Rethinking of Gini Inequality from Grouped and Individual Observations: Examples from the Vietnamese Household Expenditure Data 1998.

Keio University, Policy Management
Tadashi KIKUCHI[1]

This paper introduces some useful policy-minded approaches to measure the economic impacts on social inequality. The Gini inequality or the Gini concentration ratio is calculated both from grouped and from individual data. Because of its flexibility to obtain the Gini inequality, particular attention is paid to the simplicity and the usefulness of political simulations. First the process of estimating the Lorenz curve is explained, which related to the Gini inequality as well known. Next, by utilizing individual data, we can figure out which personal character is influential to increase or decrease the Gini inequality. An empirical illustration is also presented from the Vietnamese survey data of 5,938 households' consumption expenditure, a part of Vietnam Living Standards Survey 1997-98.

JEL: C13, D33, D63, O12, O53, R20, R58

1. Introduction.

The Lorenz curve is widely used for measuring to what extent a developing country is becoming poor and losing the balance of nation's income distribution by political scientists and policy-minded economists. However, judging the inequality only from one aspect may be misleading to describe the developing countries' real problem. Economists need to seek more appropriate approaches from multi-dimensional aspects. The reason is because the gap of the rich and the poor comes from certain groups and usually they have many characteristics, for example urban/rural residential, the number of members in their households and household owner's educational level.

The equation of the Lorenz curve from grouped data is simply and efficiently derived from the density function of income distribution, as Kakwani and Podder's paper showed in 1976. The purpose of this paper is to introduce the essence of Kakwani and Podder's paper as a useful tool to find out social inequality as well as to make use of it for political analysis. Especially when analyzing the influence of state policy on inequality, utilizing individual data gives not only clear and useful vision, but also opportunity to reevaluate the policy.

[1] Lecturer of Keio University, Policy Management
mail: tkikuchi@sfc.keio.ac.jp

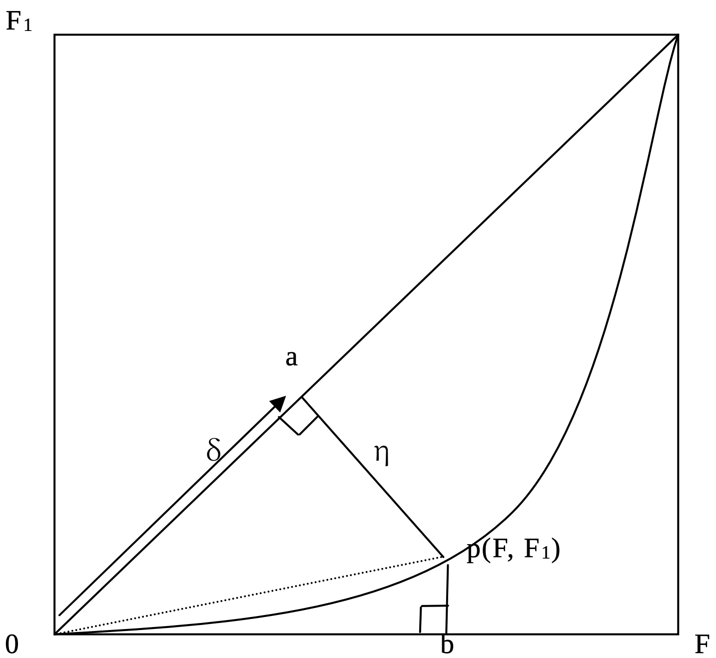
The definition and function of the Lorenz curve are provided in the next section. Section 3 describes the estimation of the Lorenz curve and the Gini inequality from grouped observations. Section 4 provides an alternative aspect of the Gini inequality, its mean and variance is also able to obtain from individual data. Section 5 reports some empirical results and simulation outputs, by using consumption data from the Vietnamese household survey 1997- 1998. The last section provides brief summary and comments.

2. Figure of the Lorenz and the Gini inequality

In this section, I owe notations below mainly to Kakwani and Podder[1976]. Here, the relationship between the Lorenz curve and the Gini inequality is introduced. Suppose that income X of a family is a random variable with probability distribution function F(x) and μ denotes the mean of the distribution. Then F_1, the first moment distribution function of X, is defined.

(2.1) $$F_1(x) = \frac{1}{\mu}\int_0^x Xg(x)\,dX$$

where g(x) is the density function. The Lorenz curve is the relationship between F(x) and $F_1(x)$. The curve is shown as below and the line $F(x) = F_1(x)$, the diagonal of the unit square, is called the egalitarian line.



Let p be any point on the curve with co-ordinates (F, F_1),and

(2.2) $$\delta = \frac{1}{\sqrt{2}}(F+F_1) \qquad \eta = \frac{1}{\sqrt{2}}(F-F_1)$$

then η will be the length of the ordinate from p to the egalitarian line and δ will be the distance of the ordinate from the origin along the egalitarian line. Using the Pythagorean theorem to △oap and △opb is helpful to understand above.

Here, income is assumed positive, the equation (2.2) imply η to be less than or equal to δ. Then, the equation of the Lorenz curve is defined in terms of δ and η.

(2.3) $\eta = f(\delta)$

where δ varies from zero to $\sqrt{2}$. Although there are some candidates of the function f(.), it is defined here, as [2]

(2.4) $\eta = a\delta^{\alpha}(\sqrt{2}-\delta)^{\beta} \qquad a>0, \qquad \alpha>0, \quad \text{and } \beta>0$

The restriction $a>0$ assures that $\eta \geqq 0$, i.e., the Lorenz curve lies below the egalitarian line. Further, from the sigh of parameters α and β, we can find as

(2.5) If $\alpha < \beta$, then the curve is skewed toward (1,1).
If $\alpha = \beta$, then the curve is symmetric.
If $\alpha > \beta$, then the curve is skewed toward (0,0).

In other words, each parameter α, β works as a kind of weight to the poor and the rich, respectively. I will investigate the relation above from the Vietnamese households' data later.

3. Estimating the Lorenz curve from grouped observations.

We can expand the equation (2.4) as a regression model to estimate the function of the Lorenz curve. However, before estimating parameters from grouped observations, preparing some notations is useful to describe.

Suppose there are N families which have been grouped into (T +1) income class, viz., (0 to x_1), (x_1 to x_2),…,(x_T to x_{T+1}). Let n_t be the number of families earning income in

[2] Kakwani and Podder[1976] also referred other function forms, using the Australian survey of consumer expenditure and finances(1967-1968).

the interval x_{t-1} and x_t; then $f_t = n_t/N$ is the relative frequency.

If x_t^* is the sample mean for the *t*th income group, then the consistent estimates of $F(x_t)$ and $F_1(x_t)$ are respectively

(3.1) $p_t = \sum_{i=1}^{T+1} f_i$ and $q_t = \frac{1}{Q}\sum_{i=1}^{T+1} x_i^* f_i$

where t=1,2,…,T and $Q = \Sigma x_i^* f_i$ is the mean income of all the families. Now using the equation (2.2), the consistent estimators of δ_t and η_t are obtained as

(3.2) $r_t = \frac{p_t + q_t}{\sqrt{2}}$ and $y_t = \frac{p_t - q_t}{\sqrt{2}}$

respectively. To estimate parameters, by taking logarithm of equation (2.4), the regression model is formed as

(3.3) $\log y_t = c + \alpha \log r_t + \beta \log(\sqrt{2} - r_t) + \varepsilon_t$

where c = log a , and ε_t is random disturbance. There are other ways to estimate equation (3.3), but here classical OLS method is used[3]. Because, as we see later, the fitness of the OLS estimation is very good and it shows no problem for the purpose in this paper.

4. The Gini inequality from individual observations

The Gini inequality can be also calculated from individual data, by using the absolute difference and the Gini mean difference in next.

(4.1) Gini mean difference $\Delta = \frac{2}{N(N-1)}\sum_{xj<xi}|x_i - x_j|$

Then

(4.2) Gini $G = \Delta/2\mu$

This is an average of the difference between the two individual's income, x_i and $x_j (i<j)$, of N(N−1) families. If we specify the distribution function F(x) with estimated mean $\hat{\mu}$ and variance $\hat{\tau}^2$, we can also calculate the mean of the Gini inequality. With the help of mathematical expansion in Shibata[1981], they are expressed in a simple form as[4]

---

[3] Kakwani and Podder[1976] also tried other estimating methods, including GLS.

[4] Note that, in Shibata[1981] normal distribution is assumed, then we need to take logarithm before using equation (4.3) here.

(4.3) $E[Gini] = \tau / \mu \sqrt{\pi}$

$$(4.4) \quad Var[Gini] = \frac{\tau^2}{N(N-1)\mu^2}\left\{\left(\frac{2\sqrt{3}-4}{\pi}+\frac{1}{3}\right)N+\frac{6-4\sqrt{3}}{\pi}+\frac{1}{3}\right\}$$

where π is the ratio of the circumference of a circle to its diameter. The implication of equation (4.3) is very interesting, because it tells that doubling income per capita to increase mean income μ is not enough to decrease the Gini inequality. The government also has to pay attention to the variance of income τ^2 to reduce the gap between the rich and the poor. The equation (4.3) assures the convergence of the Gini inequality, as the number of observations N increases[5]. Empirically, it is known that the lognormal distribution fits well to income distribution. Under the assumption of lognormal distribution, next relationship exists[6].

(4.5) $Var(X) = [E(X)]^2[\exp(\tau^2) - 1] \propto [E(X)]^2$

where $Var(X)=\tau^2$, and $E(X)=\mu$. This implies the larger mean income μ, the larger volatility τ, too. It may be rather hard to decrease the Gini index as much as we expect as state economy is developing.

From the individual data, as we see in the next section, the Gini inequality measured by the Vietnamese survey of household data in 1998 is 0.33223 and the estimation of equation (4.3) is .35050[7]. Although lognormal distribution seems to fit well with household data, the Vietnamese household data can not be drawn perfectly by lognormal distribution. This results in the difference between the two Gini inequality indexes above, which is 0.01827. We can also detect influence on the Gini inequality of a certain group, by substituting the estimated variance $\hat{\tau}^2$ from data into the equation (4.3). It is very useful to observe and investigate political impacts on national income distribution, as we will see in next section.

5. Empirical results

In this section, results of estimation of the Lorenz curve and the Gini inequality are presented. The Vietnamese survey of household expenditure data 1998 are used for this purpose, which is a part of the Vietnam Living Standards Survey 1997-98. The method of two stage random sampling is carried out in the nationwide

---

[5] It is quite natural, because we see already the Gini inequality as a kind of an average. We can apply easily the Law of large number to the Gini inequality.

[6] See Minotani[2003] more details.

[7] See footnote 4.

survey by World Bank groups and General Statistical Office in Vietnam. The nature of the survey has been extensively discussed elsewhere like [3]. The data set is made in the form of individual observation from 5,938 households, after excluding outliers out of 6,000 original data and adjusting the price at 1998 price. Here, in stead of household income, household consumption expenditure is used to analyze. Because, in developing countries, including Vietnam, people's living standards is composed of the amount of consumption goods, but not the amount of money they have. Along with this thinking, before analyzing, I also divide each household's consumption expenditure by the appropriate number of household members, which is called equivalence scale in economics, to measure representative household's consumption expenditure. The weight used for this purpose in calculating equivalent scale is 1.00 for an equal or more than sixteen year old adult, 0.445 for between a six and sixteen year old young person and 0.226 for a less six year old child, respectively[8].

By using Individual observation, it is possible to compute not only the actual value of the Gini inequality, but also figure out group's characteristics and impacts to the inequality. This is very useful to examine and evaluate the impacts of the policy in advance, especially when studying on the redistribution of nation's wealth. Two sets of the grouped data, a thirteen and a twenty four classified class respectively, are presented in Table I.

Table I

Household consumption expenditure distribution in 1998

＜attached around here ＞

Table II presents estimated parameters of the Lorenz curve (3.3) and the Gini inequality by using both grouped and individual data.

Table II

Results of the Lorenz curve estimation

＜attached around here ＞

The fitness of regression model is very good and it also assured that assuming density function in equation (2.4) is successful in the analysis. Note also that the ratio of

[8] Tsakloglou[1993] is using 1.00, 0.40 and 0.25 in calculating equivalence scale of Greece in 1974 and 1982. These weight values are interestingly almost the same with those of here, which are calculated originally by author.

estimated parameter $\hat{\beta} / \hat{\alpha}$ is more than twice. This means that household consumption expenditure of Vietnam is much skewed to the rich. The Lorenz curve in 1998 is drawn below.

Figure 1

The Lorenz curve of Vietnamese household consumption expenditure in 1998

＜attached around here ＞

Table III

Actual and estimated y in equation (3.3)

＜attached around here ＞

Table IV

Actual and estimated frequency distribution of household consumption expenditure

＜attached around here ＞

And also in order to compare with actual data, the estimation of y_i, relative frequency and the mean income of each class are given in Table III and Table IV.

Table V

Share of household consumption expenditure:

the poorest and the richest 5, 10 %

＜attached around here ＞

The estimated share of household consumption expenditure occupied by the poorest and the richest 5 or 10 percent of total population is presented in Table V. It is clear in the table that the estimated income share is quite close to the actually calculated value from individual observations. The classical estimation by OLS is successful, here.

Next, by manipulating the individual data set, and re-calculating the Gini inequality, it is able to simulate and measure the influence of the government redistribution policy. Now, assume that taxation is put on the richest group's consumption expenditure at 1% or 10% rate[9]. Then, the amount of tax, which is measured by the richest peoples' consumption power, is redistributed equivalently to the poorest group. The result of this simulation is shown in Table VI.

Table VI

Results of simulation (1):

Transferring consumption power from the richest group to the poorest group

＜attached around here ＞

From the result of Table VI, we can seek for effective means to reduce the gap between the rich and the poor. If putting tax on broader range of the richest group, this purpose is more easily achieved. Note that the effect on decreasing inequality is almost the same in both two cases where imposing 10% tax on the richest 5% group and the case where imposing only 1% tax on the richest 10% group. In this simulation, we assumed that the total amount of household consumption expenditure neither increase nor decrease and only considered on transferring problem among current households. We can also assume that consumption power increases by economic development. In next Table VII, assuming that urban or rural area is developing, then impacts to the Gini inequality is calculated from the 5,938 household data.

Table VII

Results of simulation (2):

Increasing consumption power by urban/rural residential

＜attached around here ＞

In Table VII, we can find out that unbalanced development in urban or rural area might make social equality worse, even when the average of national household consumption power increases. Note that, only by developing in urban area, it seems hard to achieve both reducing the inequality and doubling the national income, which is here measured as household consumption expenditure. These finding results imply us that regional development might get worse social balance and increase the gap between the rich and the poor more than we expect, unless the government monitors properly impacts of regional development on social inequality.

[9] Here, total tax value is set equally to 1% or 10% of the total amount of the richest consumption expenditure. And each rich household pay this kind of tax in proportion to the ratio, his consumption expenditure/ total tax value.

Finally, we can also see the multiple effects on inequality when both social and economic programs are adapted to a certain group. For instance, imagine an educational program is subsidized from abroad like ODA for young people in order that they can increase their household consumption expenditure in future and rural and urban areas keep developing under a state political scheme. Remember that, in this paper, every household is represented by using equivalence scale, so we can consider this situation, as if each household would be representative by the young with his new and progressive educational level in next future.

Table VIII
Output of simulation (3):
Double impacts both from educational promotion and from regional development
＜attached around here ＞

Here, educational level of household owner is categorized into five groups, 1) primary education No completed, 2) primary education completed, 3) lower secondary education completed, 4) upper secondary education completed, 5) university/graduated school completed. From Table VIII, we can observe multiple effects of both from regional development and educational promotion to the Gini inequality. Although, each state program assumed here increases all household's consumption power, however the influence to inequality is different, positive or negative, by each case. Observing results in Table VIII and also comparing them with simulated results in Table VII provide useful implications, as well as, problems for Vietnam economy to achieve "sustainable growth" in future. I guess that one of crucial problems, in Vietnam, lies in that most high education school facilities are located mainly in urban area. In the market economy, if ordinary people would like to get higher wage, they have choice to go to cities under development. However, an unskilled labor because of the lack of his education is always facing the same problem wherever he is.

6. Conclusion

In this paper, I introduced briefly the concept of the Lorenz curve and interpretation of the Gini inequality in a different way from both grouped and individual data, with the help of elaborated pre-studying by Kakwani and Podder[1976]. In the last section, by using individual 5,938 household data of Vietnam in 1998, I have tried to demonstrate the possibility of utilizing the Gini inequality as a clue to reduce the inequality gap in Vietnam. I emphasized on the easiness and simplicity of calculating the Gini inequality, I also hope some information provided in this paper would be useful for policy-minded researchers too. However to draw another accurate and concrete conclusion, I have to confess that modeling dynamic population and other researches such as describing an exact income distribution, checking missing data by utilizing nonparametric methods and rebuilding enough sampling in order to get rid of sampling error, are necessary. These are not mentioned here and remained as my next studying challenges in future.

7. References

[1] BHATTACHARYA, N. and B. MAHALANOBIS: "Regional Disparities in Household Consumption in India", *American Statistical Association Journal*, March (1967), 143 – 161.

[2] COWELL, F. A.: *Measuring Inequality* 2nd edition, 1995, Prentice Hall Harvester Wheatsheaf.

[3] DOLLAR, D., P. GLEWWE, J. LITVACK: *Household welfare and Vietnam's transition*, 1998, World Bank.

[4] FRITZEN, S: "Growth, inequality and the future of poverty reduction in Vietnam", *Journal of Asian Economics*, 13 (2002), 635 – 657.

[5] KAKWANI, N.C., and N.PODDER: "On the Estimation of the Lorenz Curves from Grouped Observations", *International Economic Review*, 14 (1973), 278-291.

[6] KAKWANI, N.C., and N.PODDER: "Efficient Estimation of the Lorenz Curve and Associated Inequality Measures from Grouped Observations", *Econometrica*, 44 (1976), 137 – 148.

[7] KAKWANI, N. C.: "Applications of Lorenz Curves in Economic Analysis", *Econometrica*, 44 (1976), 137 – 148.

[8] MINOTANI, C.: Toukeibunpu Handbook(in Japanese), Asakura shoten(2003).

[9] RAO, V.M.: "Two Decompositions of Concentration Ratio", *Journal of the Royal Statistical Society, Series A*, 132 (1969), 418 – 425.

[10] SHIBATA, Y.: *Seikibunpu* (in Japanese), Tokyodaigaku shuppankai(1981).

[11] TSAKLOGLOU, P.: "Aspects of inequality in Greece Measurement, decomposition and intertemporal change: 1974, 1982", *Journal of development Economics*, 40(1993), 53 – 74 .

[12] YITZHAKI, S.: "On an Extension of the Gini Inequality Index", *International Economic Review*, 24 (1983), 619 – 628.

Table I

Household consumption expenditure distribution in 1998*

class	Income Range $*	Number of Families	Mean income $	p	q	r	y
1	Below 100	270	81.8	0.04547	0.01302	0.04136	0.02294
2	100 – 149	913	128.6	0.19923	0.08229	0.19906	0.08268
3	150 – 199	1258	175.2	0.41108	0.21229	0.44079	0.14057
4	200 – 249	1017	223.2	0.58235	0.34615	0.65655	0.16702
5	250 – 299	645	273.3	0.69097	0.45011	0.80687	0.17031
6	300 – 349	457	322.7	0.76794	0.53710	0.92280	0.16322
7	350 – 399	313	373.4	0.82065	0.60603	1.00882	0.15175
8	400 – 449	223	423.9	0.85820	0.66178	1.07479	0.13889
9	450 – 499	176	473.9	0.88784	0.71097	1.13053	0.12507
10	500 – 549	127	524.4	0.90923	0.75025	1.17343	0.11241
11	550 –599	91	573.9	0.92455	0.78105	1.20605	0.10147
12	600 – 649	80	626.5	0.93803	0.81061	1.23647	0.09009
13	650 – 699	64	673.4	0.94880	0.83603	1.26207	0.07974
14	700 – 749	49	724.0	0.95706	0.85696	1.28270	0.07078
15	750 –799	54	772.7	0.96615	0.88157	1.30653	0.05981
16	800 – 849	32	829.2	0.97154	0.89722	1.32141	0.05255
17	850 –899	29	871.7	0.97642	0.91213	1.33541	0.04546
18	900 –949	33	922.6	0.98198	0.93008	1.35203	0.03670
19	950 –999	21	975.3	0.98552	0.94216	1.36308	0.03066
20	1000 –1049	19	1030.1	0.98872	0.95371	1.37350	0.02476
21	1050 – 1099	16	1074.0	0.99141	0.96384	1.38257	0.01949
22	1100 –1149	10	1132.5	0.99310	0.97052	1.38849	0.01596
23	1150 –1199	16	1178.3	0.99579	0.981641	1.39825	0.01000
24	1200 and over	25	1245.1	1.0000	1.00000	1.41421	0.00000

* $ 1.0 = 13,288 Dong at 98 price

13 class

class	Income Range $*	Number of Families	Mean income $	p	q	r	y
1	Below 100	270	81.8	0.04547	0.01302	0.04136	0.02294
2	100 – 199	2171	155.6	0.41108	0.21229	0.44079	0.14057
3	200 – 299	1662	242.6	0.69097	0.45011	0.80687	0.17031
4	300 – 399	770	343.3	0.82065	0.60603	1.00882	0.15175
5	400 – 499	399	445.9	0.88784	0.71097	1.13053	0.12507
6	500 – 599	218	545.1	0.92455	0.78105	1.20605	0.10147
7	600 – 699	144	647.3	0.94880	0.83603	1.26207	0.07974
8	700 – 799	103	749.5	0.96615	0.88157	1.30653	0.05981
9	800 – 899	61	849.4	0.97642	0.91213	1.33541	0.04546
10	900 – 999	54	943.1	0.98552	0.94216	1.36308	0.03066
11	1000 –1099	35	1050.2	0.99141	0.963842	1.38257	0.01949
12	1100–1199	26	1160.7	0.99579	0.981641	1.39825	0.01000
13	1200 and over	25	1245.1	1.00000	1.00000	1.41421	0.00000

* The price is adjusted at 1998. As calculating representative household consumption, equivalence scale by author is used.

Table II

Results of the Lorenz curve estimation*

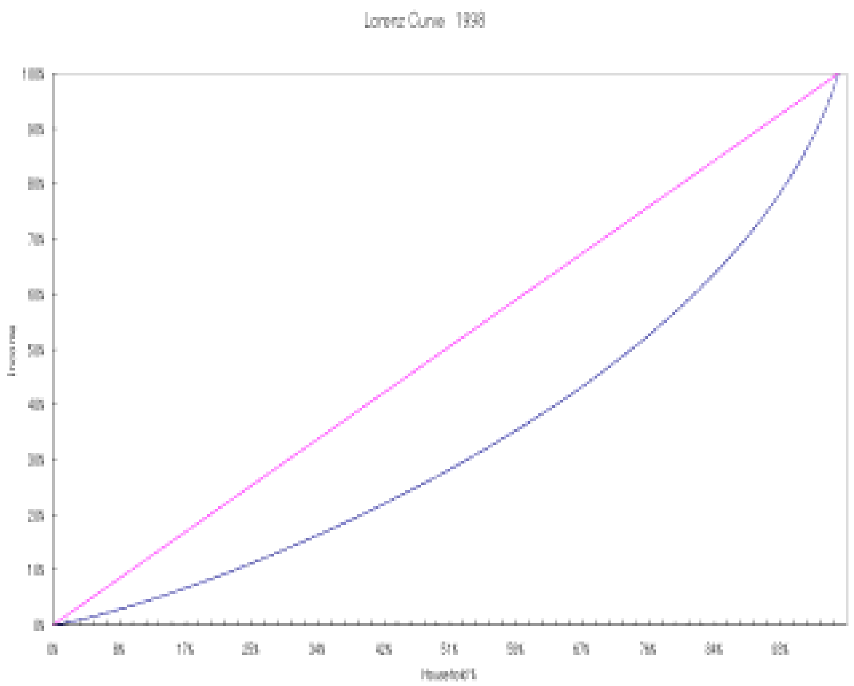
Number of income class	Coefficient estimation		R square	Gini inequality	estimated parameter a in eq(2.4)
24	constant =	-1.07046			
		(0.02046)	0.997664	0.33672	0.34285
	α =	0.950625			
		(0.01550)			
	β =	2.108686			
		(0.02277)			
13	constant =	-1.06756			
		(0.0319)	0.99768	0.31352	0.34384
	α =	0.960161			
		(0.02025)			
	β =	2.103773			
		(0.03403)			
Actual Value of Inequality from Household data				0.33223	

* All estimated parameter are significant at 1% in Table II.

Figure in a parenthesis is standard error.

Figure 1

The Lorenz curve of Vietnamese household consumption expenditure in 1998*



* The price is adjusted at 1998 and equivalence scale is used to draw the Lorenz curve.

Table III

Actual and estimated y in equation (3.3)

24 class

at 1998 price

class	Income Range $	actual y	estimated y
1	Below 100	0.02294	0.02150
2	100 − 149	0.08268	0.08696
3	150 − 199	0.14057	0.15234
4	200 − 249	0.16702	0.17764
5	250 − 299	0.17031	0.17691
6	300 − 349	0.16322	0.16584
7	350 − 399	0.15175	0.15153
8	400 − 449	0.13889	0.13691
9	450 − 499	0.12507	0.12199
10	500 − 549	0.11241	0.10885
11	550 −599	0.10147	0.09783
12	600 − 649	0.09009	0.08673
13	650 − 699	0.07974	0.07674
14	700 − 749	0.07078	0.06823
15	750 −799	0.05981	0.05785
16	800 − 849	0.05255	0.05105
17	850 −899	0.04546	0.04442
18	900 −949	0.03670	0.03620
19	950 −999	0.03066	0.03051
20	1000 −1049	0.02476	0.02496
21	1050 − 1099	0.01949	0.01995
22	1100 −1149	0.01596	0.01658
23	1150 −1199	0.01000	0.01079
24	1200 and over	0.00000	0.00000

13 class

class	Income Range $	actual y	estimated y
1	Below 100	0.02294	0.02225
2	100 − 199	0.14057	0.15397
3	200 − 299	0.17031	0.17760
4	300 − 399	0.15175	0.15167
5	400 − 499	0.12507	0.12188
6	500 − 599	0.10147	0.09762
7	600 − 699	0.07974	0.07648
8	700 − 799	0.05981	0.05760
9	800 − 899	0.04546	0.04418
10	900 − 999	0.03066	0.03032
11	1000 −1100	0.01949	0.01980
12	1100−1200	0.01000	0.01069
13	1000 and over	0.00000	0.00000

Table IV

Actual and estimated frequency distribution of household's consumption expenditure

24 class

at 1998 price

class	Income Range $	Mean Income actual $	estimated $	relative frequency actual	estimated
1	Below 100	81.8	94.6	0.01302	0.01506
2	100 – 149	128.6	113.6	0.06927	0.06118
3	150 – 199	175.2	160.9	0.12999	0.11940
4	200 – 249	223.2	225.9	0.13387	0.13550
5	250 – 299	273.3	288.2	0.10396	0.10965
6	300 – 349	322.7	343.6	0.08699	0.09262
7	350 – 399	373.4	395.2	0.06893	0.07295
8	400 – 449	423.9	442.7	0.05575	0.05823
9	450 – 499	473.9	488.8	0.04919	0.05073
10	500 – 549	524.4	533.8	0.03928	0.03998
11	550 –599	573.9	575.7	0.03080	0.03090
12	600 – 649	626.5	618.3	0.02956	0.02917
13	650 – 699	673.4	659.9	0.02542	0.02491
14	700 – 749	724.0	702.1	0.02092	0.02029
15	750 –799	772.7	746.3	0.02461	0.02377
16	800 – 849	829.2	794.8	0.01565	0.01500
17	850 –899	871.7	834.3	0.01491	0.01427
18	900 –949	922.6	882.8	0.01796	0.01718
19	950 –999	975.3	934.6	0.01208	0.01158
20	1000 –1049	1030.1	986.7	0.01154	0.01106
21	1050 – 1099	1074.0	1036.2	0.01013	0.00978
22	1100 –1149	1132.5	1093.3	0.00668	0.00645
23	1150 –1199	1178.3	1153.3	0.01112	0.01088
24	1200 and over	1245.1	1320.5	0.01836	0.01947

13 class

class	Income Range $	Mean Income actual $	estimated $	relative frequency actual	estimated
1	Below 100	81.8	88.0	0.01302	0.01401
2	100 – 199	155.6	140.0	0.19926	0.17932
3	200 – 299	242.6	251.5	0.23783	0.24648
4	300 – 399	343.3	366.3	0.15592	0.16635
5	400 – 499	445.9	464.6	0.10494	0.10932
6	500 – 599	545.1	552.4	0.07008	0.07102
7	600 – 699	647.3	637.4	0.05498	0.05414
8	700 – 799	749.5	725.2	0.04553	0.04406
9	800 – 899	849.4	812.8	0.03056	0.02924
10	900 – 999	943.1	901.2	0.03004	0.02870
11	1000 –1099	1050.2	1006.3	0.02168	0.02077
12	1100–1199	1160.7	1125.4	0.01780	0.01726
13	1200 and over	1245.1	1311.1	0.01836	0.01933

Table V

Share of household consumption expenditure:

the poorest and the richest 5, 10 %

Shares of income			Estimated from 13 class	Estimated from 24 class	Actual from Individual observations
Poorest	case 1)	5%	1.44%	1.52%	1.457%
	case 2)	10%	1.88%	1.64%	3.452%
Richest	case 1)	5%	15.92%	15.97%	16.106%
	case 2)	10%	28.28%	33.15%	26.725%

* Two cases 1) and 2) are also referred in Table VI.

Table VI

Results of simulation (1):

Transferring consumption power from the richest group to the poorest group *.

Case 1) the richest(poorest) group is the highest(lowest) 5%

		Gini Inequality		
area	rate of taxation on the richest people	before	after	Improvement change %
Nationwide	1%	0.33223	0.33126	0.29%
	10%	0.33223	0.32268	2.87%

Case 2) the richest(poorest) group is the highest(lowest) 10%

Nationwide	1%	0.33223	0.32421	2.41%
	10%	0.33223	0.31019	6.63%

*Here, the richest (poorest) group means the group occupying the highest(lowest) 5% or 10% of the household consumption expenditure distribution, see more in Table V.

Table VII

Results of simulation (2):

Increasing consumption power by urban/rural residential *

area	growth rate of consumption power: by area	growth rate of consumption power: nationwide	Gini Inequality: before	Gini Inequality: after	Gini Inequality: Improvement change %
Urban	100%	45.3%	0.33223	0.45096	−35.7%
Rural	0%				
Urban	80%	63.6%	0.33223	0.36186	−8.9%
Rural	50%				
Urban	50%	66.4%	0.33223	0.30548	8.1%
Rural	80%				
Urban	0%	54.7%	0.33223	0.25157	24.3%
Rural	100%				

* In this simulation, 1,707(28.7%) households are categorized into urban group and 4,231(71.3%) households are into rural group. And consumption power is measured by household's data of consumption expenditure, here.

Table VIII

Output of simulation (3):

Double impacts both from educational promotion and from regional development*

area	growth rate of consumption power by area	growth rate of consumption power by education level of household		nationwide	Gini Inequality before	Gini Inequality after	Improvement change (1)%	Improvement change (2)% compared to Table VII
Urban	100%	Primary education NO completed	100%					
		Primary education completed	0%					
		Lower secondary education	0%	74.6%	0.33223	0.40605	-22.22%	11.05%
		Upper secondary education	0%					
Rural	0%	University/Graduated school	0%					
Urban	50%	Primary education NO completed	50%					
		Primary education completed	40%					
		Lower secondary education	30%	100.6%	0.33223	0.29756	10.44%	2.65%
		Uppers econdary education	20%					
Rural	80%	University/Graduated school	10%					
Urban	50%	Primary education NO completed	10%					
		Primary education completed	20%					
		Lower secondary education	30%	92.2%	0.33223	0.32335	2.67%	-5.54%
		Uppers econdary education	40%					
Rural	80%	University/Graduated school	50%					
Urban	0%	Primary education NO completed	0%					
		Primary education completed	0%					
		Lower secondary education	0%	61.9%	0.33223	0.26903	19.02%	-6.50%
		Upper secondary education	0%					
Rural	100%	University/Graduated school	100%					

*In this simulation, 2,191(36.9%) households are categorized into primary education NO completed, 1,293(21.7%) households are into primary education completed, 1,221(20.5%) households are into lower secondary education, 1,032(17.3%) households are into upper secondary education and 201(3.3%) households are into university/graduated school, respectively.

see also Table VII.