $L_{i,t}$ is the labor input, and $M(B_{i,t})$ represent the intermediate inputs (the industry index is omitted to avoid complicating the equation). Here, the infrastructure stock $B_{i,t}$ influences production through two channels: the effect on output through intermediate inputs $M(B_{i,t})$ and the effect on production through term $A(B_{i,t}, t)$ expressing the Hicks-neutral technical change. As discussed in detail in Gibbons and Overman (2009), compared to $M(B_{i,t})$, the effect of $B_{i,t}$ on $A(B_{i,t}, t)$ is often unclear. The former is the straightforward effect of reducing logistics costs (called the market-mediated effect by Hulten et al. (2006)), and the latter is due to factors such as the geographical relocation of firms and changes in industrial structure (when i is a regional unit, as in this study, and not a company unit) adding to pure productivity growth (called the indirect effect by Hulten et al. (2006)). Although vigorous efforts have been made in areas such as quantitative spatial economics to isolate the impact of pure productivity growth (Redding and Rossi-Hansberg, 2017), isolation requires firm-level microdata, for which the bar to obtain is still high in many countries, including Japan.

Assuming multiplicative structure $A_{i,0}e^{\lambda_i t}B_{i,t}^{\gamma}$ for the term $A(B_{i,t}, t)$ expressing Hicks-neutral technical change, Eq. (1) could be expressed as follows:

$$Q_{i,t} = A_{i,0}e^{\lambda_i t}B_{i,t}^{\gamma}F_i\left(K_{i,t}, L_{i,t}, M(B_{i,t})\right); \tag{2}$$

where $A_{i,0}$ is productivity in the base year, $e^{\lambda_i t}$ is a time trend term, and $B_{i,t}^{\gamma}$ represents the infrastructure stock term.

The object of interest in this study is the parameter γ related to the infrastructure stock $B_{i,t}$. Hulten et al. (2006), following Solow (1957), attempted to estimate γ via growth rate estimation of productivity. In Eq. (1), the term expressing productivity is $A(B_{i,t}, t)$, which is computed as the ratio of total output $Q_{i,t}$ to the inputs used to produce that output $F_i\left(K_{i,t}, L_{i,t}, M(B_{i,t})\right)$. Unlike Solow (1957), since F_i includes intermediate inputs, we can use this ratio as a measure of total productivity (TP), which can be defined as $\mathrm{TP}_{i,t} \equiv Q_{i,t}/F_i\left(K_{i,t}, L_{i,t}, M(B_{i,t})\right)$. In Eq. (2), $\mathrm{TP}_{i,t} = A_{i,0}e^{\lambda_i t}B_{i,t}^{\gamma}$.

Next, we consider the rate of change of $\mathrm{TP}_{i,t}$. The data used for estimation here are discrete; however, we first assume that continuous data are available, and then consider discretization as an approximation. If we take the natural logarithm of both sides of Eq. (1) and differentiate it with respect to the variable t , it could be rewritten as follows.

$$\frac{\dot{Q}_{i,t}}{Q_{i,t}} = \frac{\dot{A}(B_{i,t},t)}{A(B_{i,t},t)} + \frac{\frac{\partial F_i}{\partial K_{i,t}}K_{i,t}}{F_i}\frac{\dot{K}_{i,t}}{K_{i,t}} + \frac{\frac{\partial F_i}{\partial L_{i,t}}L_{i,t}}{F_i}\frac{\dot{L}_{i,t}}{L_{i,t}} + \frac{\frac{\partial F_i}{\partial M(B_{i,t})}M(B_{i,t})}{F_i}\frac{\dot{M}(B_{i,t})}{M(B_{i,t})}; \tag{3}$$

where $\dot{A}(B_{i,t}, t) = \frac{dA(B_{i,t},t)}{dt}$; $\dot{K}_{i,t} = dK/dt$; $\dot{L}_{i,t} = dL/dt$; $and\ \dot{M}(B_{i,t}) = dM(B_{i,t})/dt$. In addition, the function $F_i(\cdot)$ is sufficiently smooth and all variables are assumed to be sufficiently smooth for t. To estimate the $\frac{\dot{A}(B_{i,t},t)}{A(B_{i,t},t)}$ term, we assume that each firm acts in the factor of production market with price as a given (Price Taker assumption). In addition, $F_i(\cdot)$ is assumed to be first-order and linearly homogenous with respect to the argument.

$$F_i\left(K_{i,t}, L_{i,t}, M(B_{i,t})\right) = \frac{\partial F_i}{\partial K_{i,t}} K_{i,t} + \frac{\partial F_i}{\partial L_{i,t}} L_{i,t} + \frac{\partial F_i}{\partial M(B_{i,t})} M(B_{i,t}). \tag{4}$$

Assuming that the costs of production factors *K*, *L*, and *M* are respectively p_K, p_L, p_M, they may be written as follows from the minimization conditions of total cost $p_K K_{i,t} + p_L L_{i,t} + p_M M(B_{i,t})$.

$$\frac{p_K}{\frac{\partial Q_{i,t}}{\partial K_{i,t}}} = \frac{p_L}{\frac{\partial Q_{i,t}}{\partial L_{i,t}}} = \frac{p_M}{\frac{\partial Q_{i,t}}{\partial M(B_{i,t})}}. \tag{5}$$

By substituting Eq. (5) into Eq. (3), we obtain

$$\frac{\dot{A}(B_{i,t},t)}{A(B_{i,t},t)} = \frac{\dot{Q}_{i,t}}{Q_{i,t}} - \pi_{K_{i,t}} \frac{\dot{K}_{i,t}}{K_{i,t}} - \pi_{L_{i,t}} \frac{\dot{L}_{i,t}}{L_{i,t}} - \pi_{M_{i,t}} \frac{\dot{M}(B_{i,t})}{M(B_{i,t})}. \tag{6}$$

This is provided that $\pi_{K_{i,t}} = \frac{p_K K_{i,t}}{p_K K_{i,t} + p_L L_{i,t} + p_M M(B_{i,t})}$; $\pi_{L_{i,t}} = \frac{p_L L_{i,t}}{p_K K_{i,t} + p_L L_{i,t} + p_M M(B_{i,t})}$; $\pi_{M_{i,t}} = \frac{p_M M(B_{i,t})}{p_K K_{i,t} + p_L L_{i,t} + p_M M(B_{i,t})}$.

The discretization procedure for Eq. (6) is explained next. To calculate TP (or TFP) using the growth accounting method, it is necessary to estimate TP using a discrete set of data such as annual data. Therefore, the term expressed using the derivative according to *t* in Eq. (6) is discretely approximated by the difference. In addition, $\pi_{K_{i,t}}$, $\pi_{L_{i,t}}$, $\pi_{M_{i,t}}$ which represents the cost share of each production factor, is discretely approximated using the average value of the cost share of the previous period and the cost share of the current period. This can be written as:

$$\Delta \ln\left(A(B_{i,t}, t)\right) = \Delta \ln(Q_{i,t}) - \bar{\pi}_K \Delta \ln(K_{i,t}) - \bar{\pi}_L \Delta \ln(L_{i,t}) - \bar{\pi}_M \Delta \ln\left(M(B_{i,t})\right); \tag{7}$$

provided that

$$
\begin{aligned}
\Delta\ln(A(B_{i,t},t)) &= \ln\left(A(B_{i,t},t)\right) - \ln\left(A(B_{i,t-1},t-1)\right);\\
\Delta\ln(Q_{i,t}) &= \ln(Q_{i,t}) - \ln(Q_{i,t-1});\\
\Delta\ln(K_{i,t}) &= \ln(K_{i,t}) - \ln(K_{i,t-1});\\
\Delta\ln(L_{i,t}) &= \ln(L_{i,t}) - \ln(L_{i,t-1});\\
\Delta\ln(M(B_{i,t})) &= \ln(M(B_{i,t})) - \ln(M(B_{i,t-1}));\\
\bar{\pi}_{K_{i,t}} &= \frac{\pi_{K_{i,t}} + \pi_{K_{i,t-1}}}{2};\\
\bar{\pi}_{L_{i,t}} &= \frac{\pi_{L_{i,t}} + \pi_{L_{i,t-1}}}{2};\\
\bar{\pi}_{M_{i,t}} &= \frac{\pi_{M_{i,t}} + \pi_{M_{i,t-1}}}{2}.
\end{aligned}
$$

3.2. Estimation of TP

Let us consider the estimation of $\mathrm{TP}_{i,t}$. The variables on the right-hand side of Eq. (7) are observable from statistical data, and by calculating the right-hand side, we obtain $\Delta\ln(\mathrm{TP}_{i,t})$ as:

$$\Delta\ln(\mathrm{TP}_{i,t}) = \ln(\mathrm{TP}_{i,t}) - \ln(\mathrm{TP}_{i,t-1}) \approx \frac{\mathrm{TP}_{i,t}-\mathrm{TP}_{i,t-1}}{\mathrm{TP}_{i,t-1}}. \tag{8}$$

Therefore, the right-hand side approximately expresses the rate of change in $\mathrm{TP}_{i,t}$. Accordingly, by setting the base year, normalizing the TP in the base year to 1, and sequentially calculating the right-hand side of Eq (7), it is possible to obtain time series $\{\mathrm{TP}_{i,t}\}_{t=0,1,2\ldots}$.

3.3. Relativization of TP

If we follow the steps described in Subsections 3.1 and 3.2, we can obtain the regional TP time series $\{\mathrm{TP}_{i,t}\}_{t=0,1,2\ldots}$ of an industry for each region *i*. However, in this study, we investigate how the infrastructure stock contributes to the growth of TP of each industry. Accordingly, we use the $\mathrm{TP}_{i,0}$ data for the base year per each region *i*, and consider the relative contribution of infrastructure stock to the growth of TP for each industry. Here, we use a method based on the translog index of Jorgenson and Nishimizu (1978) and Caves et al. (1982), and perform relativization of $\mathrm{TP}_{i,t}$. Specifically, we standardize the geometric mean TP_0^* of the base year $\mathrm{TP}_{i,0}$, and then calculate $\mathrm{TP}_{i,0}$ according to the following equation:

$$\ln\left(\frac{\mathrm{TP}_{i,0}}{\mathrm{TP}_0^*}\right) = \ln\left(\frac{Q_{i,0}}{Q_0^*}\right) - \bar{\pi}_{K_{i,0}}\ln\left(\frac{K_{i,0}}{K_0^*}\right) - \bar{\pi}_{L_{i,0}}\ln\left(\frac{L_{i,0}}{L_0^*}\right) - \bar{\pi}_{M_{i,0}}\ln\left(\frac{M(B_{i,0})}{M_0^*}\right); \tag{9}$$

provided that

$$\ln(\mathrm{TP}_0^*) = \frac{\sum_i \ ln\left(\mathrm{TP}_{i,0}\right)}{\#I}; \qquad \mathrm{TP}_0^* = \sqrt[\#I]{\Pi_i \mathrm{TP}_i};$$

$$\ln(Q_0^*) = \frac{\sum_i \ ln\left(Q_{i,0}\right)}{\#I}; \qquad Q_0^* = \sqrt[\#I]{\Pi_i Q_i};$$

$$\ln(K_0^*) = \frac{\sum_i \ ln\left(K_{i,0}\right)}{\#I}; \qquad K_0^* = \sqrt[\#I]{\Pi_i K_i};$$

$$\ln(L_0^*) = \frac{\sum_i \ ln\left(L_{i,0}\right)}{\#I}; \qquad L_0^* = \sqrt[\#I]{\Pi_i L_i};$$

$$\ln(M_0^*) = \frac{\sum_i \ ln\left(M\left(B_{i,0}\right)\right)}{\#I}; \qquad M_0^* = \sqrt[\#I]{\Pi_i M\left(B_{i,0}\right)};$$

$$\bar{\pi}_{K_{i,0}} = \frac{\pi_{K_{i,0}} + \pi_{K_0}^*}{2}; \qquad \pi_{K_0}^* = \frac{\sum_i \ \pi_{K_{i,0}}}{\#I};$$

$$\bar{\pi}_{L_{i,0}} = \frac{\pi_{L_{i,0}} + \pi_{L_0}^*}{2}; \qquad \pi_{L_0}^* = \frac{\sum_i \ \pi_{L_{i,0}}}{\#I};$$

$$\bar{\pi}_{M_{i,0}} = \frac{\pi_{M_{i,0}} + \pi_{M_0}^*}{2}; \qquad \pi_{M_0}^* = \frac{\sum_i \ \pi_{M_{i,0}}}{\#I};$$

where $\#I$ expresses the total number of regions i.

3.4. Panel data analysis

In the steps in Subsection 3.2, the $\mathrm{TP}_{i,0}$ of each region i, normalized by the national average in the base year, and the time series $\{\mathrm{TP}_{i,t}\}_{t=1,2\ldots}$ of TP normalized by the base year in each region i, were obtained. Here, the value obtained by $\mathrm{TP}_{i,0} \times \mathrm{TP}_{i,t}$ is replaced by a new value $\mathrm{TP}_{i,t}$.

Using the TP index calculated in this way, we estimate the parameter γ of infrastructure stock using a panel data analysis similar to Hulten et al. (2006). In Eq. (2), which assumes a multiplicative structure for the term expressing Hicks-neutral technical change, we now assume that the time series on the right-hand side $\mathrm{TP}_{i,t}$ and the infrastructure stock on the left-hand side $B_{i,t}$ are known. Taking the natural logarithm of both sides of Eq. (2) and adding the error term, we obtain

$$\ln\left(\mathrm{TP}_{i,t}\right) = \ln\left(A_{i,0}\right) + \lambda_i t + \gamma \ln\left(B_{i,t}\right) + \varepsilon_{i,t} \ ; \tag{10}$$

as the estimating equation. In the analysis, $\ln(A_{i,0})$ is a constant term expressing the fixed effect of each region, and $\lambda_i t$ denotes the linear time trend term.

Here, when $\ln\left(\mathrm{TP}_{i,t}\right)$ or $\ln\left(B_{i,t}\right)$ is a time series that does not satisfy stationarity, it needs to be analyzed by taking the difference and making it stationary, or using an econometric model. Here, it is important to note that the former approach looks at the short-term effects of infrastructure investment on changes in TFP. In other words, the former approach destroys the long-term equilibrium relationship (Munnell, 1992). In the latter approach, if $\ln\left(\mathrm{TP}_{i,t}\right)$ and $\ln\left(B_{i,t}\right)$ are cointegrated in I(1), fully modified OLS (Pedroni, 2001) and dynamic OLS (Kao and Chiang, 2001) can be used (e.g., Okubo, 2008). However, as will be verified later, in this study, there were a small number of sectors that suggested I(0), or stationarity, for $\ln\left(\mathrm{TP}_{i,t}\right)$. The panel ARDL model can be used even when I(0) and

I(1) variables are mixed, provided that the I(2) variable must not be present (Pesaran et al., 1999, 2001).

The panel ARDL model can be formulated as follows:

$$\ln(\mathrm{TP}_{i,t}) = \ln(A_{i,0}) + \sum_{j=1}^{p} \lambda_{ij} \ln(\mathrm{TP}_{i,t-j}) + \sum_{j=0}^{q} \delta_{ij} \ln(B_{i,t-j}) + \varepsilon_{i,t}. \quad (11)$$

If we express this as an error correction equation, then

$$\Delta\ln(\mathrm{TP}_{i,t}) = \ln(A_{i,0}) + \phi_i \left(\ln(\mathrm{TP}_{i,t-1}) - \theta_i \ln(B_{i,t})\right) + \sum_{j=1}^{p-1} \lambda_{ij}^{*} \Delta\ln(\mathrm{TP}_{i,t-j}) + \sum_{j=0}^{q-1} \delta_{ij}^{*} \Delta\ln(B_{i,t-j}) + \varepsilon_{i,t}; \quad (12)$$

is obtained, where $\Delta\ln(\mathrm{TP}_{i,t}) = \ln(\mathrm{TP}_{i,t}) - \ln(\mathrm{TP}_{i,t-1})$; $\phi_i = -(1 - \sum_{j=1}^{p} \lambda_{ij})$; $\theta_i = \sum_{j=0}^{q} \delta_{ij} / (1 - \sum_k \lambda_{ik})$ $\lambda_{ij}^{*} = -\sum_{m=j+1}^{p} \lambda_{im}$, $j = 1, 2, ..., p-1$, and $\delta_{ij}^{*} = -\sum_{m=j+1}^{q} \delta_{im}$, $j = 1, 2, ..., q-1$. Here, ϕ_i is the error-correcting speed of adjustment term. The parameter is expected to be significantly negative under the prior assumption that the variables show a return to a long-term equilibrium (Blackburne III and Frank, 2007). Our main interest is θ_i, which contains the long-term relationships between the variables. According to Murthy and Okunade (2016), the endogeneity problem does not arise in ARDL modelling when estimating both the short- and long-term coefficients simultaneously and with lagged dependent and explanatory variables. The ARDL coefficients estimates are super-consistent even for small samples.

Pesaran et al. (1999) have proposed a pooled mean group (PMG) estimator for Eq. (12). According to Blackburne III and Frank (2007), this estimator allows the intercept, short-term coefficients, and error variances to differ across the groups but constrains the long-term coefficients from being equal across groups $(\theta_i = \theta)$. They developed a maximum likelihood method to estimate the parameters[7]. See Pesaran et al. (1999) and Blackburne III and Frank (2007) for more details about the likelihood function that we maximize, and the full covariance matrix for inferences. Pesaran et al. (1999) showed that PMG is very robust to outliers and lag orders[8]. Martínez-Zarzoso and Bengochea-Morancho (2004) empirically showed in terms of an environmental Kuznets curve that a fixed effects estimator, which imposes homogeneity of slope while allowing only the intercepts to vary across individuals, may produce suspicious results.

3.5. R-JIP

This subsection outlines the Regional-Level Japan Industrial Productivity Database (R-JIP), the dataset used in this study. First, we provide an overview of the R-JIP and explain that it is an appropriate dataset for analyzing productivity by industry and region. Next, we describe the features of the latest version of R-JIP2017 used in this study. Finally, another dataset used in this study, R-JIP social infrastructure data, is described.

R-JIP is a database for policy analysis published by The Research Institute of Economy, Trade and Industry (RIETI), and is regarded as a basic resource for analyzing interregional productivity disparities and industrial structure in Japan[9]. Globally, for

[7] Stata's xtpmg command was used for the estimation.
[8] See Cho et al. (2022) for details about the ARDL model.
[9] https://www.rieti.go.jp/jp/database/r-jip.html (accessed on March 11, 2022)

example, to analyze the causes of various economic disparities among its member states and their changes over time the EU KLEMS project has compiled a database of trends in physical and human capital accumulation and TFP. In Japan, the Japan Industrial Productivity Database (JIP) has been published by RIETI as basic data for the analysis of Japan's economic growth and changes in its industrial structure. The R-JIP consists of annual data needed for estimating TFP for 47 prefectures and 23 industries, including capital and labor investments accounting for nominal and real added value, and differences in quality.

In this study, we use R-JIP2017, the latest of the available R-JIP databases. In R-JIP2017, the available annual data period has been extended from the period (1970-2009) in R-JIP2014 to (1970-2012). In addition, the data period extended by R-JIP2017 includes the Great East Japan Earthquake in 2011, with estimation of the damaged capital stock being reflected in the database.

3.6. R-JIP2017

In R-JIP2017, annual data is given for estimating TFP for 47 prefectures and 23 industries (1. Agriculture, forestry, and fishing; 2. Mining; 3. Food products and beverages; 4. Textiles; 5. Pulp, paper, and paper products; 6. Chemicals; 7. Petroleum and coal products; 8. Non-metallic mineral products; 9. Basic metal; 10. Fabricated metal products; 11. Machinery; 12. Electrical machinery; 13. Transport equipment; 14. Precision instruments; 15. Other manufacturing; 16. Construction; 17. Electricity, gas, and water supply; 18. Wholesale and retail; 19. Finance and insurance; 20. Real estate; 21. Transport and communications; 22. Private non-profit services; 23. Government services), In this study, we use data on real value added (price in 2000), nominal value added, real capital stock (price in 2000), nominal cost of capital, quality index (capital, common nationwide), man-hour (workers × total annual working hours per worker/1000), nominal labor cost, and quality index (labor). Notably, unlike the JIP database, the R-JIP database uses the output based on gross value added, as it does not have information on intermediate inputs due to the limitation of available data. However, By using the output based on gross value added, real value added items receive a negative value. Table 1 lists these negative values. The industries whose real value added is negative are 5. Pulp, paper, and paper products; 7. Petroleum and coal products; 9. Basic metals; and 14. Precision instruments. These four industries were excluded, leaving us with 19 industries for analysis.

Table 1: List of industries with negative real value added, around here

In the R-JIP database, labor quality indices are given for labor input, which enables us to analyze labor input considering differences in quality. For the calculation of the labor quality index for each prefecture, we consider the factors of education, age, gender, and industry. We refer to Tokui et al. (2013; 2019) for the details of the calculation method. The R-JIP database also provides a quality index for capital stock, which allows us to analyze capital inputs with different quality levels. As for the quality of capital, we obtain the real capital stock series by industry and capital service input from the JIP database. Additionally, we obtain the capital quality index from the ratio of the two databases. Again, we refer to Tokui et al. (2013; 2019) for the details of the calculation method of the capital quality index.

The R-JIP database also contains data on social infrastructure. According to the summary of the R-JIP database on its website, social infrastructure in the R-JIP database is based on the estimation of social infrastructure stock by the Cabinet Office[10]. In the R-JIP database, similar to the nationwide JIP database, the capital stock that can be judged to be used in the production activities of each sector is calculated as the capital service input of each sector regardless of whether the investment entity belongs to the private or public sector. For example, in the agriculture, forestry, and fisheries sectors, many public capital improvements such as agricultural roads and irrigation channels are carried out. The same applies to water supply facilities in the electricity, gas, and water industries, and toll roads in the transportation industry. In addition, the service industry (public) is included in the sectoral classification, which includes school facilities, cultural facilities, airports, and harbors. Furthermore, there are cases where regional productivity differences are considered, with the regional development of social infrastructure being the focus. In such cases, "social infrastructure" data defined by investment entities have often been used. However, when using "social infrastructure" data defined by investment entities, combined with the R-JIP data defining capital categories according to use, some social infrastructure data could double counted. Therefore, public capital inputs that cannot be associated with the economic activities of individual sectors, except for those already counted as capital service inputs of each sector in the R-JIP database, are referred to as "social infrastructure consistent with the R-JIP database" and are provided as ancillary data. The "social infrastructure consistent with the R-JIP database" includes roads other than toll roads, urban parks, flood control, mountain control, and coastal maintenance. In the "social infrastructure consistent with the R-JIP database", "toll roads" are classified as an input of the transportation and communication sector, and are classified separately from "roads other than toll roads", which are defined as social infrastructure. Therefore, the sum of "toll roads" and "roads other than toll roads" is used as road stock data for the empirical research in this study.

4. *Empirical analysis*

4.1. Construction of panel datasets

In this section, we conduct an empirical analysis based on the Hulten et al.'s (2006) method, described in section 3, using the R-JIP2017 database described in Subsection 3.6. However, as described below, some modifications are made to the method explained in Section 3 to use real value added in this study.

4.1.1. TFP estimation using real value added

Here, we explain the assumptions about the production function that are additionally required for TFP estimation using R-JIP2017.

In Section 3, we outlined the method of estimating productivity using growth accounting based on the assumption that the dataset includes both gross output and intermediate inputs. However, owing to the limitation of available data, the R-JIP database does not include information on intermediate inputs, and instead uses output data based on gross value added. Hulten et al. (2006) pointed out that although real value added data are generally used because they are easier to obtain than gross output data, using real value added data requires a weak separability assumption in the production function (Goldman

[10] https://www5.cao.go.jp/keizai2/ioj/index.html (accessed on March 11, 2022)

and Uzawa, 1964) and that using gross output data is preferable. In reality, however, studies using real value added have been widely conducted due to the availability of data, and Hulten et al. (2006) conducted an analysis using real value added to enable comparison with those studies. Hulten and Schwab (1984) and Hulten et al. (2006) assumed that the production function is weakly separable into value added and intermediate inputs, and that Hicks-neutral technical change is included in the value added function. In this sense, TP in the Section 3 is synonymous with TFP in this study. The deflator common to all countries was used for the real value added because of the availability of data.

4.1.2. Estimating TFP by considering labor and capital quality

We explain TFP estimation considering quality index using R-JIP2017. The rate of increase in TFP in industry s and prefecture i at time t, $\Delta\ln(\mathrm{TFP}_{i,s,t})$, can be obtained from the following equation:

$$\Delta\ln(\mathrm{TFP}_{i,s,t}) = \Delta\ln(V_{i,s,t}) - \frac{1}{2}\left(S^K_{i,s,t} + S^K_{i,s,t-1}\right)\Delta\ln(K_{i,s,t}) - \frac{1}{2}\left(S^L_{i,s,t} + S^L_{i,s,t-1}\right)\Delta\ln(L_{i,s,t}); \qquad (13)$$

where i (i=1, ..., 47) is an index indicating the prefecture, s (s=1,19 ...) is an index indicating the industry, t (t=1972, ..., 2012) is an index indicating time, $V_{i,s,t}$ is the real value added, $S^K_{i,s,t}$ is the capital cost share, $S^L_{i,s,t}$ is the labor cost share, $K_{i,s,t}$ represent capital inputs, and $L_{i,s,t}$ represent labor input. In addition, if $Q^K_{i,s,t}$ is the quality index of capital, $Q^L_{i,s,t}$ is the labor quality index, $Z_{i,s,t}$ is real capital stock, and $H_{i,s,t}$ is man-hours, then $K_{i,s,t} = Q^K_{i,s,t}Z_{i,s,t}$; $L_{i,s,t} = Q^L_{i,s,t}H_{i,s,t}$ holds. Therefore, Eq. (11) can be rewritten as follows:

$$\begin{aligned}\Delta\ln(\mathrm{TFP}_{i,s,t}) = &\\ &\Delta\ln(V_{i,s,t}) - \frac{1}{2}\left(S^K_{i,s,t} + S^K_{i,s,t-1}\right)\left(\Delta\ln(Z_{i,s,t}) + \Delta\ln(Q^K_{s,t})\right) \\ &- \frac{1}{2}\left(S^L_{i,s,t} + S^L_{i,s,t-1}\right)\left(\Delta\ln(H_{i,s,t}) + \Delta\ln(Q^L_{i,s,t})\right).\end{aligned} \qquad (14)$$

This is provided that the quality of capital $Q^K_{i,s,t}$ takes the same value for all prefectures, i, within the same industry, s. Additionally, $Q^K_{s,t}$ values are available for the right-hand side of Eq. (1). Since statistical values are available for the right-hand side of Eq. (14), it is possible to obtain a time series for TFP standardized to 1 $\mathrm{TFP}_{i,s,1972}$ per each industry s and prefecture i.

4.1.3. Relativization of TFP

We relativize TFP using $\mathrm{TFP}_{i,s,1972}$ for industry s in each prefecture i. First, we denote the national geometric mean of each variable with 1972 as the base year as $\ln(\bar{V}_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}\ln(V_{i,s,1972})$ and $\ln(\bar{K}_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}\ln(K_{i,s,1972})$ and $\ln(\bar{L}_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}\ln(L_{i,s,1972})$. In addition, we denote the national arithmetic mean of the cost shares of capital and labor as $(\bar{S}^K_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}S^K_{i,s,1972}$ and $(\bar{S}^L_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}S^L_{i,s,1972}$, respectively. Here, we standardize the geometric mean of TFP for each prefecture i and industry s in base year $t = 1972$ to 1. Thus, $\mathrm{TFP}_{i,s,1972}$ for each prefecture i can be obtained from the following equation:

$$\ln\left(\frac{TFP_{i,s,1972}}{\overline{TFP}_{s,1972}}\right) = \ln\left(\frac{V_{i,s,1972}}{\bar{V}_{s,1972}}\right) - \frac{1}{2}\left(S^K_{i,s,1972} + \bar{S}^K_{s,1972}\right)\ln\left(\frac{K_{i,s,1972}}{\bar{K}_{s,1972}}\right) - \frac{1}{2}\left(S^L_{i,s,1972} + \bar{S}^L_{s,1972}\right)\ln\left(\frac{L_{i,s,1972}}{\bar{L}_{s,1972}}\right). \tag{15}$$

Here, when we denote the national mean of real capital stock, capital quality, man-hours, and labor quality as$\ln(\bar{Z}_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}\ln(Z_{i,s,1972})$, $\ln(\bar{Q}^K_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}\ln(Q^K_{i,s,1972})$, $\ln(\bar{H}_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}\ln(H_{i,s,1972})$, and $\ln(\bar{Q}^L_{s,1972}) = \frac{1}{47}\sum_{i=1}^{47}\ln(Q^L_{i,s,1972})$. Hence, $K_{i,s,1972} = Q^K_{i,s,1972}Z_{i,s,1972}$; $L_{i,s,1972} = Q^L_{i,s,1972}H_{i,s,1972}$ holds, so we substitute these into the Eq. (15), and we obtain

$$\ln\left(\frac{TFP_{i,s,1972}}{\overline{TFP}_{s,1972}}\right) = \ln\left(\frac{V_{i,s,1972}}{\bar{V}_{s,1972}}\right) - \frac{1}{2}\left(S^K_{i,s,1972} + \bar{S}^K_{s,1972}\right)\left(\ln\left(\frac{Z_{i,s,1972}}{\bar{Z}_{s,1972}}\right) + \ln\left(\frac{Q^K_{i,s,1972}}{\bar{Q}^K_{s,1972}}\right)\right) - \frac{1}{2}\left(S^L_{i,s,1972} + \bar{S}^L_{s,1972}\right)\left(\ln\left(\frac{Z_{i,s,1972}}{\bar{Z}_{s,1972}}\right) + \ln\left(\frac{Q^K_{i,s,1972}}{\bar{Q}^K_{s,1972}}\right)\right). \tag{16}$$

Since $Q^K_{i,s,1972} = Q^K_{s,1972} \quad \forall i$, we obtain

$$\ln\left(\frac{TFP_{i,s,1972}}{\overline{TFP}_{s,1972}}\right) = \ln\left(\frac{V_{i,s,1972}}{\bar{V}_{s,1972}}\right) - \frac{1}{2}\left(S^K_{i,s,1972} + \bar{S}^K_{s,1972}\right)\ln\left(\frac{Z_{i,s,1972}}{\bar{Z}_{s,1972}}\right) - \frac{1}{2}\left(S^L_{i,s,1972} + \bar{S}^L_{s,1972}\right)\left(\ln\left(\frac{H_{i,s,1972}}{\bar{H}_{s,1972}}\right) + \ln\left(\frac{Q^K_{i,s,1972}}{\bar{Q}^K_{s,1972}}\right)\right). \tag{17}$$

From this equation, it is possible to stipulate the TFP level ($\mathrm{TFP}_{i,s,1972}$) for base year 1972 for each industry s and prefecture i. Thus, it is possible to calculate $\mathrm{TFP}_{i,s,t}$ at any other year: $t = 1973, \ldots, 2012$ in industry s and prefecture i, using the sequential equation of $\Delta\ln(\mathrm{TFP}_{i,s,t}) = \ln(\mathrm{TFP}_{i,s,t}) - \ln(\mathrm{TFP}_{i,s,t-1})$.

Fig. 1 shows the change in relativized TFP (TFP index) for each industry and prefecture (grouped based on region: Hokkaido/Tohoku; Kanto; Chubu; Kinki; Chugoku; Shikoku; Kyushu). Some industries, such as the service industry, showed an upward trend during the analysis period, while others, such as the real estate industry, deteriorated considerably. In addition, we found that the transition pattern differed among prefectures; however, a certain degree of similarity was observed. Since Okinawa has historically exhibited a TFP level that differs considerably from that of other prefectures, in the following analysis, used data for the years 1972 to 2012, excluding Okinawa.

Fig. 1 Changes in TFP index, around here

4.2. Panel data analysis

4.2.1. Overview of Panel data analysis

We conducted a panel data analysis to estimate the effect of infrastructure on TFP. The response variables used in the analysis were the time series $\mathrm{TFP}_{i,s,t}$ for the TFP of each industry and prefecture obtained using Eqs. (17) and (14), and the explanatory variables are the road stock explained in subsection 4.1. Here, we computed Eq. (10) using Hulten et al.'s

(2006) method. Hulten et al. (2006), in the Indian context, indicated relatively monotonic growth when looking at TFP over the long term; thus, it seems reasonable to assume linear exogenous growth. However, as can be seen from Fig. 1, it is difficult to consider the existence of linear exogenous growth with time in the Japanese context. The fact that the change in road stock (the sum of "toll roads" and "roads other than toll roads") in each prefecture shows a relatively monotonic increasing trend, as shown in Fig. 2, suggests that it is highly likely that the term exogenously changing TFP with time is changing in the same way as TFP.

Fig. 2: "Toll roads" + "non-toll roads", around here

Therefore, in this study, we adopted the following two-ways panel data model in consideration of the features of data in Japan[11]:

$$\begin{aligned}\ln\left(\mathrm{TFP}_{i,s,t}\right) &= \alpha_s + \gamma_s \ln\left(B_{i,t}\right) + u_{i,s,t},\\ u_{i,s,t} &= \mu_{i,s} + \lambda_{s,t} + \nu_{i,s,t},\end{aligned} \tag{18}$$

where γ_s is a parameter related to the infrastructure stock $B_{i,t}$, and expresses the relationship between infrastructure stock and TFP in industry s. In addition, $\mu_{i,s}$ expresses the specific effect of prefecture i in industry s, $\lambda_{s,t}$ expresses the unique effect of time t in industry s, and $\nu_{i,s,t}$ is the error term. Since the random effects model requires a rather strong assumption that there is no correlation between $\mu_{i,s}$ and the explanatory variables, in this study, we chose to use the 2FE model. As described in Subsection 3.4, when TFP or infrastructure stock series do not satisfy stationarity, the 2FE model suffers from the problem of spurious correlation, which may lead to erroneous policy implications. In this study, we conducted a unit root test on panel data. Then, we employed the panel ARDL model shown in Eq. (11), in addition to the 2FE model.

4.2.2 Results of the panel unit root test

Although there are several methods for panel unit root tests (Baltagi, 2005), in this study, we chose to use the most commonly used methods of Im et al. (2003) (IPS) and Levin et al. (2002) (LLC), as Shafique et al. (2021)[12].

Im et al. (2003) proposed the following panel model:

$$\Delta y_{it} = \rho_i y_{i,t-1} + \boldsymbol{z}'_{it}\boldsymbol{\gamma}_i + \varepsilon_{i,t}, \tag{19}$$

where $y_{i,t}$ is a series of TFPs or road stock. If we consider the constant term and trend, then $\boldsymbol{z}'_{it} = (1, t)$, so that $\boldsymbol{z}'_{it}\boldsymbol{\gamma}_i$ represents panel-specific means and linear time trends. Im et al. (2003) assumed that $\varepsilon_{i,t}$ is independently and normally distributed for all values of i and t, allowing $\varepsilon_{i,t}$ to have heterogeneous variances σ_i^2 across panels. Here, the null hypothesis H_0: $\rho_i = 0$ for all values of i and the alternative hypothesis H_a: $\rho_i < 0$. We report the IPS's

[11] Another option may include the use of interactive fixed effects model (Bai, 2009).

[12] Note that in this study, we used Stata's xtunitroot command to conduct the unit root tests.

$W_{t\text{-bar}}$ statistic (Im et al., 2003, Eq.(4.10)), which has an asymptotically standard normal distribution as $T \to \infty$ followed by $I \to \infty$.

Levin et al. (2002) assumed $\rho_i = \rho$ for all i, and put forward the following model:

$$\Delta y_{it} = \rho y_{i,t-1} + \boldsymbol{z'}_{it}\boldsymbol{\gamma}_i + \sum_{j=1}^{p} \theta_{ij}\, \Delta\, y_{i,t-j} + u_{i,t}; \qquad (20)$$

Since there is a possibility that serial correlation exists in $\varepsilon_{i,t}$, we introduce the $\sum_{j=1}^{p} \theta_{ij}\, \Delta\, y_{i,t-j}$ term, and $u_{i,t}$ is assumed to be white noise. We formulate the null hypothesis as H_0: $\rho = 0$ and the alternative hypothesis as H_a: $\rho < 0$. In other words, the null hypothesis is that the panel dataset contain unit roots. Levin et al. (2002) proposed adjusted *t*-statistic (Adjusted t*) that has a standard normal distribution when $T \to \infty$ followed by $I \to \infty$, which we use in this study. The lag p should be determined by the information criterion. In this study, we use the Akaike information criterion (AIC).

The results of the panel unit root tests are shown in Tables 2 and 3. Table 2 shows the test results for the road stock variable, while Table 3 shows the test results for TFP. Table 3 separately shows the results for the sectors in which the null hypothesis is rejected at the 5% significance level for both IPS and LLC tests, and the results for sectors in which the null hypothesis is not rejected in both tests. For the latter, we also present the results for the first-difference equation as well as for the level equation. In the tests, we assume that $\boldsymbol{z'}_{it} = (1, t)$ (introducing the constant term and trend). In addition, the average lag for p, estimated based on AIC, is included in these tables.

First, Table 2 shows that the null hypothesis of road capital stock was not rejected at the 5% significance level for the level equation in both IPS and LLC tests. However, the null hypothesis was rejected for the first-difference equation, from which we can conclude that road capital stock is in the I(1) series.

Next, we examine the results of the tests for TFP. Table 3 lists the implied I(0) variables for the sectors of Agriculture, forestry, and fishing; Mining; Food products and beverages; Machinery; Electrical machinery; Transport equipment; Other manufacturing; Construction; Transport, and communications; and the implied I(1) variables for Textiles; Chemicals; Non-metallic mineral products; Fabricated metal products; Electricity, gas, and water supply; Wholesale and retail; Finance and insurance; Real estate; Private non-profit services; and Government services. Thus, owing to the mixture of I(1) and I(0) variables, we decided to adopt the panel ARDL in this study.

Table 2: Results of panel unit root tests (road infrastructure), around here

Table 3: Results of panel unit root tests (TFP), around here

4.2.3. Estimation results

As discussed in subsection 4.2.2, the TFP of infrastructure stock and some sectors was suggested to be I(1). Such cases may be easily made stationary by taking the difference of each variable. However, the cost of this approach is a loss of information regarding long-term relationships. Table 4 below shows the estimation results for the 2FE, first-difference

(FD), and panel ARDL model (PMG estimation)[13]. For panel ARDL model, we assumed ARDL (1,1), with setting $p = q = 1$.

From Table 4, using the 2FE model, we find that the long-term effect of road infrastructure on TFP is negative in six sectors (i.e., Non-metallic mineral products; Other manufacturing; Construction; Finance and insurance; Real estate; and Private non-profit services) while being positively significant in six other sectors (i.e., Agriculture, forestry, and fishing; Food products and beverages; Fabricated metal products; Machinery; Electrical machinery; and Government services) at the 5% significance level. However, in the case of FD, the effect is not statistically significant for private non-profit services, construction, and other manufacturing; and is negative for real estate, transport, and communications, and government services, but it is positively significant in the remaining 13 sectors at the 5% significance level. Although the 2FE and FD models both consider time-invariant unit specific effects, there are relatively large differences in the estimated values at the sign level. The large difference in the coefficient estimates between the 2FE and FD models is possibly due to the non-consideration of non-stationarity (or serial correlation). In fact, when we look at the correlation coefficient between estimates for 18 industries[14], the panel ARDL and FD models, which accounted for stationarity, showed a relatively similar trend of 0.61 (between panel ARDL and 2FE: –0.23, between FD and 2FE: 0.26). This indicates the risk of using the 2FE model blindly in a series with non-stationarity. Nevertheless, it is important to note that FD only looks at short-term effects (Munnell, 1992).

The results of the PMG estimation of the panel ARDL model verify the existence of a significantly positive impact at the 5% level in 11 sectors (i.e., Mining; Textiles; Chemicals; Non-metallic mineral products; Fabricated metal products; Machinery; Transport equipment; Other manufacturing; Electricity, gas and water supply; Wholesale and retail; and Finance and insurance), and a significantly negative effect in four sectors (i.e., Electrical machinery; Transport and communications; Private non-profit services; and Government services). In particular, in industries for which the existence of unit roots is suggested, as shown in Table 3, the coefficient of the effect of infrastructure is negative in 2FE but positive in PMG in many cases (e.g., Textiles; Chemicals; Non-metallic mineral products; Electricity, Electricity, gas and water supply; and Finance and insurance)[15].

These results indicate that road infrastructure had a positive impact on TFP in many industries during the analysis period, from 1972 to 2012. Japan's rapid economic growth is generally said to have lasted from 1954 to 1973, and the analysis period of this study corresponds to the period after that. The fact is that, under the principle of “balanced development of national land”, since the early 1970s, Japan has been making administrative investments with preferential treatment to rural areas over the three major metropolitan areas, and positive effects have been achieved even during this period.

Table4: Results of the panel analysis, around here

[13] Shows estimation results only for the coefficients of long-term effects. Although the report of results is omitted, the error-correcting speed of adjustment term ϕ_i was negative for all sectors and significant at 0.1%

[14] This is calculated for 18 industries other than real estate, for which the maximum likelihood estimates did not converge in the panel ARDL model

[15] Such quite different results between static fixed effects model and PMG are also found in the findings of Martínez-Zarzoso and Bengochea-Morancho (2004).

5. Conclusions

In this study, we analyzed the relationship between road infrastructure stock and TFP using the "R-JIP2017" database of productivity by industry and prefecture in Japan. In this study, we estimated TFP based on the growth accounting frame introduced by Hulten et al. (2006), and thereby derived estimates without specifying the functional form of a production function in a certain degree. We performed long-term analysis by industry for each prefecture in Japan for the period of 1972 to 2012, which roughly corresponds to the period following Japan's rapid economic growth. In addition, the possibility of the existence of unit roots was explicitly considered in the panel data for road infrastructure stock and TFP to eliminate spurious correlations. The results of the panel unit root tests indicated that the road infrastructure variable was I(1), and that of TFP was either I(1) or I(0), depending on the sector, and therefore a panel ARDL model was used, in which a mixture of I(0) and I(1) variables was allowed, and the PMG method was used for estimation (Pesaran et al.,1999).

The estimation results of the panel ARDL model showed that TFP is positively and significantly related to road stock at the 5% level for the following industries: Mining, Textiles and Chemicals; Non-metallic mineral products; Fabricated metal products; Machinery; Transport equipment; Other manufacturing; Electricity, gas and water supply; Wholesale and retail; and Finance and insurance. In addition, road stock has a positive and significant relationship with TFP at the 5% level for the following industries: Electrical machinery; Transport and communications; Private non-profit services; and a significantly negative relationship for government services. The results of the panel ARDL model differed greatly from those of the 2FE model, which does not consider non-stationarity, with the latter tending to underestimate (negatively) the effect of road infrastructure as a whole. The blind use of the 2FE model is customary in current empirical research; however, caution should be taken with respect to causal statistical inference (Imai et al., 2021), as our study empirically demonstrated that non-stationarity in macro panels might lead to erroneous policy implications.

The results of the analysis showed that road infrastructure had a positive impact on TFP in many industries during the analysis period of 1973 to 2012. Japan's rapid economic growth is generally said to have occurred from 1954 to 1973, and the analysis period in this study corresponds to the period after that. The fact is that, under the principle of "balanced development of national land", since the early 1970s, Japan has been making administrative investments with preferential treatment to rural areas over the three major metropolitan areas, and positive effects have been achieved even during this period, which may provide useful reference for national land planning in developed and developing countries.

Future studies should validate the results of the current study by different methods (e.g., vector autoregressive model, Annala et al., 2008) using the same dataset, R-JIP2017, and then compare and discuss the results. Additionally, in the medium to long term, it is necessary to conduct empirical analysis including data before the period of high economic growth in a manner consistent with the R-JIP database. Moreover, it is also worthwhile to examine the results in spatial units as argued by Munnell (1992) in the context of Japan. Finally, as the theoretical models for structural estimation of endogenous economic growth have not made much progress, future research should explore the use of methods from quantitative spatial economics, which links theory and evidence, to advance our knowledge.

References

- Alam, K.M., Li, X., Baig, S., Ghanem, O., and Hanif, S. (2020) Causality between transportation infrastructure and economic development in Pakistan: An ARDL analysis, *Research in Transportation Economics*, 100974.
- Annala, C.N., Batina, R.G. and Feehan, J.P. (2008) Empirical impact of public infrastructure on the Japanese economy, *The Japanese Economic Review*, 59 (4), 419–437.
- Anwar, A., Ahmad, N. and Madni, G.R. (2020) Industrialization, freight transport and environmental quality: Evidence from belt and road initiative economies, *Environmental Science and Pollution Research*, 27 (7), 7053–7070.
- Aschauer, D.A. (1989) Is public expenditure productive? *Journal of Monetary Economics*, 23 (2), 177–200.
- Bai, J. (2009) Panel data models with interactive fixed effects. *Econometrica*, 77 (4), 1229–1279.
- Baird, B.A. (2005) Public infrastructure and economic productivity: A transportation-focused review, *Transportation Research Record*, 1932 (1), 54–60.
- Baltagi, B.H. (2005) *Econometric Analysis of Panel Data*, *3rd Edition*, John Wiley and Sons, Chichester.
- Barro, R.J. (1990) Government spending in a simple model of endogenous growth, *Journal of Political Economy*, 98 (5, Part2), S103–S125.
- Barzin, S., D'Costa, S. and Graham, D.J. (2018) A pseudo-panel approach to estimating dynamic effects of road infrastructure on firm performance in a developing country context, *Regional Science and Urban Economics*, 70, 20–34.
- Blackburne III, E.F. and Frank, M.W. (2007) Estimation of nonstationary heterogeneous panels, *The Stata Journal*, 7 (2), 197–208.
- Calderón, C., Moral-Benito, E. and Servén, L. (2015) Is infrastructure capital productive? A dynamic heterogeneous approach, *Journal of Applied Econometrics*, 30 (2), 177–198.
- Cass, D. (1965) Optimum growth in an aggregative model of capital accumulation, *Review of Economic Studies*, 32 (3), 233–240.
- Caves, D.W., Christensen, L.R. and Diewert, W.E. (1982) The economic theory of index numbers and the measurement of input, output, and productivity, *Econometrica*, 50 (6), 1393–1414.
- Cho, J.S., Greenwood-Nimmo, M. and Shin, Y. (2022) Recent developments of the autoregressive distributed lag modelling framework, *Journal of Economic Surveys*, in print.
- Deng, T. (2013) Impacts of transport infrastructure on productivity and economic growth: Recent advances and research challenges, *Transport Reviews*, 33 (6), 686–699.
- Ejiri, R., Okumura, M. and Kobayashi, K. (2001) The productivity of infrastructure and economic growth: A state-of-the-art review, *Journal of Japan Society of Civil Engineers*, 688 (IV-53), 75–87 (in Japanese).
- Elburz, Z., Nijkamp, P. and Pels, E. (2017) Public infrastructure and regional growth: Lessons from meta-analysis, *Journal of Transport Geography*, 58, 1–8.
- Estache, A. and Fay, M. (2007) Current debates on infrastructure policy, The World Bank, Paper prepared for the Growth Commission.
- Futagami, K., Morita, Y., and Shibata, A. (1993) Dynamic analysis of an endogenous growth model with public capital, *The Scandinavian Journal of Economics*, 95 (4), 607–625.

- Gibbons, S. and Overman, H.G. (2009) Productivity in transport evaluation studies, Department for Transport, UK.
- Goldman, S.M. and Uzawa, H. (1964) A Note on separability in demand analysis, *Econometrica*, 32 (3), 387–398.
- Holmgren, J. and Merkel, A. (2017). Much ado about nothing?–A meta-analysis of the relationship between infrastructure and economic growth, *Research in Transportation Economics*, 63, 13–26.
- Hulten, C.R., Bennathan, E. and Srinivasan, S. (2006) Infrastructure, externalities, and economic development: A study of the Indian manufacturing industry, *World Bank Economic Review*, 20 (2), 291–308.
- Hulten, C.R. and Schwab, R.M. (1984) Regional productivity growth in U.S. manufacturing: 1951–78, *American Economic Review*, 74 (1), 152–162.
- Im, K. S., Pesaran, M. H. and Shin, Y. (2003) Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, 115 (1), 53–74.
- Imai, K. and Kim, I.S. (2021) On the use of two-way fixed effects regression models for causal inference with panel data, *Political Analysis*, 29 (3), 405–415.
- Jorgenson, D.W. and Nishimizu, M. (1978) U.S. and Japanese economic growth, 1952–1974: An international comparison, *The Economic Journal*, 88 (352), 707–726.
- Kao, C. and Chiang, M.-H. (2001) On the estimation and inference of a cointegrated regression in panel data, In Baltagi, B.H., Fomby, T.B., Hill, R.C. (Eds). *Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics*, Emerald Group Publishing Limited, 15, 179–222.
- Kawaguchi, D., Ohtake, F. and Tamada, K. (2009) The productivity of public capital: Evidence from Japan’s 1994 electoral reform, *Journal of the Japanese and International Economies*, 23 (3), 332–343.
- Kawakami, T. and Doi, M. (2004) Port capital formation and economic development in Japan: A vector autoregression approach, *Papers in Regional Science*, 83 (4), 723–732.
- Khanna, R. and Sharma, C. (2021) Does infrastructure stimulate total factor productivity? A dynamic heterogeneous panel analysis for Indian manufacturing industries, *The Quarterly Review of Economics and Finance*, 79, 59–73.
- Koopmans, T.C. (1965) On the concept of optimal economic growth, Cowles Foundation Discussion Papers 1, Cowles Foundation for Research in Economics, Yale University.
- Kydland, F.E. and Prescott, E.C. (1982) Time to build and aggregate fluctuations, *Econometrica*, 50 (6), 1345–1370.
- Levin, A., Lin, C. F. and Chu, C.S. J. (2002) Unit root tests in panel data: Asymptotic and finite-sample properties, *Journal of Econometrics*, 108 (1), 1–24.
- Lucas, R.E. (1988) On the mechanics of economic development, *Journal of Monetary Economics*, 22 (1), 3–42.
- Magazzino, C. and Maltese, I. (2021) Transport infrastructures: Investments, evaluation and regional economic growth, *Research in Transportation Economics*, 88, 101125.
- Martínez-Zarzoso, I. and Bengochea-Morancho, A. (2004) Pooled mean group estimation of an environmental Kuznets curve for CO_2, *Economics Letters*, 82 (1), 121–126.
- Melo, P.C., Graham, D.J. and Brage-Ardao, R. (2013) The productivity of transport infrastructure investment: A meta-analysis of empirical evidence, *Regional Science and Urban Economics*, 43 (5), 695–706.

- Mera, K. (1973) II. Regional production functions and social overhead capital: An analysis of the Japanese case, *Regional and Urban Economics*, 3 (2), 157–185.
- Miyagawa, T., Kawasaki, K. and Edamura, K. (2013) Re-examination of the productivity effect of public capital, *RIETI Discussion Paper Series*, 13-J-081 (in Japanese)
- Miyara, I. and Fukushige, M. (2008) Types of public capital and their productivity in Japanese prefectures, *The Japanese Economic Review*, 59 (2), 194–210.
- Munnell, A.H. (1992) Policy watch: Infrastructure investment and economic growth, *Journal of Economic Perspectives*, 6 (4), 189–198.
- Murthy, V.N. and Okunade, A.A. (2016) Determinants of US health expenditure: Evidence from autoregressive distributed lag (ARDL) approach to cointegration, *Economic Modelling*, 59, 67–73.
- Na, K.Y., Han, C. and Yoon, C.H. (2013) Network effect of transportation infrastructure: a dynamic panel evidence, *The Annals of Regional Science*, 50 (1), 265–274.
- Okubo, M. (2008) Public capital and productivity: A nonstationary panel analysis, *Applied Economics Letters*, 15 (2), 95–99.
- Pedroni, P. (2001) Fully modified OLS for heterogeneous cointegrated panels, In Baltagi, B.H., Fomby, T.B., Hill, R.C. (Eds). *Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics*, Emerald Group Publishing Limited, 15, 93–130.
- Pesaran, M. H., Shin, Y. and Smith, R.P. (1999) Pooled mean group estimation of dynamic heterogeneous panels, *Journal of the American Statistical Association*, 94 (446), 621–634.
- Pesaran, M. H., Shin, Y. and Smith, R.J. (2001) Bounds testing approaches to the analysis of level relationships, *Journal of Applied Econometrics*, 16 (3), 289–326.
- Petrin, A. and Levinsohn, J. (2012) Measuring aggregate productivity growth using plant-level data, *The RAND Journal of Economics*, 43 (4), 705–725.
- Ramsey, F.P. (1928) A mathematical theory of saving, *The Economic Journal*, 38 (152), 543–559.
- Redding, S.J. and Rossi-Hansberg, E. (2017) Quantitative spatial economics, *Annual Review of Economics*, 9, 21–58.
- Romer, P.M. (1986) Increasing returns and long-run growth, *Journal of Political Economy*, 94 (5), 1002–1037.
- Romer, P.M. (1990) Endogenous technological change, *Journal of Political Economy*, 98 (5), S71–S102.
- Solow, R.M. (1956) A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70 (1), 65–94.
- Solow, R.M. (1957) Technical change and the aggregate production function, *The Review of Economics and Statistics*, 39 (3), 312–320.
- Straub, S. (2008) Infrastructure and development: A critical appraisal of the macro level literature, The World Bank.
- Swan, T.W. (1956) Economic growth and capital accumulation, *Economic Record*, 32 (2), 334–361.
- Tokui, J., Makino, T., Fukao, K., Miyagawa, T., Arai, N., Arai, S., Inui, T., Kawasaki, K., Kodama, N. and Noguchi, N. (2013) Compilation of the Regional-Level Japan Industrial Productivity Database (R-JIP) and analyses of productivity differences across prefectures, *Economic Review*, 64 (3), 218–239.
- Tokui, J. and Mizuta, T. (2019) Japan's prefectural-level KLEMS: Productivity comparison and service price differences, *International Productivity Monitor*, 36, 142–160.

- Tsukai, M. and Kobayashi, K. (2009) Infrastructure productivity with a long persistent effect, In Karlsson, C., Andersson, A.E., Cheshire, P.C., and Stough, R.R. (Eds). *New Directions in Regional Economic Development*, Springer, Berlin, Heidelberg,197–219.
- Välilä, T. (2020) Infrastructure and growth: A survey of macro-econometric research, *Structural Change and Economic Dynamics*, 53, 39–49.
- Van Beveren, I. (2012) Total factor productivity estimation: A practical review, *Journal of Economic Surveys*, 26 (1), 98–128.
- Venables, A., Laird, J.J. and Overman, H.G. (2014) Transport investment and economic performance: Implications for project appraisal, Department for Transport, UK.

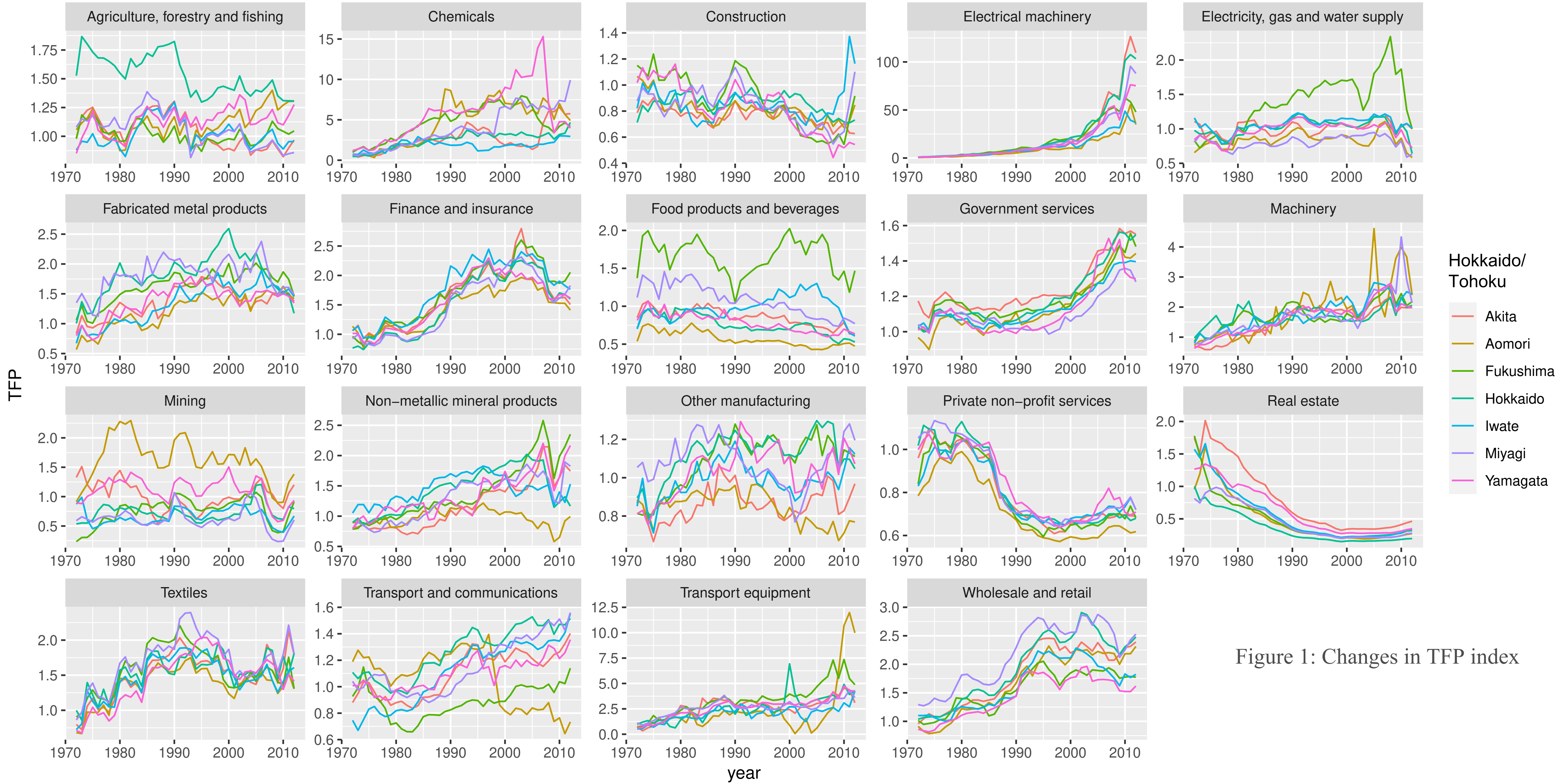

Figure 1: Changes in TFP index

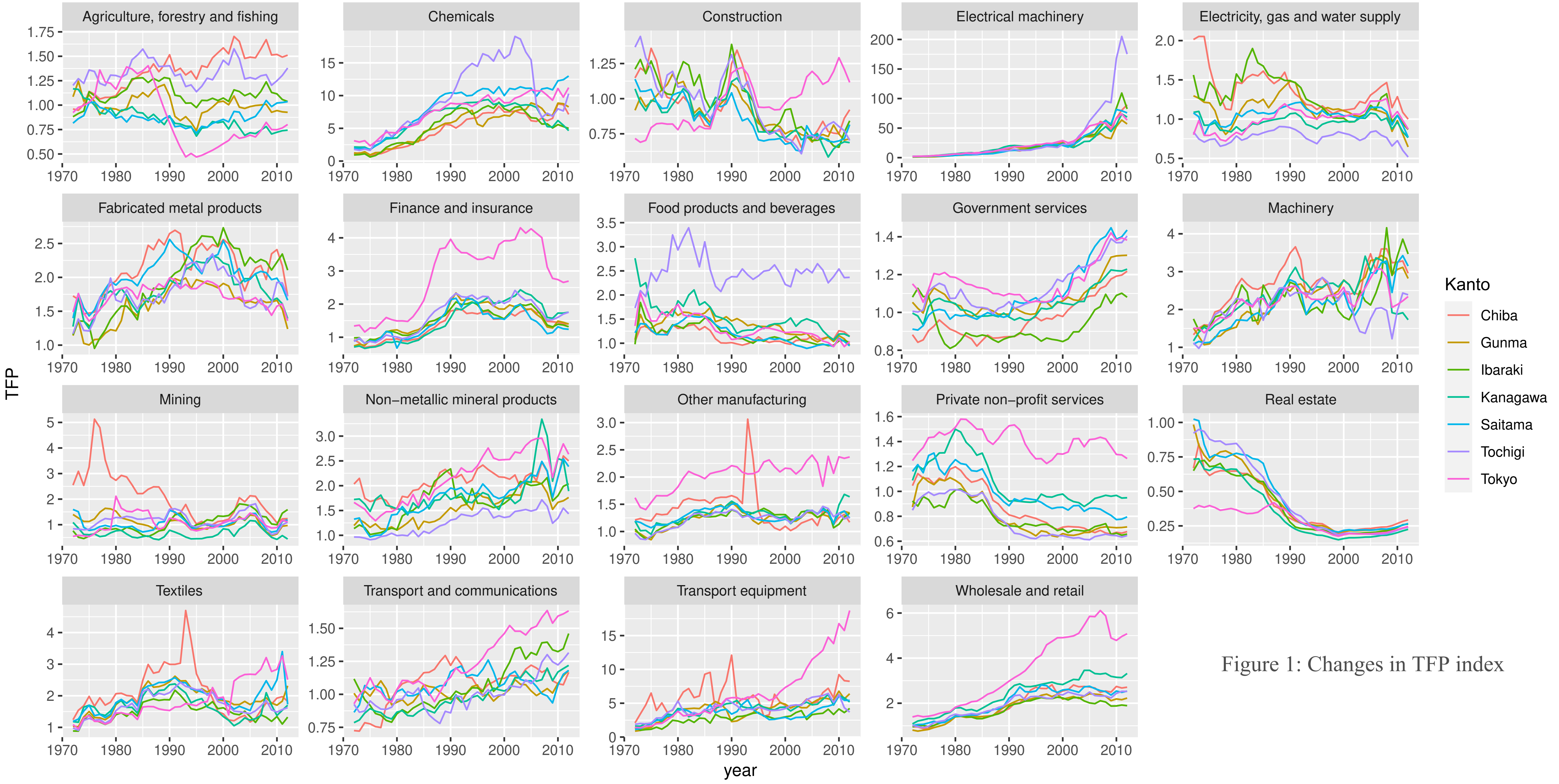

Figure 1: Changes in TFP index

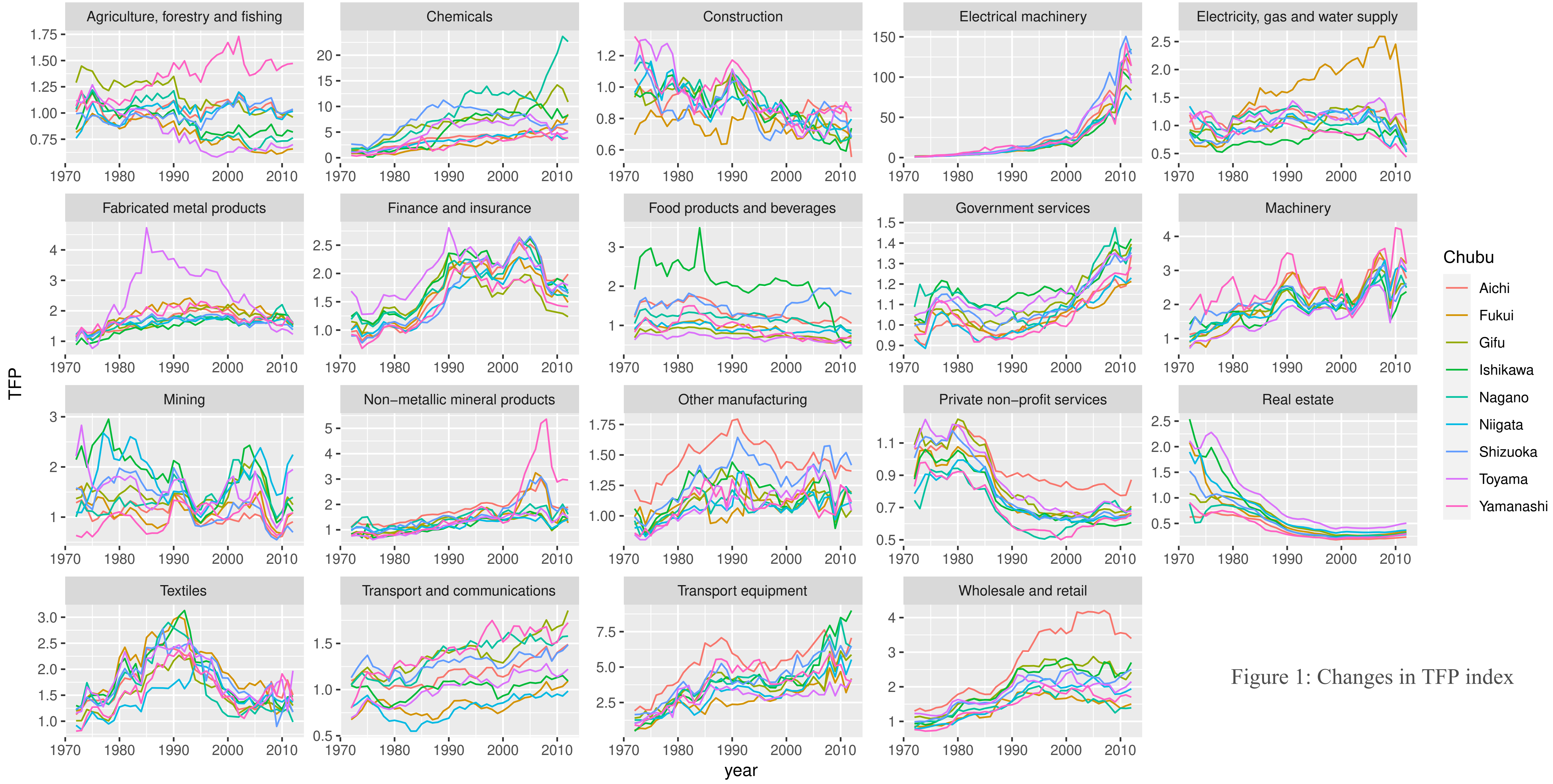

Figure 1: Changes in TFP index

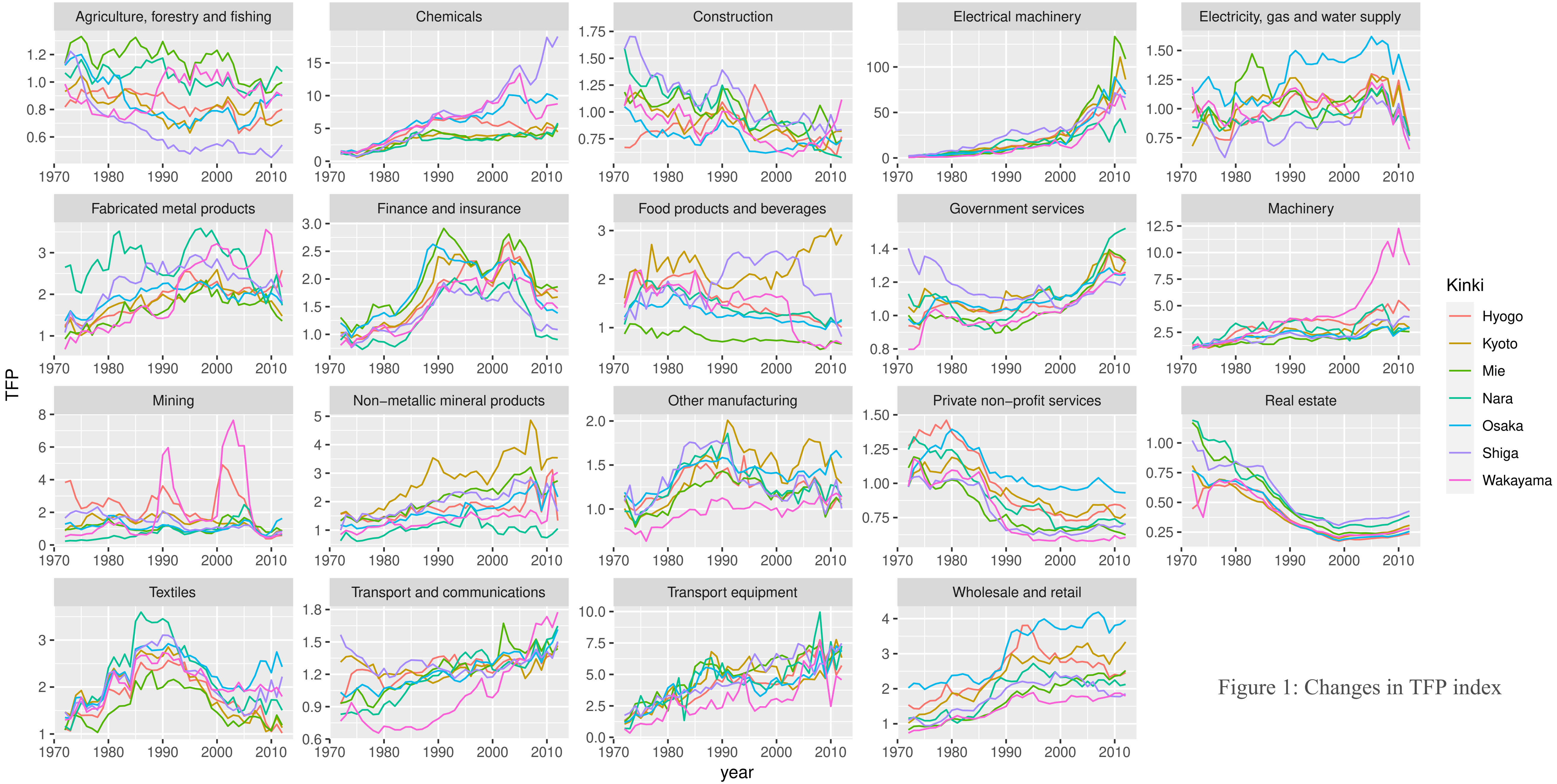

Figure 1: Changes in TFP index

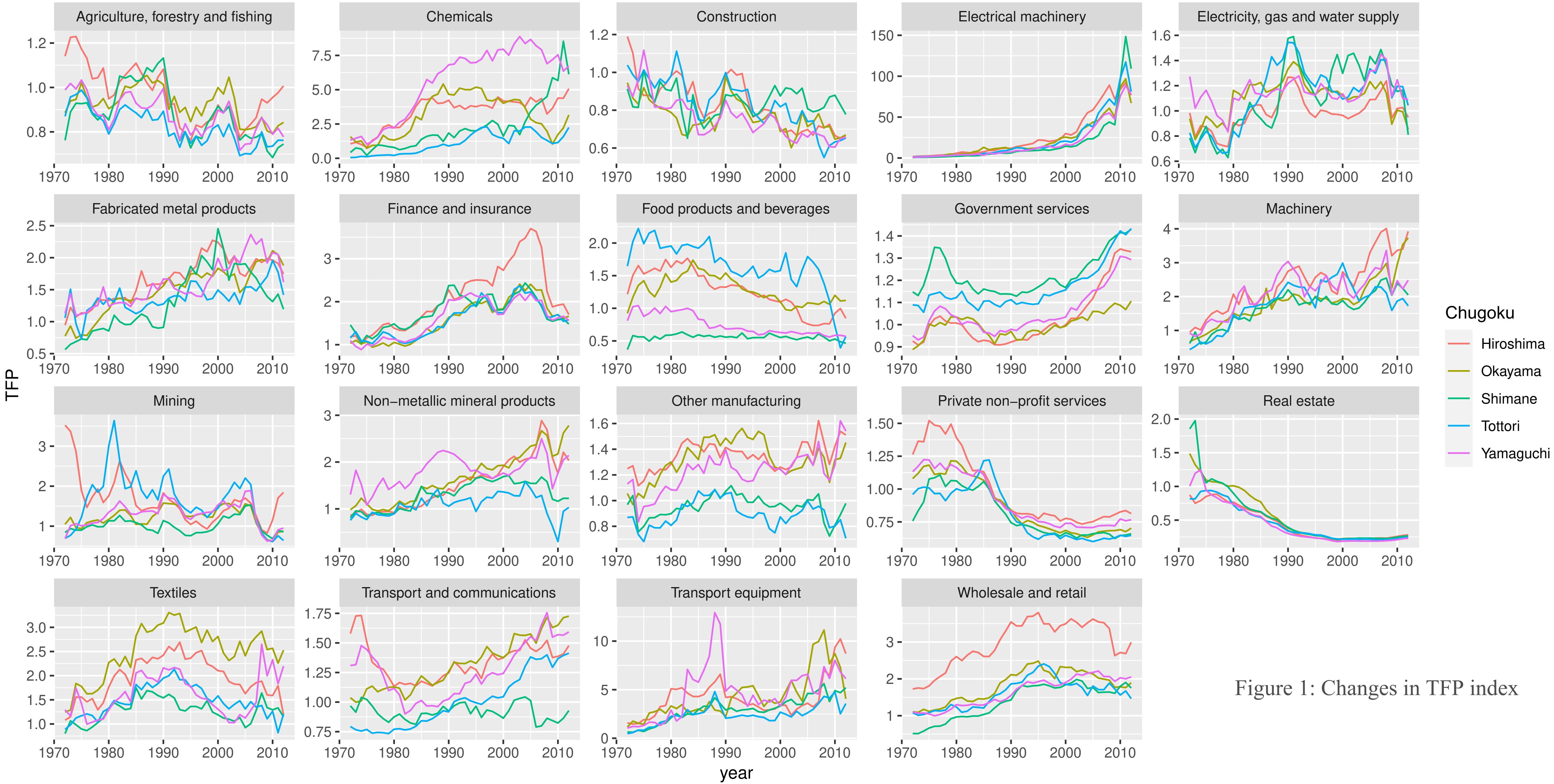

Figure 1: Changes in TFP index

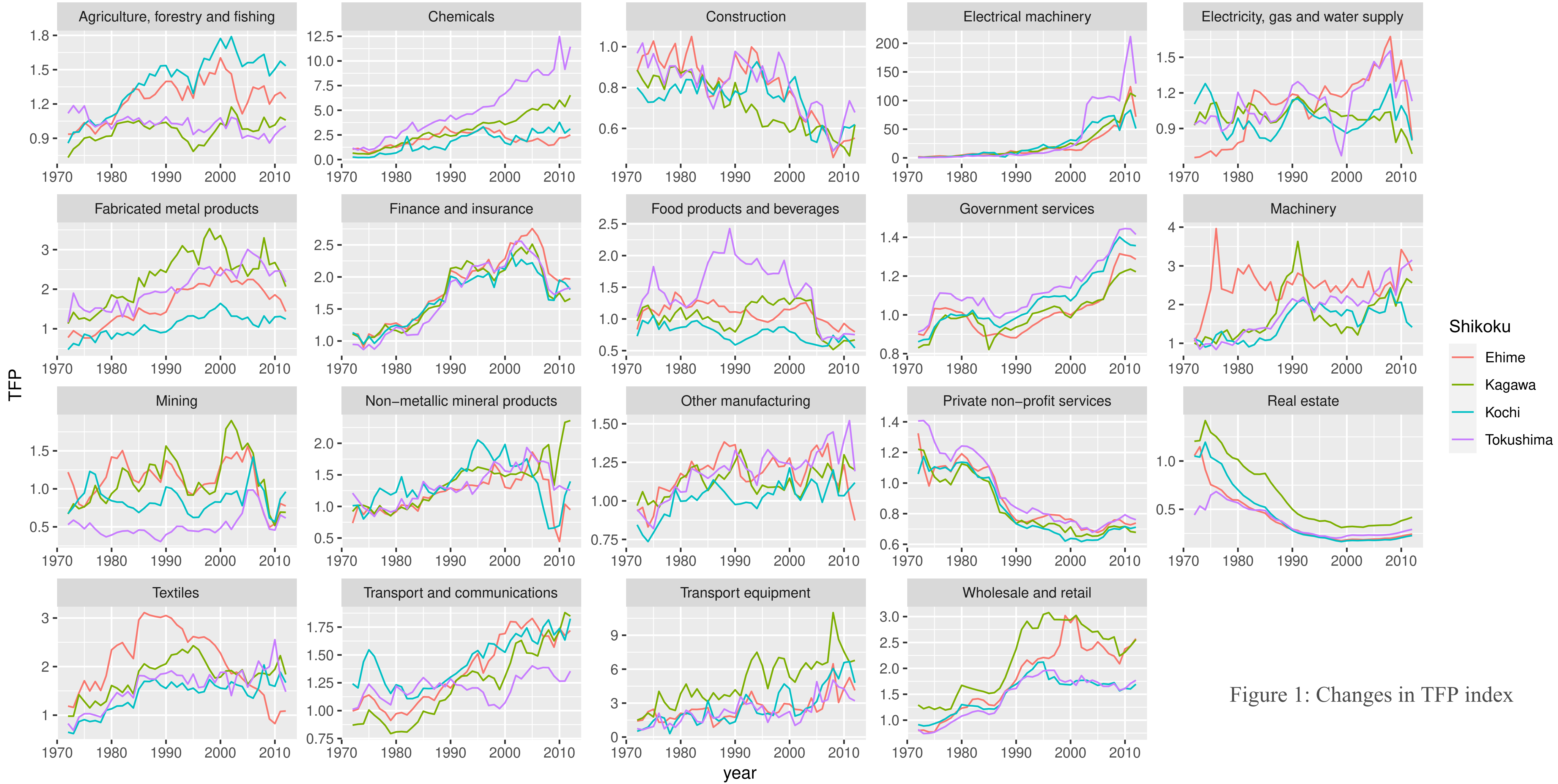

Figure 1: Changes in TFP index

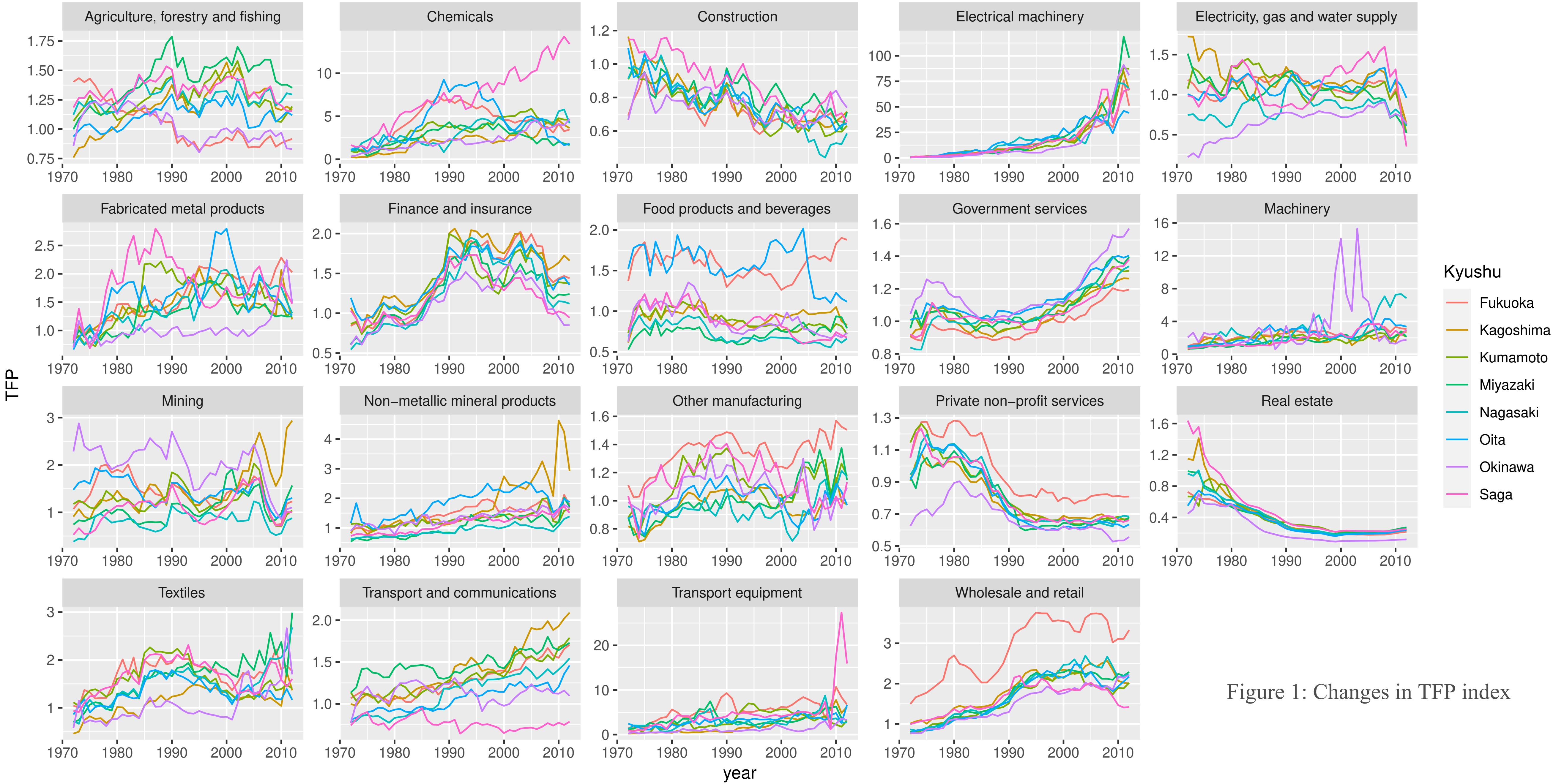

Figure 1: Changes in TFP index

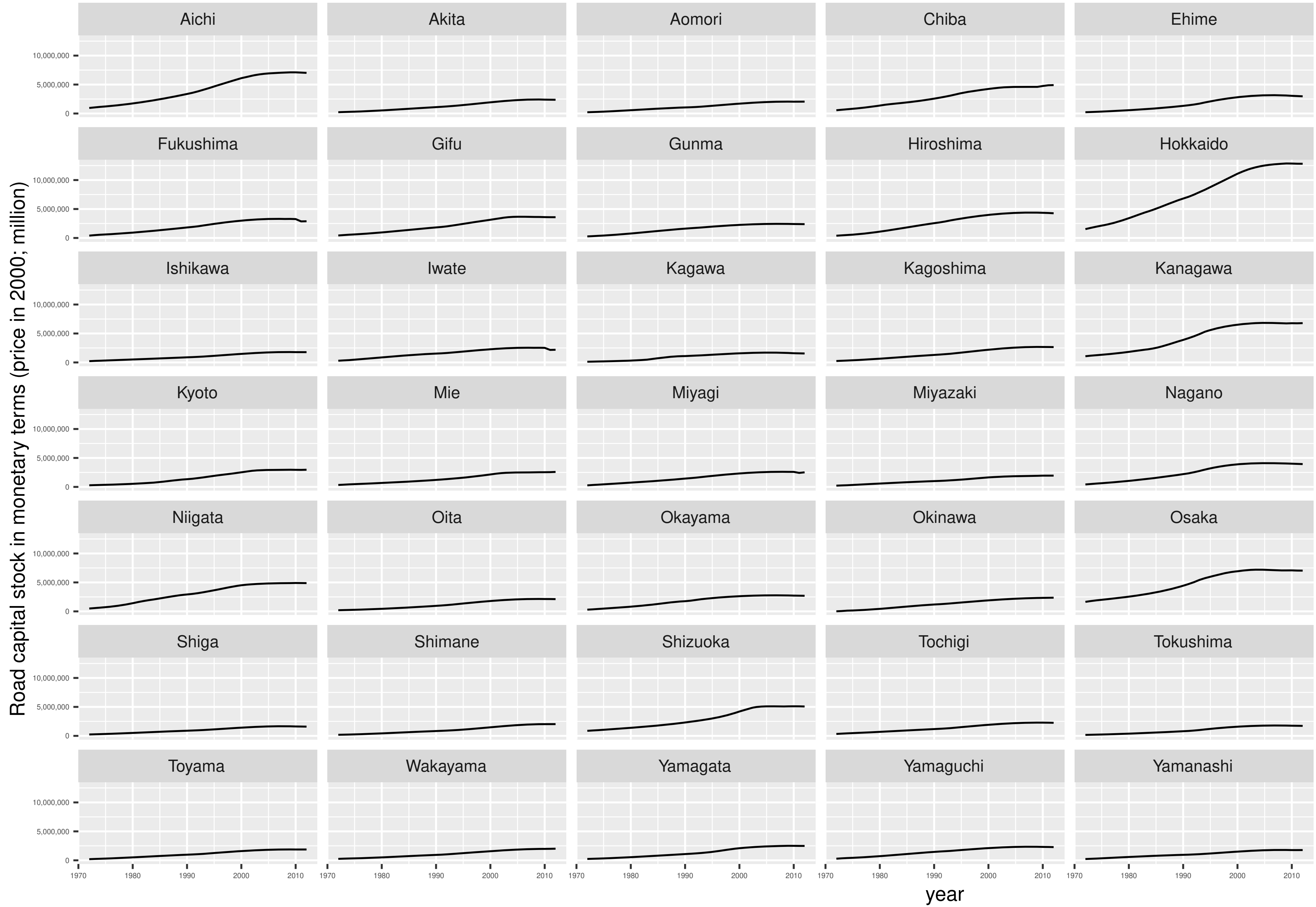

Figure 2: "Toll roads" + "non-toll roads"

Tables

Table 1: List of industries with negative real value added

Year	Prefecture	Sector	Real value added (price in 2000; million)
1974	Ehime	Petroleum and coal products	–33446.54
1982	Mie	Petroleum and coal products	–17413.19
2011	Tottori	Pulp, paper and paper products	–1726.92
1986	Yamaguchi	Petroleum and coal products	–381841.57
2008	Ehime	Petroleum and coal products	–11350.18
2009	Ehime	Basic metal	–6111.88
1972	Kochi	Precision instruments	–59.28
2008	Kochi	Precision instruments	–3815.36
2009	Kochi	Precision instruments	–3354.57
2010	Kochi	Precision instruments	–8706.00
2011	Kochi	Precision instruments	–13354.76
2012	Kochi	Precision instruments	–7061.16
1972	Oita	Basic metal	–1364.17
1988	Okinawa	Petroleum and coal products	–17601.48
1972	Okinawa	Precision instruments	–73.58

Table 2: Results of panel unit root tests (road infrastructure)

Variable: Road capital stock		Level			First-difference		
Method		lags average	Statistic	p-values	lags average	Statistic	p-values
IPS	W-t-bar	2.26	14.0	1.000	1.39	–8.03	0.000
LLC	Adjusted t*	2.26	1.10	0.865	1.39	–4.86	0.000

Table 3: Results of panel unit root tests (TFP)

Sector	Variable: TFP		Level equations		
	Method		lags average	Statistic	p-values
Agriculture, forestry and fishing	IPS	W-t-bar	0.24	–4.67	0.000
	LLC	Adjusted t*	0.24	–6.68	0.000
Mining	IPS	W-t-bar	1.35	–9.73	0.000
	LLC	Adjusted t*	1.35	–10.9	0.000
Food products and beverages	IPS	W-t-bar	1.2	–8.33	0.000
	LLC	Adjusted t*	1.2	–4.34	0.000
Machinery	IPS	W-t-bar	0.78	–6.86	0.000
	LLC	Adjusted t*	0.78	–7.34	0.000
Electrical machinery	IPS	W-t-bar	1.22	–5.39	0.000
	LLC	Adjusted t*	1.22	–4.74	0.000
Transport equipment	IPS	W-t-bar	0.72	–7.15	0.000
	LLC	Adjusted t*	0.72	–7.05	0.000
Other manufacturing	IPS	W-t-bar	0.96	–1.74	0.041
	LLC	Adjusted t*	0.96	–3.91	0.000
Construction	IPS	W-t-bar	2.04	–5.60	0.000
	LLC	Adjusted t*	2.04	–4.33	0.000
Transport and communications	IPS	W-t-bar	0.87	–5.01	0.000
	LLC	Adjusted t*	0.87	–5.47	0.000

Table 3: Results of panel unit root tests (TFP) *cont.*

Sector	Variable: TFP		Level			First-difference		
	Method		lags average	Statistic	p-values	lags average	Statistic	p-values
Textiles	IPS	W-t-bar	1.15	0.67	0.749	1.33	-29.5	0.000
	LLC	Adjusted t*	1.15	-4.24	0.000	1.33	-22.0	0.000
Chemicals	IPS	W-t-bar	1.61	-1.04	0.150	1.13	-31.7	0.000
	LLC	Adjusted t*	1.61	-5.79	0.000	1.13	-27.2	0.000
Non-metallic mineral products	IPS	W-t-bar	1.57	-2.96	0.002	1.91	-28.3	0.000
	LLC	Adjusted t*	1.57	-1.26	0.104	1.91	-19.6	0.000
Fabricated metal products	IPS	W-t-bar	0.98	2.73	0.997	0.93	-39.4	0.000
	LLC	Adjusted t*	0.98	-2.40	0.008	0.93	-31.9	0.000
Electricity, gas and water supply	IPS	W-t-bar	1.8	2.93	0.998	1.67	-20.0	0.000
	LLC	Adjusted t*	1.8	9.02	1.000	1.67	-13.7	0.000
Wholesale and retail	IPS	W-t-bar	0.39	9.04	1.000	0.76	-22.8	0.000
	LLC	Adjusted t*	0.39	1.19	0.883	0.76	-21.6	0.000
Finance and insurance	IPS	W-t-bar	0.65	11.4	1.000	0.28	-26.2	0.000
	LLC	Adjusted t*	0.65	5.77	1.000	0.28	-26.4	0.000
Real estate	IPS	W-t-bar	2.15	17.3	1.000	1.28	-14.6	0.000
	LLC	Adjusted t*	2.15	13.8	1.000	1.28	-11.6	0.000
Private non-profit services	IPS	W-t-bar	1.87	4.44	1.000	0.63	-21.1	0.000
	LLC	Adjusted t*	1.87	5.38	1.000	0.63	-19.5	0.000
Government services	IPS	W-t-bar	0.52	8.00	1.000	0.98	-25.4	0.000
	LLC	Adjusted t*	0.52	1.62	0.948	0.98	-22.4	0.000

Table4: Results of the panel analysis

Sector	Two-way fixed effects Coef.	Std.Err.	t	P>|t|	First-difference Coef.	Std.Err.	t	P>|t|	PMG (ARDL (1,1)) Coef.	Std.Err.	z	P>|z|
Agriculture, forestry and fishing	0.2408	0.03107	7.75	0.000	0.1037	0.03936	2.63	0.008	-0.01148	0.02122	-0.54	0.589
Mining	0.04173	0.06636	0.63	0.530	0.2445	0.1047	2.34	0.020	0.2099	0.08031	2.61	0.009
Food products and beverages	0.08406	0.03849	2.18	0.029	0.5210	0.05748	9.06	0.000	-0.01566	0.02190	-0.71	0.475
Textiles	-0.03988	0.03775	-1.06	0.291	0.4443	0.07302	6.09	0.000	0.4746	0.03893	12.2	0.000
Chemicals	-0.09111	0.07517	-1.21	0.226	0.6308	0.1178	5.35	0.000	1.730	0.07910	21.9	0.000
Non-metallic mineral products	-0.09861	0.04054	-2.43	0.015	0.1628	0.07427	2.19	0.029	0.2619	0.03251	8.06	0.000
Fabricated metal products	0.2663	0.03633	7.33	0.000	0.6127	0.06306	9.72	0.000	0.4932	0.02679	18.4	0.000
Machinery	0.1229	0.04751	2.59	0.010	0.3201	0.08804	3.64	0.000	0.1303	0.04429	2.94	0.003
Electrical machinery	0.2679	0.06136	4.37	0.000	0.3770	0.1071	3.52	0.000	-0.1613	0.04748	-3.4	0.001
Transport equipment	-0.02693	0.07789	-0.35	0.730	0.5368	0.1665	3.22	0.001	0.3489	0.05694	6.13	0.000
Other manufacturing	-0.06369	0.02189	-2.91	0.004	0.04658	0.04650	1.00	0.317	0.3115	0.02305	13.5	0.000
Construction	-0.1839	0.02747	-6.70	0.000	-0.05906	0.04731	-1.25	0.212	-0.03074	0.02335	-1.32	0.188
Electricity, gas and water supply	-0.04459	0.03553	-1.26	0.210	0.3059	0.05851	5.23	0.000	0.9006	0.06293	14.3	0.000
Wholesale and retail	0.03426	0.02414	1.42	0.156	0.3016	0.03495	8.63	0.000	0.7365	0.02986	24.7	0.000
Finance and insurance	-0.07961	0.02496	-3.19	0.001	0.3480	0.04943	7.04	0.000	1.156	0.05512	21.0	0.000
Real estate	-0.1535	0.03198	-4.80	0.000	-0.6402	0.03886	-16.5	0.000	Does not converge			
Transport and communications	0.02795	0.02728	1.02	0.306	-0.06546	0.03079	-2.13	0.034	-0.08366	0.01946	-4.30	0.000
Private non-profit services	-0.1042	0.01584	-6.58	0.000	0.01097	0.02508	0.44	0.662	-0.6866	0.03567	-19.3	0.000
Government services	0.04829	0.01064	4.54	0.000	-0.1334	0.01620	-8.23	0.000	-0.3440	0.01352	-25.4	0.000
Year dummy	Yes				No				Yes			
Prefecture dummy	Yes				No				No			
Constants	Yes				Yes				Yes			