# Income Mobility and Its Implication on Government Welfare Expenditure

**Myung Jae Sung***

**Abstract:** This paper analyzes income mobility between two consecutive periods and estimates the potential effects of introducing differentiated poverty treatment on direct government welfare expenditure. This is done by using a constructed pseudo-panel under the pre-existing two conditions: log-normality of household income and the stability of the household income distribution over time. This paper finds that income mobility shrank from the late 1990s onward. As a result, the probability that a poor household escapes from poverty dwindles. This is partly because the labor market becomes slightly more rigid and also partly because the population is rapidly aging: the share of the elderly who are mostly retirees is growing rapidly. A pseudo-panel study shows that total subsidies to support all poor households amount to 7.0 trillion won a year. It also shows that the lifelong poverty rate is 3.4%, approximately one third of the short-run poverty rate of 10.9%. A differential treatment on short-run and long-run poverty can save roughly half the fiscal burden of supporting the poor.

**Keywords:** Income Mobility, Poverty, Welfare Expenditure, Pseudo-Panel, Gini

JEL Classification: H31, H53, I32

## INTRODUCTION

This paper analyzes income mobility in two consecutive periods and estimates the potential effects of introducing differentiated poverty treatment on direct welfare expenditure by the government. Since the Korean economic crisis, income inequality has been increasing and the poverty problem has been getting worse. Together with a rapidly aging population, poverty has become one of the most influential factors in the country's rapid fiscal expansion. This paper aims to provide insights toward finding a parsimonious way to alleviate the fiscal burden through unequal treatment of unequal

* Myung Jae Sung(Ph. D. in economics, University of Wisconsin-Madison) is a Senior Fellow at the Korea Institute of Public Finance. E-mail: sung@kipf.re.kr.

Manuscript received July 2008; out for review July 2008; review completed August 2008; accepted Augutst 2008.

*The Korean Journal of Policy Studies*, Vol. 23, No. 1, 127-151 (2008)

poverty states.

The first part of the paper is devoted to analyzing income mobility by use of the income transition rule, which is extended to construct a pseudo-panel to estimate welfare expenditure. In the second part, which scrutinizes poverty issues, the amount of pecuniary welfare expenditure to support the poor is estimated. The effects of differential treatment based on the duration of poverty are also analyzed. Estimates of poverty rates and redistributive effects of government interventions such as taxes and fiscal expenditures are presented in the Appendix.

## INCOME MOBILITY AND A PSEUDO-PANEL

In this paper, income mobility is measured by changes in relative income ranks between two consecutive periods. The model analyzes household income using log-normal distribution (Kim & Sung, 2003, 2004) to construct an income transition rule, which is extended to a pseudo-panel. These are illustrated below.

### Literature Survey: Pseudo-Panels

Panel data sets usually provide important information that enables us to trace longitudinal changes in decision-making processes and their outcomes. Long-run analyses provide findings that are difficult to obtain in short-run analyses using cross-section data sets.

Even with panel data sets, we often face a number of difficulties in analyzing lifetime income analyses. The demographic composition of a household changes over time, and data attrition increases with time. Even if the original set of data contains a large number of observations, only a few of these observations will remain at the end of the study period. Unless the attrition occurs randomly, any inference based on the remaining sample generally leads to biased inference due to the sample selection. Therefore, an alternative statistical method needs to be posed to resolve this problem. Sometimes, a real panel data set does not deliver necessary information in analyzing life-long income analysis. In Korea, a panel known as the Korean Labor and Income Panel Study (KLIPS hereafter) is widely used in empirical studies. However, it has not been compiled for long enough to infer characteristics of lifelong income distribution. Even if it had been compiled for a sufficiently long time, it may not be suitable because of severe data attrition. In these cases, a pseudo-panel approach might be more suitable. This paper investigates the suitability of a pseudo-panel by use of statistical properties on income distribution.

There have been a handful of studies on pseudo-panels; however, none of them is micro-based. Gassner (1998) estimated the telephone access demand in the UK using independent cross-sections. He constructed a set of time-series variables that are averaged over each cross-section and named it a pseudo-panel. McKenzie (2004) studied asymptotic properties of a heterogeneous pseudo-panel that was constructed in a similar way. Many other studies have used a pseudo-panel constructed from independent or repeated cross-sections (e.g., Alessie, Devereux, & Weber, 1997; Girma, 2000; Pelzer, Eisinga, & Franses, 2001; Verbeek & Nijman, 1993; Verbeek & Vella, 2005).

The pseudo-panels adopted in the papers mentioned above are not micro-based panels but collections of aggregate time-series variables. In this case, it is almost impossible to infer relative income inequality with a pseudo-panel of aggregate variables, such as changes in Gini, poverty rates, poverty gap, and relative quintile income ratios. Therefore, a micro-based pseudo-panel is recommended, especially for longitudinal studies on individual distribution. In order to generate a pseudo-panel, the probability distribution structure of income, for instance, needs to be stable in the sense that it remains in the same probability distribution family over time.

An approach using parametric information on distribution function of a variable of interest can simplify the pseudo-panel generating process. For example, household income is distributed log-normal. A number of studies have found important and useful statistical properties on income distribution that give a wide range of applications in economic analyses. Formal studies on the log-normality of income distribution have long been rooted in Western countries. Kalecki (1945) found that the log-income was distributed normal in the UK. Aitchison and Brown (1957) found that income distribution followed a log-normal distribution. Recently, the log-normality of income distribution has been tested. Kim and Sung (2003, 2004) showed that annual household income in Korea has been distributed log-normal in most years during the past two decades and that the deviation from log-normal distribution is more or less negligible in other years. They also found that the log-normality of household income distribution holds in each population subgroup in terms of ages of household heads and in terms of household size[1] as well. They showed that the structure of income distribution has been stable over time within the family of log-normal distributions during the past two decades, except for the Korean economic crisis period 1997-1998.

1. Based on National Survey of Household Income and Expenditures, a household whose size is two or larger is tested to be distributed log-normal in terms of its income. Single households are also distributed log-normal, to the extent that a partition based on age is concerned: one is for the households whose heads are 35 years old or younger and the other is for those over 35.

## Model Identification

Consider a partition of population that consists of 41 age groups from 25 to 65: age group 25 includes 25 years of age and under, and age group 65 includes 65 and over. Suppose that the income rank of a household in age group A (A = 25, . . . ,65) in the current period, t, is the p-th percentile within its own age group. Assume that the household income is realized around the p-th percentile income with a certain probability distribution in age group (A+1) at time (t+1). The rationale for this assumption is based on the presumption that income flow is generally persistent over time. It will turn out later that this presumption is examined and supported by an empirical hypothesis test.

Let $Y_{tA} \in R^{l}$ be a random variable that denotes household income in age group A at t. Also, let $X_{tA}^{p} \in R^{l}$ be the realization of $Y_{tA}$ at the p-th percentile within age group A at t. In this case, $X_{(t+1),(A+1)}^{p'}$ denotes its realized income with the rank p' in the age group (A+1) at (t+1). Suppose that is the p-th percentile income level in the age group (A+1) at time (t+1)[2]. The income transition rule between the two consecutive years is defined as follows:

$$X_{(t+1),(A+1)}^{p'} = (1+\gamma_{(t+1),(A+1)}^{p}) \cdot X^{*p}_{(t+1),(A+1)} \text{ for some } \gamma \in R^{l},\ \gamma > -1 \tag{1}$$

Taking logarithm over equation (1) yields the following equation:

$$Z_{(t+1),(A+1)}^{p'} = \Gamma_{(t+1),(A+1)}^{p} + Z^{*p}_{(t+1),(A+1)} \tag{2}$$

where $Z = \ln (X_{(t+1),(A+1)}^{p'})$, $Z^{*} = \ln (X^{*p}_{(t+1),(A+1)})$ and $\Gamma = (1+\gamma_{(t+1),(A+1)}^{p})$.

We may call equation (2) an income transition rule. By the same token, Γ is the income transition variable. For simplicity of discussion, we may drop superscripts and subscripts that denote time (t), age (A), and income percentile (p), and regard Z, Z*, and Γ as random variables. Note that the distributions of Z, Z*, and Γ depend on t, A, and p. Thus, we will retrieve subscripts and superscripts when needed.

Assume that the log-income of each age group, Z, is distributed normal. Since Z is Z* that is reshuffled through Γ, note that both Z and Z* are normally distributed with identical means and variances. Under the normality of Z*, the normality of Γ becomes both sufficient and necessary conditions for that of Z. Consider the following expectation and variance.

2. Note that age group (A+1) at (t+1) is identical to age group A at t.

$$E(Z) = E(\Gamma) + E(Z^*) \tag{3}$$

$$Var(Z) = Var(\Gamma) + Var(Z^*) + 2Cov(\Gamma, Z^*) \tag{4}$$

Note that Z and Z* have identical distributions. Therefore, E(Γ) = 0 or, equivalently, E(γ·X*) = 0. This implies that γ is orthogonal to X*. This does not necessarily mean either E(γ | X*) = 0 or E(γ) = 0. The expectation of γ should not be zero. Otherwise, equation (3) does not hold. Since E(Z)=E(Z*) and Var(Z)=Var(Z*),

$$\mu_\Gamma = 0 \tag{5}$$

$$\sigma_{\Gamma Z^*} = -\frac{\sigma_\Gamma^2}{2} \tag{6}$$

where $\mu_\Gamma = E(\Gamma)$ and $\sigma_{\Gamma Z^*} = Cov(\Gamma, Z^*)$.

The covariance of Γ and Z* depends only on the variance of Γ, $\sigma_\Gamma^2$. Since Var(Z) = Var(Z*), $\sigma_\Gamma^2$ in the right hand side of equation (4) is always and exactly annihilated by the covariance of Γ and Z* multiplied by two. Otherwise, the resulting distribution of Z in the left hand side of equation (4) will either explode or be degenerate, depending on the relative size of covariance and variance over time. So, Z can be a legitimate normal variable through equation (6). In this respect, we may call equation (6) the stability condition.

Estimation of $\sigma_\Gamma^2$ requires information on income transition patterns between two consecutive years. KLIPS contains necessary information to estimate $\sigma_\Gamma^2$. It is estimated in three steps. In the first step, relative income ranks or cumulative relative frequencies are calculated for each household. In the second step, the relative income ratios of realized income levels to the imaginary income levels corresponding to the ranks in the previous year are calculated household by household. In the final step, $\sigma_\Gamma^2$ is estimated based on taking natural logarithm on these ratios.

There are two ways of estimating $\sigma_\Gamma^2$ under the normality of Γ. We may directly calculate the variance of Γ using the following formula (a direct method):

$$\hat{\sigma}_\Gamma^2 = \sum_{i,t} w_i \Gamma_i^2 \tag{7}$$

Since $\mu_\Gamma = E(\Gamma) = 0$, $\bar{\Gamma}_i$ is dropped from the right hand side of equation (7) by the analogy principle.

The second way of estimating $\sigma_\Gamma^2$ is to find a value of $\sigma_\Gamma^2$ which minimizes the mean squared error of the estimated densities of Γ within the distributional family of normal distributions with zero means. This is often called an indirect method.

It is not unusual to observe extremely high values of $\Gamma$ in KLIPS especially in very low income groups. Therefore, indirect methods are preferred, as they minimize distortive effects caused by the outliers[3]. In what follows, we will use indirect methods only. Equations (2) and (7) are used to measure income mobility[4].

## Estimation of Income Mobility: Estimation of Variance of $\Gamma$

Before estimating the variance of $\Gamma_{t,A}^{p}$, $\Gamma_{t,A}^{p}$ needs to be tested whether to have zero mean. This is to empirically check whether the income transition rule proposed in equation (2) is valid and therefore well specified.

### Hypothesis Test for Zero Mean of Income Transition Variable $\Gamma$

Table 1 shows that absolute values of $\bar{\Gamma}_{t,A}^{p}$ are less than one without any exception for all time (t), income percentiles ($p \in [0,1]$), and ages (A=25,...,65)[5]. Likewise, as shown in Table 2, absolute t-values of $\bar{\Gamma}_{t,A}^{p}$ are much less than 1. We cannot reject the null hypothesis of zero mean of $\Gamma_{t,A}^{p}$. We can conclude that $\Gamma$ has zero mean for all t, p and A. Therefore, we can conclude that equation (2) is empirically supported.

### Income Mobility: Variance of $\Gamma$

As discussed earlier, income mobility can be measured by the variance of $\Gamma(\sigma_{\Gamma}^{2})$ under the income transition rule, which allows shifts in income ranks with a probability distribution under the log-normality of income. Higher values of $\sigma_{\Gamma}^{2}$ yield more variability in income shifting in the next period, and, therefore, income mobility is positively correlated with $\sigma_{\Gamma}^{2}$.

Table 3 shows the estimates of $\sigma_{\Gamma}^{2}$ for several income percentiles and ages for the years 2005-2006. $\sigma_{\Gamma}^{2}$ varies with income percentiles and ages. Generally speaking, $\sigma_{\Gamma}^{2}$

3. $\Gamma$ is distributed normal with mean 0 as shown in equation (5). This does not necessarily mean that the sample average, $\bar{\Gamma}$, is zero. Nevertheless, the probability that $\bar{\Gamma}$ is zero is zero in the continuous probability distributions. However, according to the Kolmogorov law of large numbers, $\bar{\Gamma}$ converges to $\Gamma$ as the sample size increases to infinity.
4. There are two kinds of income changes over time: income mobility and income risk. Income mobility is the income shift caused by the structural change. Income risk is the temporary income shift without causing structural change.
5. The hypothesis is tested using Chi-square test statistics. Since $\bar{\Gamma}_{t,A}^{p}$ is distributed normal for all t, p and A, the sum of squared $\bar{\Gamma}_{t,A}^{p}$ over the whole sample yields Chi-square statistic with the degree of freedom N, the sample size. In fact, the p-value of this test statistic is infinitesimal. Therefore, the null hypothesis is not rejected.

**Table 1.** Means of Income Transition Variable Γ (Based on 2005-2006 KLIPS)

| Percentile\Age | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 |
|---|---|---|---|---|---|---|---|---|---|
| 0.05 | -0.9157 | -1.0462 | -0.9083 | -0.8672 | -0.9118 | -0.8685 | -0.8187 | -0.8988 | -0.9595 |
| 0.10 | -0.9130 | -1.0430 | -0.9047 | -0.8640 | -0.9087 | -0.8643 | -0.8150 | -0.8961 | -0.9583 |
| 0.15 | -0.9103 | -1.0399 | -0.9010 | -0.8608 | -0.9055 | -0.8602 | -0.8112 | -0.8933 | -0.9572 |
| 0.20 | -0.9077 | -1.0367 | -0.8974 | -0.8576 | -0.9023 | -0.8561 | -0.8074 | -0.8906 | -0.9560 |
| 0.25 | -0.9050 | -1.0335 | -0.8937 | -0.8543 | -0.8991 | -0.8519 | -0.8037 | -0.8878 | -0.9549 |
| 0.30 | -0.9024 | -1.0303 | -0.8900 | -0.8511 | -0.8959 | -0.8478 | -0.7999 | -0.8851 | -0.9537 |
| 0.35 | -0.8998 | -1.0271 | -0.8863 | -0.8478 | -0.8927 | -0.8436 | -0.7961 | -0.8823 | -0.9525 |
| 0.40 | -0.8972 | -1.0239 | -0.8826 | -0.8446 | -0.8895 | -0.8394 | -0.7923 | -0.8796 | -0.9514 |
| 0.45 | -0.8945 | -1.0207 | -0.8789 | -0.8413 | -0.8863 | -0.8352 | -0.7885 | -0.8768 | -0.9502 |
| 0.50 | -0.8919 | -1.0175 | -0.8751 | -0.8381 | -0.8831 | -0.8310 | -0.7847 | -0.8741 | -0.9490 |
| 0.55 | -0.8894 | -1.0143 | -0.8714 | -0.8348 | -0.8798 | -0.8268 | -0.7810 | -0.8713 | -0.9479 |
| 0.60 | -0.8868 | -1.0111 | -0.8677 | -0.8315 | -0.8766 | -0.8226 | -0.7772 | -0.8685 | -0.9467 |
| 0.65 | -0.8842 | -1.0079 | -0.8639 | -0.8282 | -0.8733 | -0.8183 | -0.7734 | -0.8658 | -0.9455 |
| 0.70 | -0.8816 | -1.0047 | -0.8602 | -0.8250 | -0.8701 | -0.8141 | -0.7696 | -0.8630 | -0.9444 |
| 0.75 | -0.8791 | -1.0015 | -0.8565 | -0.8217 | -0.8668 | -0.8098 | -0.7658 | -0.8602 | -0.9432 |
| 0.80 | -0.8766 | -0.9984 | -0.8527 | -0.8184 | -0.8635 | -0.8056 | -0.7620 | -0.8575 | -0.9420 |
| 0.85 | -0.8740 | -0.9952 | -0.8489 | -0.8151 | -0.8602 | -0.8013 | -0.7583 | -0.8547 | -0.9408 |
| 0.90 | -0.8715 | -0.9920 | -0.8452 | -0.8119 | -0.8569 | -0.7971 | -0.7545 | -0.8519 | -0.9396 |
| 0.95 | -0.8690 | -0.9888 | -0.8414 | -0.8086 | -0.8537 | -0.7928 | -0.7507 | -0.8492 | -0.9385 |

**Table 2.** t-Values of Means of Γ (Based on 2005-2006 KLIPS)

| Percentile\Age | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 |
|---|---|---|---|---|---|---|---|---|---|
| 0.05 | -0.907 | -0.945 | -0.900 | -0.899 | -0.911 | -0.863 | -0.875 | -0.897 | -0.904 |
| 0.10 | -0.906 | -0.945 | -0.899 | -0.898 | -0.910 | -0.861 | -0.873 | -0.896 | -0.904 |
| 0.15 | -0.906 | -0.944 | -0.898 | -0.898 | -0.909 | -0.859 | -0.872 | -0.895 | -0.904 |
| 0.20 | -0.905 | -0.944 | -0.897 | -0.897 | -0.908 | -0.857 | -0.871 | -0.894 | -0.903 |
| 0.25 | -0.904 | -0.943 | -0.896 | -0.896 | -0.907 | -0.855 | -0.869 | -0.893 | -0.903 |
| 0.30 | -0.904 | -0.942 | -0.895 | -0.895 | -0.906 | -0.854 | -0.868 | -0.892 | -0.903 |
| 0.35 | -0.903 | -0.942 | -0.894 | -0.894 | -0.905 | -0.852 | -0.867 | -0.891 | -0.902 |
| 0.40 | -0.902 | -0.941 | -0.893 | -0.893 | -0.904 | -0.850 | -0.865 | -0.890 | -0.902 |
| 0.45 | -0.902 | -0.941 | -0.892 | -0.892 | -0.903 | -0.848 | -0.864 | -0.889 | -0.902 |
| 0.50 | -0.901 | -0.940 | -0.891 | -0.891 | -0.902 | -0.846 | -0.863 | -0.888 | -0.902 |
| 0.55 | -0.900 | -0.939 | -0.890 | -0.890 | -0.901 | -0.844 | -0.861 | -0.887 | -0.901 |
| 0.60 | -0.900 | -0.939 | -0.889 | -0.889 | -0.900 | -0.843 | -0.860 | -0.886 | -0.901 |
| 0.65 | -0.899 | -0.938 | -0.888 | -0.888 | -0.899 | -0.841 | -0.859 | -0.885 | -0.901 |
| 0.70 | -0.899 | -0.938 | -0.886 | -0.887 | -0.898 | -0.839 | -0.857 | -0.884 | -0.900 |
| 0.75 | -0.898 | -0.937 | -0.885 | -0.886 | -0.897 | -0.837 | -0.856 | -0.883 | -0.900 |
| 0.80 | -0.897 | -0.937 | -0.884 | -0.885 | -0.896 | -0.835 | -0.854 | -0.882 | -0.900 |
| 0.85 | -0.897 | -0.936 | -0.883 | -0.884 | -0.895 | -0.833 | -0.853 | -0.881 | -0.899 |
| 0.90 | -0.896 | -0.935 | -0.882 | -0.883 | -0.894 | -0.831 | -0.852 | -0.880 | -0.899 |
| 0.95 | -0.895 | -0.935 | -0.881 | -0.882 | -0.893 | -0.829 | -0.850 | -0.879 | -0.899 |

Note: Shaded area denotes that the null hypothesis of zero mean of Γ is not rejected.

seems to be correlated negatively with income percentiles; higher income percentiles tend to have smaller values of $\sigma_\Gamma^2$. The correlation between $\sigma_\Gamma^2$ and ages does not seem clear enough to show a linear relationship. It seems rather U-shaped: young and old generations tend to have large values of $\sigma_\Gamma^2$ and, however, middle-aged generations tend to have relatively smaller values of $\sigma_\Gamma^2$.

**Table 3.** Variances of Γ (Based on 2005-2006 KLIPS)

| Percentile\Age | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 |
|---|---|---|---|---|---|---|---|---|---|
| 0.05 | 1.01901 | 1.22495 | 1.01890 | 0.92971 | 1.00140 | 1.01378 | 0.87622 | 1.00379 | 1.12594 |
| 0.10 | 1.01457 | 1.21903 | 1.01293 | 0.92477 | 0.99659 | 1.00828 | 0.87072 | 0.99981 | 1.12398 |
| 0.15 | 1.01015 | 1.21311 | 1.00696 | 0.91983 | 0.99178 | 1.00279 | 0.86524 | 0.99583 | 1.12202 |
| 0.20 | 1.00575 | 1.20718 | 1.00099 | 0.91488 | 0.98695 | 0.99728 | 0.85977 | 0.99185 | 1.12005 |
| 0.25 | 1.00137 | 1.20126 | 0.99502 | 0.90993 | 0.98210 | 0.99177 | 0.85431 | 0.98788 | 1.11808 |
| 0.30 | 0.99700 | 1.19534 | 0.98906 | 0.90499 | 0.97726 | 0.98625 | 0.84887 | 0.98391 | 1.11610 |
| 0.35 | 0.99265 | 1.18942 | 0.98310 | 0.90004 | 0.97240 | 0.98073 | 0.84344 | 0.97995 | 1.11412 |
| 0.40 | 0.98831 | 1.18350 | 0.97715 | 0.89510 | 0.96753 | 0.97521 | 0.83803 | 0.97599 | 1.11213 |
| 0.45 | 0.98400 | 1.17759 | 0.97120 | 0.89017 | 0.96266 | 0.96969 | 0.83263 | 0.97203 | 1.11013 |
| 0.50 | 0.97971 | 1.17168 | 0.96527 | 0.88523 | 0.95778 | 0.96416 | 0.82725 | 0.96808 | 1.10814 |
| 0.55 | 0.97543 | 1.16578 | 0.95934 | 0.88031 | 0.95290 | 0.95863 | 0.82189 | 0.96414 | 1.10614 |
| 0.60 | 0.97118 | 1.15989 | 0.95342 | 0.87538 | 0.94801 | 0.95311 | 0.81655 | 0.96021 | 1.10413 |
| 0.65 | 0.96694 | 1.15400 | 0.94751 | 0.87047 | 0.94312 | 0.94758 | 0.81123 | 0.95628 | 1.10212 |
| 0.70 | 0.96273 | 1.14813 | 0.94161 | 0.86557 | 0.93822 | 0.94206 | 0.80594 | 0.95236 | 1.10011 |
| 0.75 | 0.95853 | 1.14227 | 0.93572 | 0.86067 | 0.93333 | 0.93654 | 0.80066 | 0.94844 | 1.09809 |
| 0.80 | 0.95436 | 1.13641 | 0.92985 | 0.85578 | 0.92843 | 0.93102 | 0.79540 | 0.94454 | 1.09607 |
| 0.85 | 0.95021 | 1.13058 | 0.92399 | 0.85091 | 0.92353 | 0.92551 | 0.79017 | 0.94064 | 1.09405 |
| 0.90 | 0.94608 | 1.12475 | 0.91815 | 0.84604 | 0.91864 | 0.92000 | 0.78496 | 0.93676 | 1.09202 |
| 0.95 | 0.94197 | 1.11894 | 0.91232 | 0.84119 | 0.91374 | 0.91450 | 0.77977 | 0.93288 | 1.08999 |

We estimate $\sigma_\Gamma^2$ using KLIPS. It is not tractable to show the structural changes in $\sigma_\Gamma^2$ over time where $\sigma_\Gamma^2$ also varies with income percentiles and ages, although it is not impossible. An alternative way follows. Without loss of generality, for simplicity of discussion, consider the case where $\sigma_\Gamma^2$ is independent of ages and income percentiles. This implies that $\sigma_\Gamma^2$ can vary only with time. Based on the KLIPS data sets, income mobility measured by $\sigma_\Gamma^2$ seems to have dwindled since the economic crisis. This can be empirically tested using the following test statistics. Note that Γ follows a normal distribution. Thus, it is well known that

$$\frac{(n-1)s_\Gamma^2}{\sigma_\Gamma^2} \sim x^2(n-1) \tag{8}$$

From equation (8), by the central limit theorem, we can have the following asymptotic normal distribution of sample variance estimator of Γ.

$$s_{\Gamma}^{2} \sim_{A} N\left(\sigma_{\Gamma}^{2}, \frac{2\sigma_{\Gamma}^{4}}{n-1}\right) \tag{9}$$

From equation (9), the difference in $\sigma_{\Gamma}^{2}$ for different time t≠s asymptotically follows the following normal distribution under the null hypothesis of identical variances, i.e. $\sigma_{\Gamma t}^{2} = \sigma_{\Gamma s}^{2}$,

$$s_{\Gamma t}^{2} - s_{\Gamma s}^{2} \sim_{A} N\left(0, \frac{2\sigma_{\Gamma t}^{4}}{N_t-1} + \frac{2\sigma_{\Gamma s}^{4}}{N_s-1}\right) \text{ for } t\neq s \tag{10}$$

where $N_t$ and $N_S$ are sample sizes at t and s (t≠s), respectively. Using equation (10), we can test whether there has been a structural change in $\sigma_{\Gamma}^{2}$. Table 4 shows that since 1998 there have been three structural breaks between two adjacent years. We can conclude that the income mobility measured by $\sigma_{\Gamma}^{2}$ has dwindled since the economic crisis: $\sigma_{\Gamma}^{2}$ was 0.25646 in 1998-1999 and substantially decreased to 0.19492 in 2005-2006.

**Table 4.** Statistics Regarding Γ when Γ is distributed independently of income percentiles and Ages (Based on KLIPS)

| | 1998-1999 | 1999-2000 | 2000-2001 | 2001-2002 | 2002-2003 | 2003-2004 | 2004-2005 | 2005-2006 |
|---|---|---|---|---|---|---|---|---|
| Sample Size | 4,052 | 3,714 | 3,416 | 3,411 | 3,389 | 3,506 | 3,577 | 3,601 |
| $\sigma_{\Gamma}^{2}$ | 0.25646 | 0.23788 | 0.25123 | 0.24222 | 0.23769 | 0.23905 | 0.22271 | 0.19492 |
| t-value of Test Statistic | – | -2.342 | 1.626 | -1.066 | -0.550 | 0.167 | -2.103 | -3.976 |
| $(\bar{\Gamma})$ | 0.09719 | 0.07394 | 0.07913 | -0.04031 | -0.02574 | -0.04518 | -0.06393 | -0.03584 |

Note: Shaded area denotes that the hull hypotheses that income mobility does not change are rejected at 5% significance level.

## Extension to Construct a Pseudo-Panel

Under the normality of Z (or Z*), that is the log-normality of X, the income transition rule reduces to the variance of Γ. This also means that the dimension of parameter space reduces to one. Under the log-normality of household income distribution, estimation of $\sigma_{\Gamma}^{2}$ suffices to generate a pseudo-panel data set together with estimation of means and variances of each age groups and changes in relative shares among age groups.

Note that Γ and Z* are correlated normal variables. A simple example is a bivariate normal distribution. This relationship satisfying equation (2) is stronger than necessary and, however, simple to apply. Therefore, we adopt the bivariate normality between Γ and Z* for simplicity of discussion.

To construct a pseudo-panel we need to generate future values for Z for each period. Future values for Z can be obtained as a sum of Γ and Z*, according to equation (2). Therefore, the future values for the pairs (Γ,Z*) are obtainable by generating random numbers for the joint bivariate normal variables. The values of Z* for an age group is determined by the income ranks in the previous period and also by the income distribution in the current period. Since Γ and Z* are negatively correlated, independent draws will not suffice. Given each values of Z*, matching values of Γ are generated using the conditional probability distribution of Γ given Z*. Since (Γ,Z*) are jointly bivariate normal, BVN(0, $\mu_Z$, $\sigma_\Gamma^2$, $\sigma_\Gamma^2$, $\rho$) where $\rho = \frac{\sigma_{\Gamma Z^*}}{\sigma_\Gamma \cdot \sigma_{Z^*}}$ [6], the distribution of Γ conditional on Z* follows a normal distribution whose mean and variance are varying with Z*. Note that the expectation of a variate conditional on the other variate from bivariate normal distribution is linear in the conditioning variate. So, the conditional mean of Γ given Z* is linear in Z* such that $E(\Gamma \mid Z^*) = \mu_\Gamma + \frac{\sigma_{\Gamma Z^*}}{\sigma_{Z^*}^2}(Z^* - \mu_{Z^*})$ and its variance is $\sigma_\Gamma^2(1-\rho^2)$. Since $\mu_\Gamma$=0,

$$\Gamma \mid Z^* \sim N\left(\frac{\sigma_{\Gamma Z^*}}{\sigma_{Z^*}^2}(Z^* - \mu_{Z^*}), \sigma_\Gamma^2(1-\rho^2)\right) = N\left(\frac{\sigma_{\Gamma Z}}{\sigma_Z^2}(Z - \mu_Z), \sigma_\Gamma^2(1-\rho^2)\right) \tag{11}$$

## EFFECTS OF INCOME MOBILITY ON WELFARE EXPENDITURE

In this section, we focus on the poverty[7] dynamics of income mobility and necessary fiscal expenditure for poor households. Using an income transition rule and a

---

6. Note that $\sigma_{Z^*}^2 = \sigma_{Z^*}^2$ so, $\rho = -\frac{\sigma_\Gamma}{2\sigma_{Z^*}} = -\frac{\sigma_\Gamma}{2\sigma_Z}$
7. There have been a handful of studies on poverty issues in Korea. Park and Kim (2000), Lee and Lee (2000), Lim (2006), and Cho and Kim (2007) analyzed changes in poverty rates and poverty persistence. Keum (2006) raised issues of working poor and estimated its determinants. Park and Kim (1997) and Sim (2006) scrutinized the characteristics and determinants of poverty. Hwang (2001) estimated the poverty outflow rates. Seok (2007) studied poverty dynamics. Sung (2007) tested whether the poverty outflow rate changed.

pseudo-panel, both of which hold under the log-normality of income, we test the changes in income mobility. We estimate necessary welfare expenditure to support the poor. In the meantime, we revisit the income transition rule and then estimate the scale of government welfare expenditure required to support the poor.

## Poverty Outflow/Inflow Probabilities

Under the income transition rule and the log-normality of income, we can calculate the conditional probability of longitudinal income shifting. Poverty inflow and outflow probabilities are good examples.

The poverty rate is defined as the proportion of households whose income is below the poverty line (PL). The poverty line is usually predetermined either at the minimum subsistence level (absolute poverty) or at the equivalized median income multiplied by a constant between zero and one[8] (relative poverty). Poverty outflow probability (POP) at a certain period of time (t) is the probability that a household earns income no less than the poverty line in the following period (t+1). POP varies with the income level or income rank at time t, since the distribution of income shifting factor, $\Gamma$, at (t+1) depends upon the income percentile at t. Similarly, poverty inflow probability (PIP) at a given time (t) is defined as the probability that a household income is below the poverty line in the following period (t+1).

Note that $\Gamma$ and $Z^*$ for all t jointly follow a bivariate normal distribution and their sum, Z, follows a normal distribution. For a household in which the head is (A+1) years old at t, the household is classified into the poor group if its income (Z) is realized below the poverty line. Note that $Z^*$ has the same distribution as Z. More specifically, $Z^*$ is an imaginary or unobserved random variable that could be generated within the age group (A+1) at t at the same income percentile as in age group A at (t-1) when the household head was A years old. However, the realized income at t is not $Z^*$ but Z, which becomes different from $Z^*$ due to the income transition variable, $\Gamma$.

Therefore, we can calculate the probability of income at t being realized below or

Lee (2004, 2006) dealt with poverty issues and related policy suggestions. An and Song (2006) and Lee (2007) discussed the validity and feasible way of introducing EITC. Kim and Kwon (2007) studied the effects of National Pension system. Yoo and Kim (2002) analyzed poverty issues based on the permanent income concept. An and Song (2006) also studied the relationship between business cycles and income distribution.

Rodgers and Rodgers (1993) scrutinized long-run poverty. Bane and Ellwood (1986) studied poverty dynamics. Gardiner and Hills (1999), and Jarvis and Jenkins (1998) analyzed income mobility. Besley (1990) studied effects of policy alleviation programs.

8. Typically, it is set at 0.5.

above the poverty line for a household of which the income percentile was p in the previous period (t-1), using the bivariate normality of ($\Gamma$,Z*). This probability depends on the distribution of $\Gamma$ conditioned on Z*.

Note that ($\Gamma$, Z*) are jointly bivariate normal, BVN($\mu_\Gamma$, $\mu_Z$, $\sigma_\Gamma^2$, $\sigma_Z^2$, $p$) where $\mu_\Gamma$ = 0 and $\rho = \frac{\sigma_{\Gamma Z}}{\sigma_\Gamma \sigma_Z} = -\frac{\sigma_\Gamma}{2\sigma_Z}$. The joint and marginal densities of ($\Gamma$,Z*) and Z* are as follows:

$$f(\Gamma, Z^*) = \frac{1}{2\pi\sigma_\Gamma\sigma_{Z^*}\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{\Gamma^2}{\sigma_\Gamma^2} - \frac{2\rho\Gamma(Z^*-\mu_{Z^*})}{\sigma_\Gamma\sigma_{Z^*}} + \frac{(Z^*-\mu_{Z^*})^2}{\sigma_{Z^*}^2}\right)\right] \quad (12)$$

$$f_Z(Z^*) = \frac{1}{\sqrt{2\pi}\,\sigma_{Z^*}} \exp\left[-\frac{(Z^*-\mu_{Z^*})^2}{2\sigma_{Z^*}^2}\right] \quad (13)$$

The conditional density of $\Gamma$ given Z* is $f(\Gamma \mid Z^*) = \frac{f(\Gamma, Z^*)}{f_Z(Z^*)}$:

$$f(\Gamma \mid Z^*) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\Gamma^2(1-\rho^2)}} \exp\left[-\frac{1}{2\sigma_\Gamma^2(1-\rho^2)}\left(\Gamma - \frac{\sigma_{\Gamma Z^*}}{\sigma_{Z^*}^2}(Z^*-\mu_{Z^*})\right)^2\right] \quad (14)$$

Suppose that the income percentile of a household was p in age group A at (t-1). Denote the p-th income percentile $z_0$ in age group (A+1) at t. Suppose that PL is given at $P_0$. The poverty outflow probability is the probability that the realized log-income is larger than the log of the poverty line, given Z* = $z_0$:

$$\Pr[Z \geq \ln(P_0) \mid Z^* = z_0] = pr[\Gamma \geq \ln(P_0) - z_0 \mid z_0] = \int_{in(P_0)-Z_0}^{\infty} f(\Gamma \mid z_0) d\Gamma \quad (15)$$

The poverty inflow probability (PIP) is the complement of the poverty outflow probability and can be formulated as

$$\Pr[Z \geq \ln(P_0) \mid Z^* = z_0] = 1 - pr[Z \geq \ln(P_0) \mid Z^* = z_0] \quad (16)$$

That is, PIP=1 - POP. From this, we can infer that PIP curve is symmetric to POP curve around the horizontal line at 0.5.

Equation (15) can be rewritten by use of the cumulative standard normal distribution function F(·) as:

$$POP = Pr\,[\Gamma \geq \ln(P_0) - z_0 \mid Z^* = z_0] = F\; \frac{-\ln(P_0) + z_0 + \dfrac{\sigma_{\Gamma Z^*}}{\sigma^2_{Z^*}}(z_0 - \mu_Z)}{\sqrt{\sigma^2_{\Gamma}(1-\rho^2)}} \tag{17}$$

Given $\mu_{Z^*}$, $\sigma^2_{Z^*}$, $\sigma^2_{\Gamma}$, and $\rho$, we can calculate POP. Since F(·) is continuously differentiable in Z*, and equivalently in Z, we can obtain a continuous locus corresponding to a set of the above mentioned parameters. By the definition of income percentiles (p), p is strictly monotonic in Z. Due to the one-to-one relationship between Z and p, we may have continuous POP curves based on p instead of Z in the X-axis. As the household size increases, the income mean increases. Furthermore, $\sigma^2_{Z^*}$ also changes with the household size. Therefore, different household types (in terms of household size) have different POP curves.

Recall the basic assumption that the expected income rank in the following period is the same as in the previous period. This means that poor households in the previous period tend to remain poor in the following period, unless Γ is sufficiently large. To the contrary, high-income households in the previous period tend to remain far from the poverty line in the following period. Therefore, the higher p is in the previous period, the higher POP is in the following period. However, the lower p is, the lower POP is. This implies that POP is increasing in p in the previous period at a decelerated rate. In other words, the acceleration diminishes. Thus, POP curves are generally concave (or, upwardly convex).

The can be verified mathematically. $Z^*( = z_0)$ in equation (17) represents a proxy variable for income rank or percentile p. Partial differentiation of equation (17) with respect to $z_0$ yields

$$\frac{\partial POP}{\partial z_0} = \frac{\partial F}{\partial z_0} = \frac{1+\sigma_{\Gamma Z^*} / \sigma^2_{Z^*}}{\sqrt{\sigma^2_Z(1-\rho^2)}}\,F'. \tag{18}$$

Note that $| \sigma_{\Gamma Z^*} / \sigma^2_{Z^*} | < 1$ and $\partial F / \partial z_0 > 0$. Therefore, POP is increasing in p.

POP is the probability that a poor household in the current period escapes from poverty in the next period, or that a non-poor household in the current period continues to earn above the poverty line in the next period.

The absolute poverty line varies with the household size. Each household type, in terms of household size, has different means and variances of log-income. Therefore, it is natural that each household type has a different POP curve. Location of POP curves depends on the relative distances between means and poverty lines discounted by standard deviations. Based on the HIES data, the location of the POP curve of a larger household becomes higher until it reaches around 4.

**Table 5.** Estimates of Poverty Outflow Probability Based on the 2007 HIES

(unit: %)

| Y Percentile(100p)\H. Size | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0.00 | 0.00 | 0.35 | 79.74 | 81.85 | 77.51 |
| 5 | 0.14 | 4.57 | 60.96 | 94.64 | 92.90 | 89.09 |
| 10 | 7.65 | 29.60 | 83.37 | 97.32 | 96.37 | 96.54 |
| 15 | 25.88 | 57.45 | 91.10 | 98.53 | 97.12 | 97.50 |
| 20 | 44.04 | 72.31 | 94.84 | 99.03 | 98.08 | 97.94 |
| 25 | 59.01 | 83.83 | 96.75 | 99.24 | 98.47 | 98.57 |
| 30 | 71.61 | 89.66 | 98.02 | 99.44 | 98.70 | 98.92 |
| 40 | 89.25 | 95.51 | 99.10 | 99.68 | 99.33 | 99.08 |
| 50 | 96.49 | 98.19 | 99.60 | 99.78 | 99.56 | 99.36 |

Notes: Y percentile denotes income percentile in the previous period. 100p denotes the relative income rank multiplied by 100.
H. Size denotes household size in terms of the number of household members.

**Figure 1.** Poverty Outflow Probabilities by Household Size Based on Market Income (HIES 2007)

(unit: %)

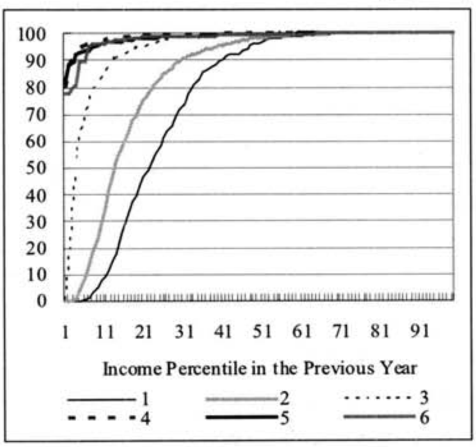



**Figure 2.** Poverty Inflow Probabilities by Household Size Based on Market Income (HIES 2007)

(unit: %)

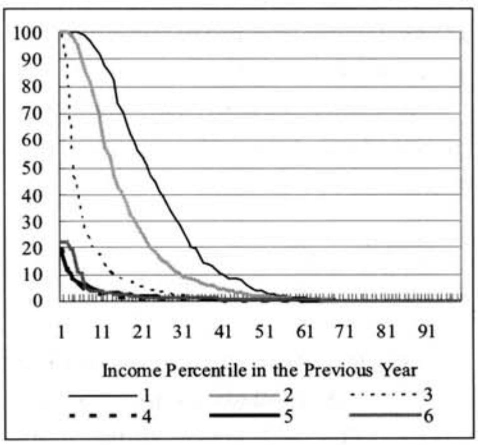



Figure 1 shows the estimated POP curves using the 2007 HIES data. The estimates of absolute poverty rates in 2007 are 23.1%, 13.6%, 3.9%, 0.8%, 1.4%, and 1.3% for household sizes of one through six, respectively. For a single or double household whose income percentile (p) was 0.01 in the previous period, the POP is very low, near zero. The POP increases with the household size. It increases dramatically up to the household size of four, especially in lower income percentiles. This is probably because heads of households of four or more are generally middle aged and also because the average number of employed or self-employed persons per household increases. Likewise, the POP increases with p in the previous period: for a single household, POP increases to 0.14% when p = 0.05, 7.65% when p = 0.10, et cetera. When p is around 0.25 or higher, POP approaches one, and it becomes almost certain that the households tend to remain far away from the poverty line in the next period.

Most single or double households comprise elderly persons, so their POPs tend to be very low. This is why their POP curves lie far below those of others. PIP curves in Figure 2 are just mirror images of the POP curves.

A decline in the poverty line leads to a rise in POP, and a decline in the mean of log-income leads to a decline in POP. Therefore, the resulting effect depends on the relative magnitudes. To illustrate this mathematically, consider the following partial derivatives of F with respect to $\ln(P_0)$ and $\mu_{Z*}$:

$$\frac{\partial F}{\partial \ln(P_0)} = -\frac{1}{\sqrt{\sigma_\Gamma^2 (1-\rho^2)}} F' < 0 \tag{19}$$

$$\frac{\partial F}{\partial \mu_{Z*}} = -\frac{\sigma_{\Gamma Z*} / \sigma_{Z*}^2}{\sqrt{\sigma_\Gamma^2 (1-\rho^2)}} F' > 0. \tag{20}$$

As the household size becomes smaller, the negative correlation between $\ln(P_0)$ and POP increases POP, since $\ln(P_0)$ becomes smaller. To the contrary, as the household size becomes smaller, the positive correlation between $\mu_{Z*}$ and POP decreases in POP, since $\mu_{Z*}$ becomes smaller. Therefore, when the household size becomes smaller, the location of POP curve depends totally on the absolute values of the effects of the two factors.

The sizes of longitudinal changes in $\ln(P_0)$ and $\mu_{Z*}$ are similar to each other in Korea. Note that $|\sigma_{\Gamma Z*} / \sigma_{Z*}^2| < 1$. Therefore, the fewer the household members are, the higher POP is.

According to the income transition rule, the level of income in the current period depends primarily on the income rank in the previous period and also on the value of income transition variable Γ. Γ, in turn, is negatively correlated with income rank. This means that there exists a mean reversing effect, whereby a household whose income is lower than the median income tends to have higher probability of Γ being positive; however, a household whose income is greater than the median income tends to have higher probability of Γ being negative.

Normality of log-income implies that mean of log-income equals its median. Thus, if the income is larger than the median income, i.e. $Z^*(=z_0) > \mu_{Z*}$, the conditional expectation of $\Gamma$ given Z* becomes negative: $E(\Gamma|Z^*)<0$. In this case, the conditional probability that $\Gamma$ given Z* is negative is greater than 0.5: $Prob(\Gamma<0|Z^*)>0.5$. As Z* becomes larger, this probability increases. Symmetrically, where the income is smaller than the median income, i.e. $Z^*(=z_0)<\mu_{Z^*}$, $E(\Gamma|Z^*)>0$ and $Prob(\Gamma>0|Z^*)>0.5$.

Longitudinal changes in income rank stem from $\Gamma$. POP (or PIP) depends on the variability of income rank, that is, $\sigma_\Gamma^2$ and $\sigma_{\Gamma Z^*}$. Since $\sigma_{\Gamma Z^*} = -\sigma_\Gamma^2/2$, POP depends only on $\sigma_\Gamma^2$. As the variability of $\Gamma$, i.e., the variance of $\Gamma$, increases, the width of longitudinal changes in income rank becomes wider. Therefore, for a household which was poor in the previous period, the probability of escaping from poverty in the following period (i.e., POP) becomes larger with higher . To the contrary, for a household whose income was greater than the poverty line in the previous period, POP diminishes as the variance of $\Gamma$ increases.

POP would not change much around the 20th percentile for a household of two people and around the 20th to 25th percentiles for a household of four, even if the variance of $\Gamma$ changes significantly. Therefore, as $\sigma_\Gamma^2$ increases, POP curves rotate clockwise around the above mentioned income rank regions, varying according to household size (see Figures 3 & 4).

**Figure 3.** Poverty Outflow Probabilities by Household Size =2 Based on Market Income (HIES)

(unit: %)

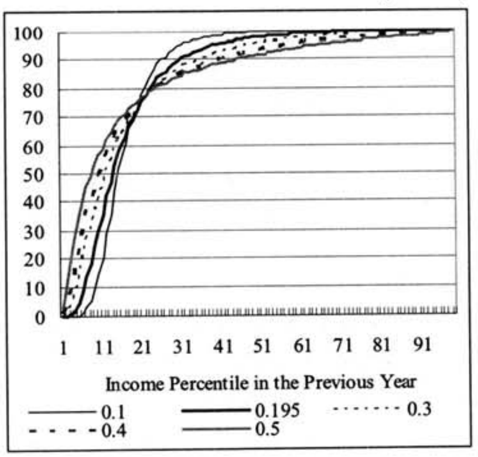
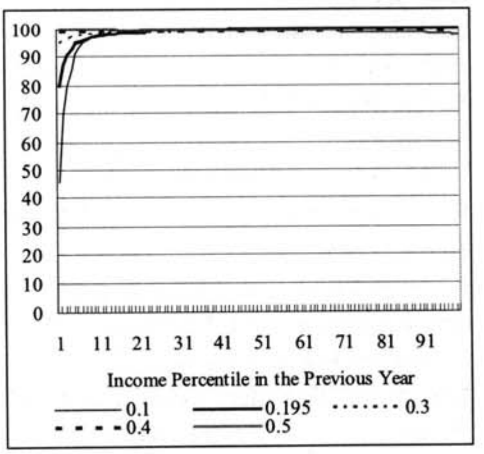



**Figure 4.** Poverty Inflow Probabilities by Household Size =4 Based on Market Income (HIES)

(unit: %)



## Direct Welfare Expenditure for the Poor with Differential Treatment[9]

A household is poor in the short run when its current household income is less than the poverty line. A household is poor in the long run when its present discount value of lifetime income flow is less than that of the lifetime poverty line. Figures in the Appendix show that in 2007 the short-run poverty rate at market income level was 10.88%. Note that because of income mobility, some households that are poor in the short run are not poor in the long run. As shown in Table 4, under the income mobility of 0.19492 measured by $\sigma_\Gamma^2$ in 2005-2006, the longer-run poverty rate diminishes as the time horizon is extended: it shrinks up to 3.415% until the ninth year. After that, it does not change any more. A long-run poverty rate lasting longer than nine years is regarded as lifelong poverty rate. Therefore, about 70% of poor households are no longer poor in the long run. These households need less support than lifelong poor households to cope with poverty. We infer from this that a slightly tighter fiscal expenditure might be sufficient to support all the poor with less fiscal burden in the long-run.

**Table 6.** Rates of Life-long Poverty Using a Pseudo-Panel

(unit: %)

| Household Size | 1 | 2 | 3 | 4 | 5 | 6 | Average |
|---|---|---|---|---|---|---|---|
| 1st Year | 2.896 | 6.774 | 4.374 | 6.866 | 7.321 | 6.634 | 5.633 |
| 2nd Year | 2.892 | 6.613 | 3.486 | 3.497 | 3.509 | 5.537 | 4.192 |
| 3rd Year | 2.764 | 6.292 | 3.573 | 2.352 | 4.058 | 5.537 | 3.837 |
| 4th Year | 2.878 | 6.354 | 3.073 | 2.374 | 3.837 | 5.537 | 3.744 |
| 5th Year | 2.764 | 6.345 | 3.204 | 2.092 | 1.887 | 5.537 | 3.548 |
| 6th Year | 2.764 | 6.584 | 3.141 | 2.077 | 2.227 | 5.537 | 3.610 |
| 7th Year | 2.764 | 6.484 | 3.1 | 2.122 | 1.933 | 5.537 | 3.569 |
| 8th Year | 2.764 | 6.484 | 2.956 | 1.884 | 1.342 | 5.537 | 3.430 |
| 9th Year~ | 2.764 | 6.484 | 2.956 | 1.846 | 1.272 | 5.537 | 3.415 |

Note: A pseudo-panel is compiled with the time horizon until year 2050 based on the income transition rule using the KLIPS data.

9. In this section, we adopt the basic structure of analyzing methods primarily from Chapter III of Sung (2007). We modify the methods by allowing that the distribution of Γ can vary with income ranks and/or ages in accordance of the result of empirical hypothesis testing, and update the available information. Based on these refined models, the estimates are totally recalculated.

We measured lifelong poverty rate by age group. For young households, it is almost zero or very low; for those 37 or younger, it is zero. However, the poverty rate of households with older people is high; for example, it is 9.8% for households of 60-year-olds and 14.5% for households of 65-year-olds. This tendency exists because those in the older generations are mostly retirees and because their remaining time horizon is relatively short: they have less chance to make up for current poverty gaps.

**Table 7.** Life-long Poverty Rate by Household Ages Measured in Current Period

(unit: %)

| Age | Life-long Poverty Rate | Age | Life-long Poverty Rate |
|---|---|---|---|
| ~37 | 0.00 | 52 | 0.96 |
| 38 | 0.41 | 53~54 | 0.00 |
| 39 | 0.00 | 55 | 1.90 |
| 40 | 1.25 | 56 | 0.76 |
| 41~43 | 0.00 | 57 | 1.65 |
| 44 | 0.62 | 58 | 8.79 |
| 45 | 0.22 | 59 | 4.50 |
| 46 | 0.59 | 60 | 5.28 |
| 47 | 0.26 | 61 | 9.75 |
| 48 | 0.48 | 62 | 16.87 |
| 49 | 0.00 | 63 | 14.40 |
| 50 | 0.63 | 64 | 21.35 |
| 51 | 3.25 | 65 | 14.54 |

Note: The pseudo-panel was compiled with the year 2050 as time horizon, based on the income transition rule and using KLIPS data.

It is necessary to support poor households. According to the National Basic Livelihood Security System (NBLSS), some current NBLSS beneficiaries are not poor in the long run but only in the short run. In this case, they will be able to repay at some time in the future. This is what this paper is focusing on. The government subsidizes all the poor households in the current period and recollects in the next period some of the subsidies from those who are poor in the current period but not in the long run. In this case, the fiscal burden of the government would be reduced roughly by the amount of repayment in the following periods[10].

10. The conditional repayment issue rather than the unconditional giving-out system like the current one is surely controversial because it requires a fundamental shift in philosophy. The discussion is not yet complete and, so, needs more careful discussion. This paper

The net costs are threefold. The first part is direct subsidies rendered to the lifelong poverty households ($C_2$). The second part is the foregone interest applicable to the repayment ($A_1$). The last part is the default: the amount not repaid by the short-run poverty households ($A_2$).

The net fiscal burden of the government can be estimated using the following equations:

$A = \int (PL\text{-}X)\cdot 1(PL>X)\cdot\{1\text{-}1(\text{Lifelong Poverty})\}f(X)dX$
interest cost: $A_1 = A \times$ interest rate (net)
default cost: $A_2 = A \times$ default rate
INDIRECT COST: $C_1 = A_1 + A_2$
DIRECT COST: $C_2 = \int (PL\text{-}X)\cdot 1(PL>X)\cdot 1(\text{Lifelong Poverty})f(X)dX$
X: household income
PL: poverty line
f(·): p.d.f. of log-normal variable
1(·): indicator function whose value is one if (·) is true, and zero otherwise.

Suppose that the government subsidizes all the poor households by making up for the whole gap between income and the poverty line. Assume that the interest rate is 5%, that the average duration of repayment is five years, and that the annual default rate is 10%. The total direct subsidies are the sum of subsidies distributed both to short-run and long-run poor households. The former are about 5.4 trillion and the latter are 1.6 trillion won. Therefore, the total direct subsidies amount to 7.0 trillion won. The foregone interest revenue distributed to the shorter-run poor households are 1.4 trillion won: the foregone interest revenue accrues from the cumulative balance for the past five years as well as from the current subsidies to the short-run poor. Since the average annual default rate is 10%, the annual default costs are 0.5 trillion won. Therefore, the net total costs are 3.5 trillion won.

In sum, about 7.0 trillion won is required to support all the poor households to guarantee a minimum subsistence level in the short-run. If lifelong poverty is distinguishable from short-run poverty, we could cut the fiscal expenditure roughly in half with differential treatment based on the duration and severity of poverty. However, this may not be practicable, not only because of the moral hazard but also because of the lack of necessary a priori information to distinguish lifelong poverty from short-run poverty. The challenge will be to provide policy tools to avert the moral hazard so

assumes that the regime already changes to the conditional repayment system and estimates its effects.

**Table 8.** Estimates of Government Direct Welfare Expenditure For the Poor

| Cost | | | | Units | Cost |
|---|---|---|---|---|---|
| # of Households in 2007 | | | H | | 16,417,423 |
| Short-run Poverty | per household | Total Subsidy | A | Won | 328,865 |
| | | Interest Cost (Cum.) | $A_1$ (=A×5%×5yrs) | | 82,216 |
| | | Annual Default Cost | $A_2$ (=A×10%) | | 32,887 |
| | | Direct Cost | $C_1$ (=$A_1$+$A_2$) | | 115,103 |
| | Total Cost | Subsidy | A × H | 100 Million Won | 53,991 |
| | | Interest (Cum.) | $A_1 \times H$ | | 13,498 |
| | | Default | $A_2 \times H$ | | 5,399 |
| | | Direct Cost | $C_1 \times H$ | | 18,897 |
| Lifelong Poverty | per household | Total Subsidy (No Repayment) | $C_2$ | Won | 95,568 |
| | Total Cost | | $C_2 \times H$ | 100 Million Won | 15,690 |
| Net Cost | Total Subsidies | | $(A+C_2) \times H$ | | 69,681 |
| | Total Cost | Total Subsidy (Net) | $C_2 \times H$ | | 15,690 |
| | | Interest | $A_1 \times H$ | | 13,498 |
| | | Default | $A_2 \times H$ | | 5,399 |
| | | Sum | C × H | | 34,587 |
| Saving | | | $(A+C_2-C) \times H$ | | 35,094 |

Note: The number of households in 2007 is from the National Statistical Office of Korea.

**Table 9.** Budget of National Basic Livelihood Security System

(unit: hundred million won)

| | National Basic Livelihood Security Expenditure | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008(a) |
| NBLSS (Total) | 32,696 | 34,034 | 35,230 | 39,127 | 46,520 | 53,758 | 65,759 | 68,440 |
| Assistance to Labor | 600 | 1,203 | 1,203 | 1,624 | 2,021 | 2,337 | 2,594 | 2,594 |
| Medical Assistance | 15,897 | 16,904 | 17,617 | 18,810 | 22,148 | 26,623 | 35,771 | 35,161 |
| Payment | | | | | | | | |
| for Disabled | 331 | 448 | 519 | 664 | 897 | 1,119 | 3,130 | 3,279 |
| for Elderly Pension | 1,999 | 2,460 | 2,145 | 2,150 | 2,126 | 2,153 | 2,175 | 15,948 |

Note: (a) denotes budget.
Source: Ministry for Health, Welfare and Family Affairs

that the potential savings can be realized.

The total budget for NBLSS expenditure in 2008 is 6.8 trillion won. Of this, the amount of cash transfer is less than half, which is far less than the estimated maximum (7.0 trillion won) needed to support all the poor households. This means that the current NBLSS budget cannot cover all poor households. However, if poor households can be classified into short-run and long-run poor, and if the short-run poor are required to repay the benefits they received, the government can financially support more poor households without any expansion of fiscal expenditure.

## CONCLUDING REMARKS

The major contributions of this paper are twofold. One is the development of a method of analyzing lifetime income mobility by constructing a pseudo-panel. This is the first time a micro-based pseudo-panel has been applied to the concept of lifelong income, enabling us to broaden the scope of empirical studies regarding income-related issues. Another contribution is the suggestion that a practical policy tool involving conditional repayment could be developed to relieve the fiscal burden and thus stabilize the long-run fiscal balance.

We estimated income mobility using the income transition rule between two consecutive years. We found that the income mobility measured by the variance of the income transition variable has diminished since the late 1990s. This is partly because the labor market became slightly more rigid; in addition, the population is aging rapidly and therefore the proportion of retirees is growing rapidly. As a result, the poverty outflow probability of a poor household has decreased.

The rate of short-run poverty, which is defined by current household income below the poverty line, is estimated to be 10.88% in 2007. About 7 trillion won a year are required to support all the poor households in the short run. The government could reasonably expect to cut the fiscal expenditure roughly by half by differentiating among poor households. Poor households can be divided into two groups: one is poor both in the short run and in the long run, and the other is poor only in the short run. The rate of long-run poverty, in which the present discount value of lifetime income flow is less than that of future poverty line values, is estimated to be 3.42% in Table 6. Unequal treatment based on the severity or duration of poverty can save the government in direct welfare expenditures to support the poor or, at least in part, leave room to expand its coverage of benefit recipients. In other words, the differentiated treatment of the poor may potentially relieve the fiscal burden of the government.

However, two problems arise: how to distinguish long-run poverty from short-run poverty, and the moral hazard involved. This paper does not provide solutions to these problems. We hope that subsequent studies will follow up on and and resolve these problems.

## REFERENCES

Aitchison, J., & Brown, J. A. C. (1957). *The lognormal distribution with special reference to its uses in economics*. Cambridge, UK: Cambridge University Press.

Alessie, R., Devereux, M. P., & Weber, G. (1997). Intertemporal consumption, durables and liquidity constraints: A cohort analysis. *European Economic Review*, 41: 37-59.

An, C. B., & Song, J. (2006). Introducing the Korean EITC: Its effects and plans [in Korean). *Korean Journal of Public Finance*, 20(2): 33-71.

Bane, M. J., & Ellwood, D. T. (1986). Slipping into and out of poverty: the dynamics of spells. *The Journal of Human Resources*, 21(1): 1-23.

Besley, T. (1990). Means testing versus universal provision in poverty alleviation programmes. *Economica, New Series*, 57(225): 119-129.

Gardiner, K., & Hills, J. (1999). Policy implications of new data on income mobilty. *The Economic Journal*, 109(453): F91-F111.

Gassner, K. (1998). An estimation of UK telephone access demand using pseudo-panel data (Cahier de recherche économique No. 9817). Lausanne, Switzerland: University of Lausanne, Ecole HEC-DEEP.

Girma, S. (2000). A quasi-differencing approach to dynamic modelling from a time series of independent cross-sections. *Journal of Econometrics*, 98: 365-383.

Hwang, D. (2001). A dynamic analysis on poverty after the economic crisis [in Korean]. *Labor Policy Studies*, 1(2): 31-59.

Jarvis, S., & Jenkins, S. P. (1998). How much income mobility is there in Britain? *The Economic Journal*, 108(447): 428-443.

Kalecki, M. (1945). On the Gibrat distribution. *Econometrica*, 13(2): 161-170.

Keum, J. (2006). A study on working poor in Korea after the financial crisis [in Korean). *Korean Labor Economic Association Dissertation*, 29(1): 41-73.

Kim, D. C., & Kwon, H. J. (2007). The impact of national pension scheme poverty reduction of the elderly [in Korean]. *Korean Social Security*, 23(3): 251-275.

Kim, J. M., & Sung, M. J. (2003). Derivation of cohort life-cycle age-income profile using characteristics of income distribution [in Korean]. *Journal of Korean Economic Analysis*, 9(3).

Kim, J. M., & Sung, M. J. (2004). *Social security policy: Long-term expenditure projections and policy directions* (Research Report 04-02) [in Korean]. Seoul: Korea Institute of Public Finance.

Lee, J.-W., & Lee, S. (2001). Economic crisis and income disparity: Income distribution and poverty before and after the 1997 crisis [in Korean]. *International Economic Journal*, 7(2): 79-109.

Lee, S. (2004). Direction of poverty policy for the poor households with work ability [in Korean]. *Korean Social Security*, 20(3): 27-56.

Lee, S. (2006). The types and effects of non-contributory income maintenance strategies for low-income people with work ability [in Korean]. *Korean Journal of Social Welfare Studies*, 30: 101-130.

Lee, S. (2007). An evaluation and policy tasks of the Korean earned income tax credit [in Korean]. *Korean Social Security*, 23(3): 25-55.

Lim, S. H. (2006). The determinants of poverty-exit possibility: By means of economical factors [in Korean]. *Korean Social Security*, 22(2): 253-277.

McKenzie, D. J. (2004). Asymptotic theory for heterogeneous dynamic pseudo-panels. *Journal of Econometrics*, 120: 235-262.

Park, C.-y., & Kim, J.-W. (2000). Poverty measurement by characteristics of household head during the economic crisis [in Korean]. *Korean Social Security*, 16(1): 1-23.

Park, S.-w., & Kim, S.-h. (2007). The characteristics and determinants to poverty since economic crisis 1997 [in Korean]. *Korean Economic Studies*, 25(3): 177-196.

Pelzer, B., Eisinga, R., & Franses, P. H. (2001). Estimating transition probabilities from a time series of independent cross-sections. *Statistica Neerlandica*, 55: 249-262.

Rodgers, Joan R., & John L. Rodgers (1993). Chronic poverty in the United States. *The Journal of Human Resources*, 28(1): 249-262.

Seok, S. H. (2007). Testing for state dependence effects in poverty dynamics [in Korean]. *Korean Social Security*, 23(3): 79-99.

Sim, S.-Y. (2006). An empirical study on the transition of the micro traits of poor salary and wage earners' households: Focus on the Family Income and Expenditure Survey 1982-2004 [in Korean]. *Korean Social Security*, 22(4): 163-189.

Sung, M. J. (2007). A study on recent poverty dynamics and estimation of public cash transfers to the poor in Korea [in Korean]. *Korean Journal of Public Finance*, 22(1): 61-88.

Sung, M. J., & Park, K. (2008). Redistributive effects of taxes and transfers including consumption taxes and in-kind benefits [in Korean]. *Public Finance Review*, 1(1): 63-94.

Verbeek, M., & Nijman, T. (1993). Minimum MSE estimation of a regression model with fixed effects from a series of cross-sections. *Journal of Econometrics*, 59: 125-136.

Verbeek, M., & Vella, F. (2005). Estimating dynamic models from repeated cross-sections. *Journal of Econometrics*, 127: 83-102.

Yoo, G. (2002). Analyses on the changes in income distribution around the economic crisis and on the income mobility in Korea [in Korean]. *KDI Policy Research*, 24(2).

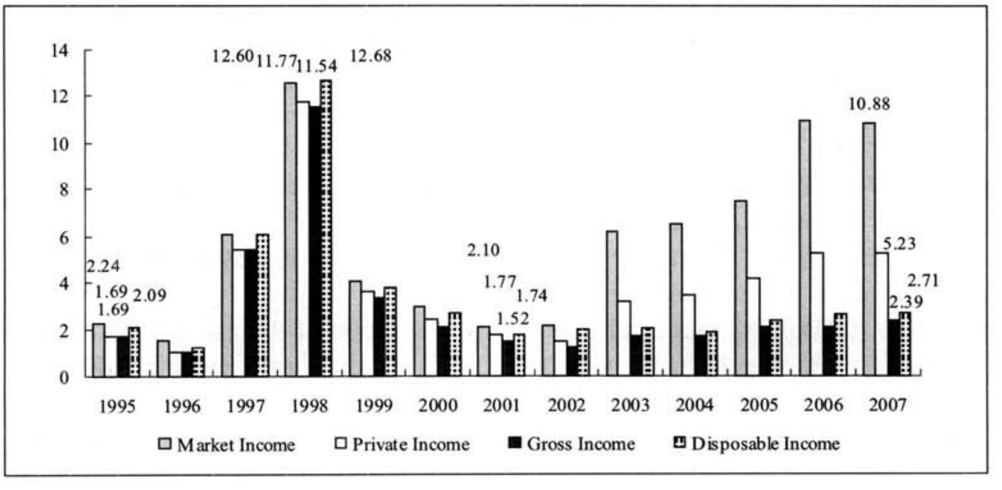
## APPENDIX: POVERTY RATES

*Poverty* is usually defined as household income being below the poverty line. In this paper, we consider the absolute poverty only[11]. Usually, disposable income is used to calculate poverty rates. The absolute poverty rate for market income was 2.24% in 1995, sharply rose to 12.60% in 1998, shrank back to 2.10% in 2001, and jumped up again to 10.88% in 2007.

The differences in poverty rates between market income and disposable income indicate the effects of taxes and transfers on poverty. The gap becomes larger with time: it was 0.15%p in 1995, 0.36%p in 2001 and 8.17%p in 2007.

**Figure A-1.** Absolute Poverty Rates

(unit: %)




Regarding the definition of income, Sung and Park (2008) proposed a definition of different types of incomes:

- Market Income (MY): Income from employment, investment, etc
- Private Income (PY): MY + Private Transfer Income
- Gross Income (GY): PY + Public Transfer Income
- Disposable Income (DY): GY - Direct taxes
- Post-tax Income (PTY): DY - Indirect Taxes

11. Relative poverty is not substantially different from absolute poverty to the extent that the HIES is concerned in Korea.

According to the HIES, private transfers are mostly pecuniary subsidies for living expenses transferred between parents and offspring or among close relatives. The social security system in Korea does not yet fully cover the whole population, and therefore private transfers have played an important role up until the present. From market income to post-tax income, the poverty rate decreased from 10.88% to 3.19% in 2007. About 70% of poor households were relieved from poverty by private transfers and government fiscal expenditures.

**Figure A-2.** Changes in Absolute Poverty Rates by Numerous Incomes in 2007

(units: %, %p)