Inequality of Opportunity in South Korea

Woojin Lee and Jinsoon Cho

Department of Economics

Korea University

Outline

1. Introduction
2. Literature Review
3. Summary of Main Findings
4. Methods
5. Data
6. Descriptive Statistics
7. Results
8. Conclusion

1. Introduction

- Economic development should be conceived of as the degree to which an economy has implemented an efficient and just distribution of economic resources.

- The standard measure of economic development is GDP per capita, which reflects a particular conception of justice. It is a utilitarian conception of justice, where individual utility is proportional to personal income, and social welfare is the average of utilities in a population.

- A more attractive conception of justice seems opportunity equalization. Roemer (1993, 1998) show that the EOP conception of justice requires that we maximize the following social welfare function:

$$\int_0^1 Min\{v^1(\pi,\varphi),\cdots,v^t(\pi,\varphi)\}d\pi$$

where $v^t(\pi,\varphi)$ is the utility of an individual at the π th percentile of type t distribution at policy φ.

1. Introduction

- The goal of the present paper is to assess the degree to which South Korean economy is developed from an EOP perspective.

- Inequality of Opportunity

– We can always separate the inequality in a society into the part due to differential circumstances and to differential efforts: Roemer(1993, 1998). We call the inequality due to differential circumstances inequality of opportunity.
– Circumstances are those factors that are beyond individuals' control and have great influence on personal achievement.

- We decompose inequality of income, education, health, etc. in South Korea into the part due to unequal circumstances and the one due to differential exercise of individual effort.

2. Literature Review

- Roemer (1998), Kim and Lee (2008)

Compute optimal redistributive tax rates that are required to correct unequal opportunities

- Ko and Lee (2011)

Taking father's education as the only circumstance variable, they compute that father's education accounts for 25% of son's education inequality and 12% of son's income inequality.

- Lefranc et al. (2004), Kang, Ju and Oh, and Jung (2016)

They take father's education as the only circumstance variable. Instead of algebraically computing the magnitude of unequal opportunities, they test unequal opportunity using the concept of stochastic dominance.

- Erikson et al. (2015)

Using a Swedish dataset and taking 6 circumstance variables, they compute that circumstances account for 31% of male income inequality and 25% of female income inequality in Sweden.

3. Summary of Main Findings in this paper

- We use 5 circumstance variables:

father's education, gender, birth year, number of siblings, grown-up region.

- We measure individual achievement by three variables: education (schooling years), income, and health (BMI).

- We use the mean-equalized Shapley-value decomposition method:

$$Sh_j^e(I,J,G)=\sum_{\substack{S\subset J \\ j\in S}} \frac{(s-1)!(k-s)!}{k!}\left[G\left(y^e(S)\right)-G\left(y^e(S-\{j\})\right)\right].$$

- The fives circumstance variables account for 47% of education inequality, 52% of income inequality, and 38% of health inequality.

4. Methods

(1) A regression specification for education equation

$$Edu_i = \alpha_0 + \alpha_1 Fedu_i + \alpha_2 Male_i + \alpha_3 Age_i + \alpha_4 Growreg_i + \alpha_5 Nsib_i + \epsilon_i$$

We interpret the effect of age in education as a circumstance.

(2) A regression specification for income equation

$$Wage_i = \beta_0 + \beta_1 Fedu_i + \beta_2 Male_i + \beta_3 Age_i + \beta_4 Growreg_i + \beta_5 Nsib_i + \beta_6 Edu_i + \eta_i$$

As Edu in (2) is affected by circumstances (see (1)), we substitute (1) for Edu: more on this later.

(3) A regression specification for health equation

$$BMI_i = \varepsilon_0 + \varepsilon_1 Fdeu_i + \varepsilon_2 Male_i + \varepsilon_3 Age_i + \varepsilon_4 Age_i^2 + \varepsilon_5 Growreg_i + \varepsilon_6 Nsib_i + \varepsilon_7 Workout_i + \tau_i,$$

4. Methods

Comments on (2)

- When we substitute equation (1) for Edu in equation (2), we have

$$Wage_i = \gamma_0 + \gamma_1 Fedu_i + \gamma_2 Male_i + \gamma_3 Age_i + \gamma_4 Growreg_i + \gamma_5 Nsib_i + \epsilon_i$$

$$\gamma_j = \beta_j + \widehat{\alpha_j}\beta_6, \text{ j}=\{0, \cdots, 5\}$$

$$\mu_i = \eta_i + \beta_6 \epsilon_i$$

- Here, $\gamma_3 Age_i = (\beta_3 + \widehat{\alpha_3}\beta_6)\, Age_i$.

- The above regression is problematic. When analyzing the effect of birth year(age) in income, we need to distinguish the part of age as a circumstance and that as the degree of accumulated experience.

- We define $Wage_i^{Adj} = Wage_i - \widehat{\beta_3} Age_i$ and run the following regression:

$$Wage_i^{Adj} = \delta_0 + \delta_1 Fedu_i + \delta_2 Male_i + \delta_3 Age_i + \delta_4 Growreg_i + \delta_5 Nsib_i + \rho_i$$

4. Methods

- Classification of Types

Circumstance Variable	Types	Criteria	Obs.
Father's Education	Group 1	0-6 years	1689
	Group 2	7-12 years	1464
	Group 3	Over 13 years	281
Gender	Group 1	Female	1824
	Group 2	Male	1610
Birth Year	Group 1	1960-1970	1903
	Group 2	1971-1980	1531
Grown-Up Region	Group 1	Other cities	2117
	Group 2	Metropolitan cities	1317
Number of Siblings	Group 1	Less than Median	1352
	Group 2	More than Median	2081

5. Data

- Korean Labor and Income Panel Study (KLIPS)

  - A register of individuals living in cities.
  - We use personal data of individuals born between 1960-1980.

- Variables for Individual Achievement

  - Education: measured by schooling years.
  - Labor income: 5 year average of before tax labor income (2009-2013).

Nominal incomes are converted into real ones taking 2013 as the base year.

- BMI : 3 year average of BMI which is defined by $\text{BMI} = \frac{\text{mass}_{\text{kg}}}{\text{height}_{\text{m}}^2}$

5. Data

- Explanatory Variables

  - Father's education (Schooling years)
  - Gender (1=male, 0=female)
  - Birth year
  - Grown-up region before age 17 (1=metropolitan city, 0=if else)
  - Number of Siblings

    We divide individuals into two groups (1960-70/ 1971-80).

    1 if an individual has more siblings than the group median, 0 otherwise.

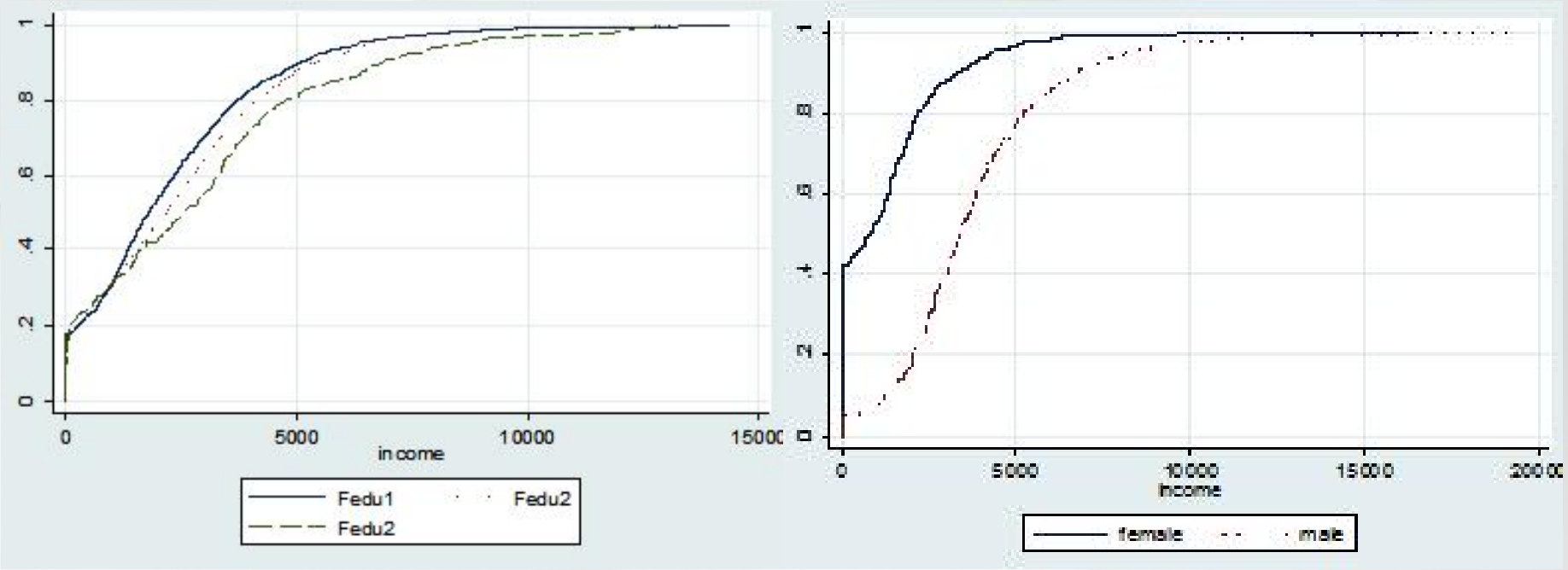
6. Descriptive Statistics

- Proportions of each category of individual education by types

	Individual's Education			
Father's Edu	0-9 years	10-12 years	Over 13 years	Total
0-6 years	0.166	0.610	0.224	1.000
7-12 years	0.042	0.436	0.522	1.000
Over 13 years	0.011	0.213	0.776	1.000
Total	0.100	0.503	0.397	1.000
Gender	0-9 years	10-12 years	Over 13 years	Total
Female	0.120	0.551	0.329	1.000
Male	0.077	0.449	0.474	1.000
Total	0.100	0.503	0.397	1.000

6. Descriptive Statistics

- Income distribution (CDF) by types



7. Result

- Regression Result

	Education	Labor Income	BMI
Fathers Education	0.1834*** (19.65)	23.22* (2.37)	−0.0191 (-1.32)
Male	0.6468*** (8.67)	1918.1*** (25.51)	1.8518*** (15.59)
Grown-up region	0.2157** (8.67)	1.0844 (0.01)	0.0926 (0.74)
Number of Siblings	−0.0592* (-2.27)	32.32 (6.27)	−0.0596 (-0.48)
Age	−0.0237** (-3.27)	45.39*** (6.27)	0.4000*** (2.63)
Age^2			−0.0043** (-2.37)
Education		269.45*** (15.63)	
Workout			−0.4834 (-0.82)
Constant	12.09*** (39.29)	−3737.1** (-10.01)	13.512*** (4.25)
Obs	3427	2851	1627

7. Result

- Decomposition result: education inequality

	GINI		CV	
	AC	RC	AC	RC
Fedu	0.030	0.309	0.052	0.279
Male	0.007	0.071	0.012	0.066
Growreg	0.002	0.024	0.004	0.021
Nsib	0.003	0.029	0.005	0.024
Age	0.004	0.039	0.006	0.032
Resid	0.051	0.529	0.108	0.578
Total	0.096	1.000	0.188	1.000

7. Result

- Decomposition result: income inequality (income>0)

	GINI		CV	
	AC	RC	AC	RC
Fedu	0.037	0.107	0.052	0.076
Male	0.138	0.396	0.262	0.384
Growreg	0.002	0.007	0.004	0.005
Nsib	0.000	0.000	0.000	0.000
Age	0.003	0.008	0.004	0.006
Resid	0.168	0.482	0.360	0.528
Total	0.348	1.000	0.682	1.000

7. Result

- Decomposition result: income inequality (total sample)

	GINI		CV	
	AC	RC	AC	RC
Fedu	0.036	0.083	0.048	0.057
Male	0.212	0.486	0.407	0.484
Growreg	0.001	0.001	0.001	0.001
Nsib	0.001	0.001	0.001	0.001
Age	0.003	0.007	0.004	0.005
Resid	0.183	0.421	0.38	0.452
Total	0.435	1.000	0.841	1.000

7. Result

- Decomposition result: income inequality (male, income>0)

	GINI		CV	
	AC	RC	AC	RC
Fedu	0.040	0.150	0.071	0.133
Growreg	0.001	0.005	0.002	0.004
Nsib	0.002	0.006	0.003	0.005
Age	0.001	0.004	0.001	0.003
Resid	0.224	0.836	0.457	0.855
Total	0.268	1.000	0.535	1.000

7. Result

- Decomposition result: health inequality

	GINI		CV	
	AC	RC	AC	RC
Fedu	0.001	0.013	0.001	0.010
Male	0.014	0.221	0.026	0.228
Growreg	0	0.005	0.001	0.004
Nsib	0	0.003	0	0.003
Age	0.007	0.141	0.015	0.133
Workout	0.002	0.001	0	0.003
Resid	0.039	0.615	0.07	0.618
Total	0.063	1.000	0.113	1.000

8. Conclusion

- In all three cases, the circumstance is found to account for 46% of the inequality.

- The father’s education is the most important contributor to education inequality, and gender is the most important contributor to income and health inequality.

- In education inequality, father's education contributes the most (about 31% of inequality) and the gender(7%), birth year(4%), the number of siblings(3%), and grown-up region(2%).

- In income inequality, gender contributes the most (40%) and father’s education (11%).

- In health inequality, gender contributes the most (23%) and age (13%).

- We interpret residuals as an individual effort. The residual’s variance, however, varies across types (type-specific residuals), which individuals are not held accountable for.