INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 400

Factor Decomposition of Income Inequality Change: Japan's Regional Income Disparity from 1955 to 1998

Takayuki HIGASHIKATA*

March 2013

Abstract

We propose a method for the decomposition of inequality changes based on panel data regression. The method is an efficient way to quantify the contributions of variables to changes of the Theil T index while satisfying the property of uniform addition. We illustrate the method using prefectural data from Japan for the period 1955 to 1998. Japan experienced a diminishing of regional income disparity during the years of high economic growth from 1955 to 1973. After estimating production functions using panel data for prefectures in Japan, we apply the new decomposition approach to identify each production factor's contributions to the changes of per capita income inequality among prefectures. The decomposition results show that total factor productivity (residual) growth, population change (migration), and public capital stock growth contributed to the diminishing of per capita income disparity.

Keywords: income inequality, factor decomposition, Japan, regional disparity

JEL classification: C63, O15, R12

---

* Southeast Asian Studies Group I, Area Studies Center, IDE (Takayuki_Higashikata@ide.go.jp)

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

Factor Decomposition of Income Inequality Change: Japan's Regional Income Disparity from 1955 to 1998

Takayuki HIGASHIKATA*

Abstract

We propose a method for the decomposition of inequality changes based on panel data regression. The method is an efficient way to quantify the contributions of variables to changes of the Theil T index while satisfying the property of uniform addition. We illustrate the method using prefectural data from Japan for the period 1955 to 1998. Japan experienced a diminishing of regional income disparity during the years of high economic growth from 1955 to 1973. After estimating production functions using panel data for prefectures in Japan, we apply the new decomposition approach to identify each production factor's contributions to the changes of per capita income inequality among prefectures. The decomposition results show that total factor productivity (residual) growth, population change (migration), and public capital stock growth contributed to the diminishing of per capita income disparity.

Keywords: income inequality, factor decomposition, Japan, regional disparity
JEL Classification: C63, O15, R12

*Institute of Developing Economies (IDE-JETRO): Takayuki_Higashikata@ide.go.jp

I Introduction

After the financial crisis in 2008, widening income inequality has again attracted worldwide public attention. In the United States, for example, mass demonstrations were conducted by Occupy Wall Street in 2011, in which participants expressed their dissatisfaction with the income gap between the top 1% of earners and the rest. Data on income inequality shows that the share of the top 1% of income earners increased from 8.0% in 1981 to 17.4% in 2011 in the United States, and from 7.1% to 9.5% in Japan.[1]

What caused these income inequality changes? Many factors, such as skill-biased technological changes, the weakening bargaining power of unions, and changes in returns to education, have been suggested as causes of the widening wage or income inequality trend (Lemieux 2010). From the macroeconomic perspective, influenced by studies on the correlation between inequality and growth,[2] and convergence in the growth literature (Barro and Sala-i-Martin 2004), growth factors such as capital stock, technology, education, and institutions have been investigated as driving forces of inequality change.

To identify and quantify the sources of inequality, researchers have been interested in the decomposition of inequality indices. This decomposition can be classified into two types: category-based (population group) decomposition and regression-based decomposition.

The category-based approach divides samples into discrete categories, for example, by province, urban/rural residence, and educational attainment, and then identifies how large the income gaps are within and between categories (Mookherjee and Shorrocks 1982). This approach can be a useful descriptive tool, though it has certain limitations, as Morduch and Sicular (2002) have pointed out. First, the decomposition can only be carried out over discrete categories. Second, handling multiple factors is often unwieldy since the number of groups increases multiplicatively with the number of categories for each factor. Finally, there is a lack of control for endogeneity.

The second approach is to decompose the inequality index into each factor's contribution after estimating a model, for example, an earning function. Starting from the development of the Oaxaca-Blinder decomposition method (Oaxaca 1973), a wide variety of decomposition methods have been used, because regression-based approaches can overcome the limitations of category-based decomposition methods (Morduch and Sicular 2002, Bourguignon, Fournier and Gurgand 2001, Fields 2003).

Morduch and Sicular (2002) examined inequality decompositions by income source through comparison of properties of indices such as the Gini index, Theil T index, coefficient of variation (CV), and squared CV.[3] They introduced the property of uniform additions which states that measured inequality should fall if everyone in the population receives a positive transfer of equal size. After

[1] Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty and Emmanuel Saez, The World Top Incomes Database, http://topincomes.g-mond.parisschoolofeconomics.eu/, accessed on January 25, 2013.

[2] Perotti (1996) surveyed on the positive correlation between equality and growth, and divided the literature into 4 approaches: the fiscal policy, sociopolitical instability, borrowing constraints/investment in education, and endogenous fertility approaches. On the other hand, using nonparametric methods, Banerjee and Duflo (2003) showed that net changes in inequality are associated with reduced growth in the next period.

[3] The Theil L index was not considered because of its unattractive properties (Shorrocks 1983).

examining inequality indices, they concluded that only the Theil T index and squared CV satisfy this principle.[4] However, we cannot directly use their method for decompositions based on regression models in logarithmic form. Because of this limitation, their method has not been fully exploited in the analyses of factor decomposition of income inequality.

Another approach which has been becoming more popular recently is a Shapley value decomposition of inequality indices as proposed by Shorrocks (1999). Wan, Lu and Chen (2007) and Wan (2004) adopted this decomposition procedure to obtain the contributions of variables to any inequality index using estimated production functions models.[5] The method is more attractive than that of Morduch and Sicular (2002) because it can be applied to any estimation model (Wan 2007). However, it has some drawbacks. For example, zero income decomposition should be avoided, and results are sensitive to the design of the income tree (Shorrocks 1999, Sastre and Trannoy 2002).

The aim of this paper is to propose a new method for factor decomposition of inequality index changes with the property of uniform addition which was defined by Morduch and Sicular (2002). The method makes it possible to apply estimators from regressions with logarithmic variables to the changes of factor decomposition between two terms. We use the method to illustrate the regional income inequality among prefectures in Japan. Japan experienced high economic growth as a whole from 1955 to 1973 (9% on average) together with declining individual and regional income inequality. The merit of analyzing Japan's experience is that we have high-quality regional data. Not only labor and private capital stock data, but also public capital stock data are available. Japan's experience could provide useful information to those economies facing high and rising income inequality among regions, especially populous developing countries with vast territories, such as China, India, Brazil, and Indonesia.[6]

The rest of the paper is organized as follows. In Section II we propose a new decomposition approach for change of inequality indices. In Section III , we apply the method to regional income disparity in Japan from 1955 to 1998 after estimating a production function using data on labor, private capital, and public capital. Section IV concludes this paper.

II Method for Factor Decomposition of Changes in Theil T Index

In this section, we propose a new method for factor decomposition of changes in the Theil T index. First, we propose a decomposition method for individual income inequality changes. Next, the method is applied to grouped data, such as regional per capita income. Following the notation of Morduch and Sicular (2002), we assume that an individual gains income from K different sources, the total income

[4] They applied their regression-based decomposition method to household data from China. The empirical results demonstrated that the Theil T decomposition and the squared CV decomposition have the same signs for each variable, though the Gini decomposition has opposite signs.

[5] For brief explanation of Shapley decomposition, see Chakravarty, Deutsch and Silber (2008) or Wan and Zhou (2005).

[6] Milanovic (2005) studied the five most populous countries in the world (China, India, Brazil, Indonesia, and the United States) from1980 to 2000, and showed that inequality between regional mean incomes, and inequality between population-weighted regional mean incomes in all Asian countries were rising in the period 1980-2000.

of individual i is $y_i = \sum_{k=1}^{K} y_{k,i}$. And the Theil T index of the total income of individual i is:

$$\mathrm{I}(\mathbf{y}) \equiv \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{\mu} \ln \frac{y_i}{\mu}$$

where $\mu \equiv \frac{1}{n} y_i$. The contribution of $\mathrm{y_k}$ to $\mathrm{I}(\mathbf{y})$ is

$$\mathrm{I}(\mathbf{y})^{y_k} \equiv \frac{1}{n} \sum_{i=1}^{n} \frac{y_{k,i}}{\mu} \ln \frac{y_i}{\mu}$$

and the proportional contribution θ_k of source k to overall inequality is

$$\theta_k \equiv \frac{\mathrm{I}(\mathbf{y})^{y_k}}{\mathrm{I}(\mathbf{y})} = \frac{\frac{1}{n} \sum_i \frac{y_{k,i}}{\mu} \ln \frac{y_i}{\mu}}{\frac{1}{n} \sum_i \frac{y_i}{\mu} \ln \frac{y_i}{\mu}} \tag{1}$$

If $y_{k,i}$ is constant and $y_{k,i} = \mu_k$, Eq.(1) is

$$\theta_k = \frac{\frac{\mu_k}{\mu} \cdot (-1) \cdot \overbrace{\frac{1}{n} \sum_{i=1}^{n} \ln \frac{\mu}{y_i}}^{Theil\ L\ index}}{\frac{1}{n} \sum_i^n \frac{y_i}{\mu} \ln \frac{y_i}{\mu}} \tag{2}$$

In this case, if μ_k is positive, θ_k becomes negative because the numerator of Eq.(2) is the Theil L index multiplied by a negative number.

As explained in the previous section, a shortcoming of Morduch and Sicular (2002) method is that all variables must be in level, not logarithmic, form. However, this limitation can be avoided as follows. Let us consider a panel data set of individual earnings in which $y_{i,t}$ represents the earnings of individual i in time period t. If we assume that the individual earning equation is

$$\ln y_{i,t} = \sum_k a_k \ln x_{i,t}^k + \sum_m b_m z_{i,t}^m \quad x \geq 0 \text{ and } z \geq 0$$

where x and z are exogenous explanatory variables. Applying the first-order Taylor expansion around the point t = 0, we have

$$y_{i,t} \simeq y_{i,0} + y_{i,0}\left(\sum_k a_k (\ln x_{i,t}^k - \ln x_{i,0}^k) + \sum_m b_m (z_{i,t}^m - z_{i,0}^m) \right) \tag{3}$$

So, we can decompose the difference in earnings inequality between times 0 and τ (> 0) as follows.

$$\begin{aligned} \Delta \mathrm{I}(\mathbf{y}) &\equiv \mathrm{I}(\mathbf{y}_\tau) - \mathrm{I}(\mathbf{y_0}) \\ &= \frac{1}{n} \sum_i^n \frac{y_{i,\tau}}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau} - \frac{1}{n} \sum_i^n \frac{y_{i,0}}{\mu_0} \ln \frac{y_{i,0}}{\mu_0} \\ &= \frac{1}{n} \sum_i \frac{y_{i,\tau} - \tilde{y}_{i,\tau}}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau} + \frac{1}{n} \sum_i \frac{y_{i,0}}{\mu_0} \left(\ln \frac{y_{i,\tau}}{\mu_\tau} - \ln \frac{\tilde{y}_{i,\tau}}{\mu_\tau} \right) \end{aligned} \tag{4}$$

where $\tilde{y}_{i,\tau} \equiv \frac{\mu_\tau}{\mu_0} y_{i,0}$, and by applying the first-order Taylor expansion,

$$\tilde{y}_{i,\tau} \equiv \frac{\mu_\tau}{\mu_0} y_{i,0} \simeq y_{i,0} + y_{i,0}\left[\sum_k a_k\left(\ln \frac{\overline{x}_t^k}{\overline{x}_0^k} x_0{}^k - \ln x_0^k\right) + \sum_m b_m\{(\overline{z}_t{}^m - \overline{z}_0{}^m + z_{i,0}^m) - z_{i,0}^m\} \right.$$
$$\left. + \underbrace{\ln \frac{\mu_t}{\mu_0} - \ln \prod_k \left(\frac{\overline{x}_t^k}{\overline{x}_0^k}\right)^{a_k} \prod_m \left(\frac{e^{\overline{z}_t^m}}{e^{\overline{z}_0^m}}\right)^{b_m}}_{(\,\equiv \ln S\,)} \right] \tag{5}$$

The quantity $\ln S$ represents the difference between a realized average income growth and a counterfactual income growth in which each income generating resource was allocated evenly among individuals. So, $\ln S$ can be regarded as the index of relative allocative efficiency. If $\ln S < 0$, then the average income could have grown more, at least, with evenly allocated resources. Substituting Eq.(5) into Eq.(4), we get

$$\begin{aligned}
\Delta \mathrm{I}(\mathbf{y}) = &\sum_k a_k\left[\left(\frac{1}{n}\sum_i \frac{y_{i,0}\,(\ln x_{i,\tau}^k - \ln x_{i,0}^k)}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau}\right) - \left(\frac{1}{n}\sum_i \frac{y_{i,0}\,(\ln \tilde{x}_{i,\tau}^k - \ln x_{i,0}^k)}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau}\right)\right] \\
&+ \sum_k a_k\left[\frac{1}{n}\sum_i \frac{w_i^k}{\overline{x}_\tau^k}\left(\ln \frac{x_{i,\tau}^k}{\overline{x}_\tau^k} - \ln \frac{\tilde{x}_{i,\tau}^k}{\overline{x}_\tau^k}\right)\right] \\
&+ \sum_m b_m\left[\left(\frac{1}{n}\sum_i \frac{y_{i,0}\,(z_{i,\tau}^m - z_{i,0}^m)}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau}\right) - \left(\frac{1}{n}\sum_i \frac{y_{i,0}\,(\tilde{z}_{i,\tau}^m - z_{i,0}^m)}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau}\right)\right] \\
&+ \omega \sum_m b_m\left[\frac{1}{n}\sum_i \frac{w_i^m}{1+\frac{\overline{z}_\tau^m}{\omega}}\left(\ln \frac{1+\frac{z_{i,\tau}^m}{\omega}}{1+\frac{\overline{z}_\tau^m}{\omega}} - \ln \frac{1+\frac{\tilde{z}_{i,\tau}^m}{\omega}}{1+\frac{\overline{z}_\tau^m}{\omega}}\right)\right] \\
&- \ln S\left[1 + \left(\frac{1}{n}\sum_i \frac{y_{i,0}}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau}\right)\right] \\
= &\sum_k a_k\left[\mathrm{I}(\mathbf{y}_\tau)^{\mathbf{y_0}(\ln \mathbf{x}_\tau^\mathbf{k} - \ln \tilde{\mathbf{x}}_\tau^\mathbf{k})} + \left(\mathrm{I}(\mathbf{x}_\tau^\mathbf{k})^{\mathbf{w^k}} - \mathrm{I}(\tilde{\mathbf{x}}_\tau^\mathbf{k})^{\mathbf{w^k}}\right)\right] + \sum_m b_m\left[\mathrm{I}(\mathbf{y}_\tau)^{\mathbf{y_0}(\mathbf{z}_\tau^\mathbf{m} - \tilde{\mathbf{z}}_\tau^\mathbf{m})} + \omega\left(\mathrm{I}(\mathbf{1}+\tfrac{\mathbf{z}_\tau}{\omega})^{\mathbf{w^m}} - \mathrm{I}(\mathbf{1}+\tfrac{\tilde{\mathbf{z}}_\tau}{\omega})^{\mathbf{w^m}}\right)\right] \\
&- \ln S\left(1 + \mathrm{I}(\mathbf{y}_\tau)^{\mathbf{y_0}}\right)
\end{aligned} \tag{6}$$

where $\tilde{x}_{i,\tau}^k \equiv \left(\frac{\overline{x}_\tau^k}{\overline{x}_0^k}\right) x_{i,0}^k$, $\tilde{z}_{i,\tau}^m \equiv \overline{z}_\tau^m - \overline{z}_0^m + z_{i,0}^m$, $w_i^k \equiv \left(\frac{y_{i,0}}{\mu_0}\right)\overline{x}_t^k$, $w_i^m \equiv \frac{y_{i,0}}{\mu_0}\left(1 + \frac{\overline{z}_\tau}{\omega}\right)$, and $\omega\ (>0)$ is a weight for the approximation of e^z.[7]

This transformation shows that a factor's contribution to the change of the Theil T index between times 0 and τ consists of three components: 1) the difference between the contribution of each factor's realized growth on the total income inequality in the second period and that of a counterfactual situation, in which each income generating factor grows at the same rate, as if a central planner allocated each additionally provided resource, 2) the difference between the realized factor endowment

[7] According to first-order Taylor expansion at $c = 0$, $e^c \simeq 1 + c$. So, if ω is large enough, we can transform $\ln e^z/e^{\overline{z}}$ as

$$\ln \frac{e^z}{e^{\overline{z}}} = \omega \ln \frac{e^{z/\omega}}{e^{\overline{z}/\omega}} \simeq \omega \ln \frac{1 + z/\omega}{1 + \overline{z}/\omega}$$

inequality and the counterfactual one with the same weight for each factor (w_i^k, or w_i^m), and 3) the effect of relative allocation efficiency weighted by $1+\mathrm{I}(\mathbf{y}_\tau)^{\mathbf{y_0}}$ as an overall contribution. And if we have time series data for income between times 0 and τ, we can sum up each decomposition result to illustrate each factor's contribution for that period ($\mathrm{I}(\mathbf{y}_\tau) - \mathrm{I}(\mathbf{y_0}) = \sum_{t=1}^{\tau} \mathrm{I}(\mathbf{y_t}) - \mathrm{I}(\mathbf{y_{t-1}})$).

Application to regional income disparity

The above inequality decomposition method for individual panel data can be applied to aggregated panel data of regional per capita income. If the per capita income in region i at time t is $y_{i,t} \equiv \frac{Y_{i,t}}{N_{i,t}}$, where $Y_{i,t}$ is the regional income, and $N_{i,t}$ is the regional population, we have a weighted mean of regional per capita income

$$\mu_t = \sum_i^m p_{i,t} y_{i,t} = \frac{\sum Y_{i,t}}{\sum N_{i,t}}$$

where $p_{i,t} \equiv \frac{N_{i,t}}{\sum_i N_{i,t}}$. If we assume that the production function has the Cobb-Douglass form with constant returns to scale

$$\ln Y_{i,t} = \alpha + (1-\beta)\ln L_{i,t} + \beta \ln K_{i,t} + a_{i,t}, \quad 0 < \beta < 1$$

where L is the number of workers, K is the capital stock, and a is total factor productivity (TFP), or the residuals, then the Theil T index for the per capita income is

$$\mathrm{I}(\mathbf{y_t}) = \sum_i^m p_{i,t} \frac{y_{i,t}}{\mu_t} \ln \frac{y_{i,t}}{\mu_t}$$

The differences in regional per capita income inequality ($\Delta \mathrm{I}(\mathbf{y})$) between times 0 and τ are decomposed as follows:

$$\begin{aligned}
\Delta \mathrm{I}(\mathbf{y}) &\equiv \mathrm{I}(\mathbf{y}_\tau) - \mathrm{I}(\mathbf{y_0}) \\
&= \sum_i^m p_{i,\tau} \frac{y_{i,\tau}}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau} - \sum_i^m p_{i,0} \frac{y_{i,0}}{\mu_0} \ln \frac{y_{i,0}}{\mu_0} \\
&= \sum_i p_{i,\tau} \frac{y_{i,\tau} - \overbrace{\frac{p_{i,0}}{p_{i,\tau}} \frac{\mu_\tau}{\mu_0} y_{i,0}}^{(\equiv \tilde{y}_{i,\tau})}}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau} + \sum_i p_{i,0} \frac{y_{i,0}}{\mu_0} \left(\ln \frac{y_{i,\tau}}{\mu_\tau} - \ln \frac{y_{i,0}}{\mu_0} \right) \\
&= (1-\beta) \sum_i p_{i,\tau} \frac{y_{i,0} (\ln l_{i,\tau} - \ln \tilde{l}_{i,\tau})}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau} + \beta \sum_i p_{i,\tau} \frac{y_{i,0} (\ln k_{i,\tau} - \ln \tilde{k}_{i,\tau})}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau} \\
&\quad + \sum_i p_{i,\tau} \frac{y_{i,0} (a_{i,\tau} - \tilde{a}_{i,\tau})}{\mu_\tau} \ln \frac{y_{i,\tau}}{\mu_\tau} \\
&\quad + (1-\beta) \sum_i p_{i,\tau} \frac{w_i^l}{\mu_\tau^l} \left(\ln \frac{l_{i,\tau}}{\mu_\tau^l} - \ln \frac{\tilde{l}_{i,\tau}}{\mu_\tau^l} \right) + \beta \sum_i p_{i,\tau} \frac{w_i^k}{\mu_\tau^k} \left(\ln \frac{k_{i,\tau}}{\mu_\tau^k} - \ln \frac{\tilde{k}_{i,\tau}}{\mu_\tau^k} \right) \\
&\quad + \omega \sum_i p_{i,\tau} \frac{w_i^{1+\frac{a}{\omega}}}{\mu_\tau^{1+\frac{a}{\omega}}} \left(\ln \frac{1 + \frac{a_{i,\tau}}{\omega}}{\mu_\tau^{1+\frac{a}{\omega}}} - \ln \frac{1 + \frac{\tilde{a}_{i,\tau}}{\omega}}{\mu_\tau^{1+\frac{a}{\omega}}} \right) - \sum_i p_{i,\tau} \frac{w_i^n}{\mu_\tau^n} \left(\ln \frac{n_{i,\tau}}{\mu_\tau^n} - \ln \frac{\tilde{n}_{i,\tau}}{\mu_\tau^n} \right)
\end{aligned}$$

$$-\underbrace{\left[\ln\frac{\mu_t}{\mu_0}-\ln\left(\frac{\mu_\tau^l}{\mu_0^l}\right)^{(1-\beta)}\left(\frac{\mu_\tau^k}{\mu_0^k}\right)^{\beta}\left(\frac{e^{\bar{a}_\tau}}{e^{\bar{a}_0}}\right)\right]}_{(\,\equiv\ln S\,)}\left(1+\sum_i p_{i,\tau}\frac{y_{i,0}}{\mu_\tau}\ln\frac{y_{i,\tau}}{\mu_\tau}\right) \tag{7}$$

where for any variable X, $\bar{X}_t \equiv \frac{1}{m}\sum X_{i,t}$, $\tilde{X}_{i,t} \equiv (\bar{X}_t/\bar{X}_0)X_{i,0}$, $x_{i,t} \equiv X_{i,t}/N_{i,t}$, $\tilde{x}_{i,t} \equiv \tilde{X}_{i,t}/N_{i,t}$, $\mu_t^x \equiv \sum p_{i,t}x_{i,t} = \sum p_{i,t}\tilde{x}_{i,t} = \bar{X}_t/\bar{N}_t$, $\bar{a}_t \equiv \frac{\sum a_{i,t}N_{i,t}}{\sum N_{i,t}}$, $\tilde{a}_{i,t} \equiv \bar{a}_t - \bar{a}_0 + a_{i,0}$ and $\omega\,(>0\,)$ is a weight.

III Regional Income Disparity in Japan from 1955 to 1998

In this section, we apply the above decomposition method to Japan's income disparity change from 1955 to 1998, and identify the contributions of the various factors in the production function.[8] Our motivation for investigating Japan's regional income inequality arises from its historical experience of the high economic growth with diminishing income inequality starting from around 1955 until 1973 when the oil shock hit the country.

The income inequality of Japan in the period from 1955 to 1998, measured by the Gini coefficient on income (before income redistribution), gradually declined to 0.354 in 1972 although it increased once in 1961 to 0.390. After the oil shock, the Gini coefficient jumped up to 0.375 in 1975, and then it declined to 0.349 in 1981. However, the Gini index gradually crept up during the period of the asset price bubble economy from 1984 and the depression after the subsequent collapse in 1991, reaching 0.472 in 1999 (Tachibanaki 2006).

[Figure1]

The same trend is captured in the data of regional per capita income inequality. Figure 1 plots the time series of regional income inequality measured by the Theil T index from 1955 to 1998. As the figure illustrates, overall inequality decreased constantly during the 1960s. From the middle of the 1970s to 1980 in the post oil shock decade, the degree of inequality hovered under 0.03, but steadily increased to reach 0.05 in 1989. After that, unlike the results using the household data, regional inequality dropped again to 0.038 and remained almost flat until the end of the period.

We analyzed the regional income inequality of Japan, especially in the 1960s, because the analysis can provide useful information to developing countries, such as China, India, and Indonesia that are facing widening income gaps among individuals and regions during their higher economic growth period. Data for the analysis and variables for estimations of the production function are available for Japan as far back as 1955, so it is possible to determine the factors behind the change of income inequality from the perspective of economic growth. We end our analysis in 1998 because a series of datasets for private and public capital stock, taking into consideration the effects of privatization of public companies, are available only until 1998.

Before starting our analysis, it would be useful to review briefly the literature on regression-based factor decomposition approaches to the analysis of regional income inequality in Japan. From

[8] We excluded Okinawa prefecture from the analysis because some data are not available for Okinawa from 1955 to 1971.

the viewpoint of economic growth theory, that is, the β convergence approach, Fukao and Yue (2000) decomposed their estimator for the convergence of labor productivity into contributions of production factors, and concluded that from 1955 to 1973 labor (migration) contributed most to the convergence, followed by public capital stock and human capital stock. But in the estimation for β convergence, some serious problems such as the endogeneity problem need to be solved in order to avoid biases in estimators (Acemoglu 2009).[9]

On the other hand, from the viewpoint of the cross-sectional dispersion of incomes (σ convergence), Yamano and Ohkawara (2000) examined the role of public capital on regional income inequalities based on simulations. Apart from these contributions, there is little work on regression-based factor decomposition of regional inequality in Japan, although we have found an enormous volume of the literature dealing with other countries from this perspective, such as the work of Wan, Lu and Chen (2007) and Tsui (2007) for China.

In the following subsections, we introduce the data used for the analysis. Then, we estimate a standard Cobb-Douglas production function model with constant returns to scale. Finally, using parameters from the estimation, we apply the method described in Section II to decompose the change of per capita income inequalities in Japan into contributions from each factor.

1 Data

Japan consists of 47 prefectures, but we have excluded Okinawa due to missing data problems: public capital stock and private capital stock data were not available until 1971 for Okinawa. GDP, public capital, and private capital stock data were all denominated in the 1990 price. Data on prefectural GDPs came from the Cabinet Office.[10]

The prefectures of Fukushima, Saitama, and Okayama lack GDP deflators for some periods.[11] So we used deflators estimated by ordinary least squares (OLS) using the deflators of surrounding prefectures as explanatory variables. Population data was taken from Fukao and Yue (2000) for the period from 1955 to 1974, which originally comes from the Annual Report on Prefectural Accounts for each year,[12] and from the Cabinet Office for the period from 1975 to 1998. Employment data came from Doi (1998) for the period 1955 to 1974 and from the Cabinet Office for the period from 1975 to 1998. Finally, private and public capital stock data were taken from Doi (2002). These data were adjusted for the effects of the privatization of three public companies and the devastating earthquake in 1995.

[9] Shioji (2001) carefully treated those problems by adopting the dynamic panel approach to identify the effects of public capital on output per capita in Japan using prefectural data.

[10] All data for which the source is referred to as the Cabinet Office were downloaded from http://www.esri.cao.go.jp/.

[11] Fukushima lacks data from 1975 to 1979, Saitama from 1975 to 1976, and Okayama from 1975 to 1984.

[12] Downloadable from http://www.ier.hit-u.ac.jp/ fukao/japanese/data/index.html

2 Production Function Estimation

The prefecture-level production function is assumed to be a standard Cobb-Douglas production function with constant returns to scale. The production function is:

$$Y_{i,t} = A_{i,t} L_{i,t}^{\beta_L} K_{i,t}^{\beta_K} G_{i,t}^{\beta_G}, \quad \beta_L + \beta_K + \beta_G = 1$$

where $Y_{i,t}$ is the gross prefecture domestic product, $A_{i,t}$ is TFP or residual, $L_{i,t}$ is the number of employed persons, $K_{i,t}$ is the private capital stock, $G_{i,t}$ is the government capital stock, i is the subscript for prefecture, and t is the year. The production function expressed in terms of GDP per employed person is

$$\ln y_{i,t} = \beta_0 t + \beta_K \ln k_{i,t} + \beta_G \ln g_{i,t} + \gamma_t + c_i + e_{i,t}$$

where $y_{i,t}$ is the GDP per employed person, $k_{i,t}$ is the private capital stock per employed person, $g_{i,t}$ is the public capital stock per employed person, γ_t is the set of year dummy variables, c_i represents time-invariant prefecture effects, and $e_{i,t}$ is the idiosyncratic shock. We estimated the first-difference equation:

$$\Delta \ln y_{i,t} = \beta_0 + \beta_K \Delta \ln k_{i,t} + \beta_G \Delta \ln g_{i,t} + \Delta\gamma_t + \epsilon_{i,t}, \quad \epsilon_{i,t} \sim N(0, \sigma^2) \tag{8}$$

[Table1]

Table 1 gives the estimation results for Eq. (8). Column (1) reports the OLS regression result for the first-difference estimator with the full sample. Taking into consideration the structural change after the oil shock in 1973, we also estimated models limiting the period from 1955 to 1972, and from 1973 to 1998. Columns (2) and (3) contain the results obtained for these limited samples.

The results in columns (1) and (2) are similar to those of Shioji (2001). He showed that the infrastructure component of public capital had significant positive effects on economic growth, implying that elasticity of output with respect to infrastructure is somewhere around 0.1 to 0.15 in the period 1955 to 1990. In column (3), the coefficient of public capital stock is not statistically significant as many studies have pointed out. Using the 1994 electoral reform as a natural experiment, Kawaguchi, Ohtake and Tamada (2009) reported that they could not reject the null hypothesis that public capital stock did not increase labor productivity. In addition, Fukao and Yue (2000) also showed that the marginal productivity of public capital was not statistically significant for the period of 1973 to 1995 in their translog production function model.

3 Factor Decomposition of changes in regional inequality

Using the results of column (1) in Table 1, we applied the decomposition method of Section II to identify the contributions of each production factor.

[Figure2]

Figure 2 displays each factor's contribution to the change of the Theil T index for each year from 1955 to 1998. It is evident from Figure 2 that TFP and population change contributed negatively to the inequality index change for almost all years until 1975. On the other hand, labor force growth widened the degree of inequality. However, it is difficult to see the contributions of private capital stock and public capital stock in the figure. In order to clarify the overall contribution from each production factor, we added up the factors changes in each year for all the years in the sample, and for the periods before and after the oil shock ($\sum_{t=1}^{T} \mathrm{I}(\mathbf{y_t}) - \mathrm{I}(\mathbf{y_{t-1}})$).

[Table2]

Table 2 shows the factor decomposition results for the periods of (1) 1955- 1998, (2) 1956-1972, and (3) 1973-1998. For the whole period (column (1)) TFP change contributed most to the decrease of regional income inequality followed by growth of population (migration) and public capital stock. This indicates that TFP and public capital stock in the poorer prefectures grew faster than the national average rate.[13] In contrast, the growth of the labor force and private capital stock increased the regional inequality. In addition, it is worth noting that the effects of the labor force canceled out those of the population growth that must have accompanied the migration from rural to urban areas.

As is also illustrated in Figure 1, the first row of Table 2 shows that total inequality tended to worsen in the later 25-year period. Although this is mainly caused by the decrease of (downward) contributions from TFP and population, the effects of the latter should not be overstated because we should take into account the contribution of labor force too. The combination of those two effects had contributions that were almost the same for the first period (0.0267) and the second period (0.0243).

Finally, it should be noted that the actual per capita income (national per capita income) growth did not reach the counterfactual growth in which each resource was evenly allocated to the population ($\ln S < 0$).[14]

IV Conclusion

In this paper we have proposed a new method for inequality change decomposition. The approach is based on regression estimations that satisfy the property of uniform addition as defined by Morduch and Sicular (2002). We applied the method to the analysis of Japan's regional income inequality changes from 1955 to 1998, and decomposed the changes of the Theil T index into production factors, such as TFP (residual), labor, private capital stock, public capital stock, population, and allocation efficiency.

Our results show that TFP and public capital stock grew faster in the less developed regions, and this led to the diminishing per capita income inequality among prefectures in Japan. On the other hand, the growth rates of private capital stock, labor, and population were higher in the more

[13] We must be especially careful when considering the public capital result because, as Table 1 shows, the coefficient for the public capital stock for the second period is not statistically significant.

[14] In our analysis, the value of $\mathrm{I}(\mathbf{y}_\tau)^{\mathbf{y_0}}$ is always positive, but small. Table 2 shows that the contribution of *resource allocation efficiency* to the overall inequality change is 0.0132, of which only 0.0005 comes from the term $-\ln S \cdot \mathrm{I}(\mathbf{y}_\tau)^{\mathbf{y_0}}$.

affluent regions. As a policy implication, we would like to emphasize the role of public capital stock. However, as showed in our estimation, the parameter for public capital stock was not statistically significant after 1973. As Shioji (2001) emphasized, we need to disaggregate public capital stock into components to clarify its contribution.

References

Acemoglu, Daron. 2009. *Introduction To Modern Economic Growth.* Princeton Univ Press.

Banerjee, Abhijit V, and Esther Duflo. 2003. "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth*, 8(3): 267–99.

Barro, Robert J., and Xavier Sala-i-Martin. 2004. *Economic Growth, 2nd Edition.* . 2 ed., The MIT Press.

Bourguignon, Francois, Martin Fournier, and Marc Gurgand. 2001. "Fast Development with a Stable Income Distribution: Taiwan, 1979-94." *Review of Income and Wealth*, 47(2): 139–63.

Chakravarty, Satya R., Joseph Deutsch, and Jacques Silber. 2008. "On the Watts Multidimensional Poverty Index and its Decomposition." *World Development*, 36(6): 1067–1077.

Doi, Takero. 1998. "Panel analysis of public capital in Japan (in Japanese)." *Kokumin Keizai*, 27–52.

Doi, Takero. 2002. *Regional Perspective of Japanese Economy and Fiscal Policy (in Japanese).* Mitsubishi Economic Research Institute.

Fields, Gary S. 2003. "Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States." *Research in labor economics*, 22: 1–38.

Fukao, Kyoji, and Xi-ming Yue. 2000. "Regional Factor Inputs and Convergence in Japan: How Much Can We Apply Closed Economy Neoclassical Growth Models? (in Japanese)." *Keizai-kenkyu*, 51(2): 136–151.

Kawaguchi, Daiji, Fumio Ohtake, and Keiko Tamada. 2009. "The productivity of public capital: Evidence from Japan's 1994 electoral reform." *Journal of the Japanese and International Economies*, 23(3): 332–343.

Lemieux, Thomas. 2010. "What Do We Really Know About Changes in Wage Inequality?" In *Labor in the New Economy: 71 (National Bureau of Economic Research Studies in Income and Wealth). NBER Chapters*, , ed. Katharine G. Abraham, James R. Spletzer and Michael Harper, 17–59. National Bureau of Economic Research, Inc.

Milanovic, Branko. 2005. "Half a World: Regional inequality in five great federations." *Journal of the Asia Pacific Economy*, 10(4): 408–445.

Mookherjee, Dilip, and Anthony F Shorrocks. 1982. "A Decomposition Analysis of the Trend in UK Income Inequality." *Economic Journal*, 92(368): 886–902.

Morduch, Jonathan, and Terry Sicular. 2002. "Rethinking Inequality Decomposition, With Evidence from Rural China." *Economic Journal*, 112(476): 93–106.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review*, 14(3): 693–709.

Perotti, Roberto. 1996. "Growth, Income Distribution, and Democracy: What the Data Say." *Journal of Economic Growth*, 1(2): 149–87.

Sastre, Mercedes, and Alain Trannoy. 2002. "Shapley inequality decomposition by factor components: Some methodological issues." *Journal of Economics*, 9(1): 51–89.

Shioji, Etsuro. 2001. "Public Capital and Economic Growth: A Convergence Approach." *Journal of Economic Growth*, 6(3): 205–27.

Shorrocks, Anthony F. 1983. "The Impact of Income Components on the Distribution of Family Incomes." *The Quarterly Journal of Economics*, 98(2): 311–26.

Shorrocks, Anthony F. 1999. "Decomposition Procedures for Distributional Analysis: A unified Framework Based on the Shapley Value." *mimeo*.

Tachibanaki, Toshiaki. 2006. "Inequality And Poverty In Japan." *The Japanese Economic Review*, 57(1): 1–27.

Tsui, Kai-yuen. 2007. "Forces Shaping China's Interprovincial Inequality." *Review of Income and Wealth*, 53(1): 60–92.

Wan, Guanghua. 2004. "Accounting for income inequality in rural China: a regression-based approach." *Journal of Comparative Economics*, 32(2): 348–363.

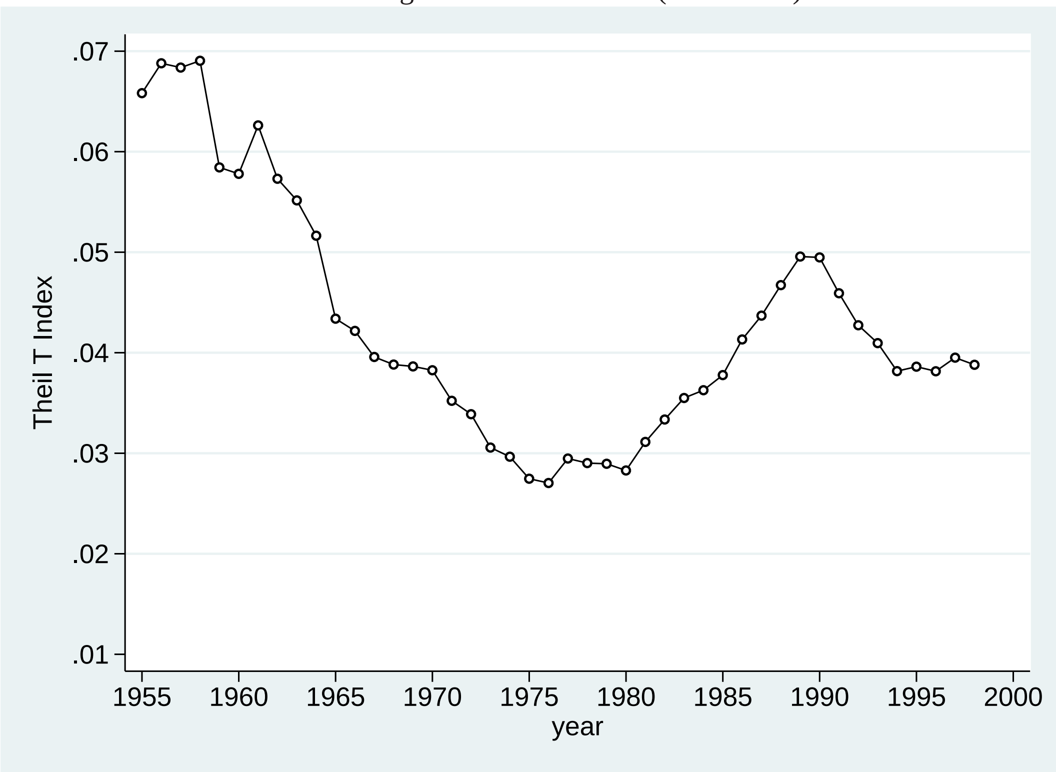
Wan, Guanghua. 2007. "Understanding Regional Poverty And Inequality Trends In China: Methodological Issues And Empirical Findings." *Review of Income and Wealth*, 53(1): 25–34.

Wan, Guanghua, and Zhangyue Zhou. 2005. "Income Inequality in Rural China: Regression-based Decomposition Using Household Data." *Review of Development Economics*, 9(1): 107–120.

Wan, Guanghua, Ming Lu, and Zhao Chen. 2007. "Globalization And Regional Income Inequality: Empirical Evidence From Within China." *Review of Income and Wealth*, 53(1): 35–59.

Yamano, Norihiko, and Toru Ohkawara. 2000. "The Regional Allocation of Public Investment: Efficiency or Equity?" *Journal of Regional Science*, 40(2): 205–229.

Figure 1: Theil T Index (1955-1998)



Source: Calculated using data from the Statistics Bureau's homepage.

Note: Excluding Okinawa.

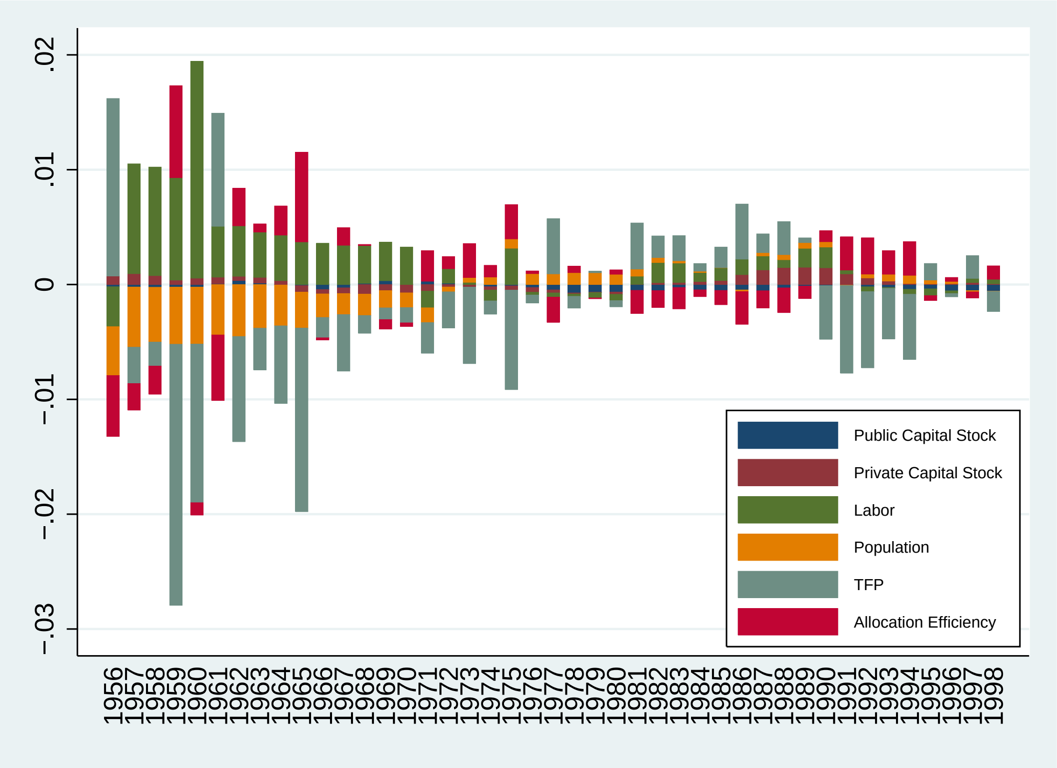
Table 1: The Estimation of Production Function (Dependent variable: $\Delta ln(Y/L)$)

	(1) 1956-98	(2) 1956-72	(3) 1973-98
$\Delta ln(K/L)$	0.254***	0.149*	0.408***
	(0.047)	(0.065)	(0.070)
$\Delta ln(G/L)$	0.104*	0.183***	0.024
	(0.041)	(0.053)	(0.066)
Constant	0.020***	0.055***	−0.003
	(0.003)	(0.006)	(0.004)
Observations	1932	736	1150
Adjusted R^2	0.628	0.274	0.410

Source: Author's estimation.

Note: Huber robust standard errors in parentheses. * significant at 5 %, ** significant at 1 %, and *** significant at 0.1 %. Results for year dummies are not shown.

Figure 2: Decomposition Results (1955-1998)



Source: Author's estimation.

Note: Allocation Efficiency includes that of weighted by $I(\mathbf{y}_\tau)^{\mathbf{y}_0}$.

Table 2: Decomposition Results

	(1) 1955 - 1998	(2) 1955 - 1972	(3) 1973-1998
Total	−0.0270	−0.0319	0.0049
Population	−0.0422	−0.0539	0.0117
Labor	0.0932	0.0806	0.0126
Private capital stock	0.0088	0.0008	0.0079
Public capital stock	−0.0104	−0.0006	−0.0099
TFP (Residual)	−0.0896	−0.0684	−0.0212
Allocation efficiency	0.0132	0.0095	0.0036

Source: Author's calculation.