Inequality and Mobility of Household Earnings over the Life Cycle and Business Cycles*

Sanguk Kwon†

November 2020

Abstract

Using a long panel of Korean household data, we examine the trend of inequality and mobility of household earnings after the Asian Financial Crisis. From Korean household data, we find that the following stylized facts: (1) income inequality is correlated with the macroeconomy, (2) households with initial low income have higher future income, and (3) income mobility is persistently high. We propose generalizations of two conventional models of income process to explain those stylized facts. We show that our proposed models outperform the conventional models based on the model selection criteria of Andrews and Lu (2001). Our generalized income models show that a significant source of income inequality is heterogeneous responses by households to aggregate shocks. The sources of income inequality differ by the education level of the head of households.

JEL Codes: D10, E20, E25

Keywords: inequality, mobility, income process

*I thank Shawn Ni, David Kaplan, Aaron Hedlund, and Rui Yao for helpful comments and discussions.
†Department of Economics, University of Missouri, United States of America.

1 Introduction

The trend in household inequality has been heavily studied in empirical and theoretical economic literature. Quantitative analysis of the household income dynamics hinges on the assumptions to identify two income components (permanent component and transitory component). Distinguishing permanent and transitory components is necessary to understand the source of income inequality, which is useful to make economic policy. If an increase in permanent inequality is a main driver of inequality, there are required public policies to improve career trajectories, such as education and health. By contrast, if an increase in transitory inequality is a main driver of inequality, social welfare can potentially be improved by better social insurance and access to credit market for short-term income shortfalls.

There are two main approaches to model the permanent component in the unexplained income. One widely assumed model for the permanent component is a random walk, which Guvenen (2009) calls a Restricted Income Profiles (RIP) process. For model identification, the shocks to the permanent component and the innovations to the transitory component are assumed to be (a) - independent over time for each household and (b) - identically and independently distributed cross households in each period (see e.g., Meghir and Pistaferri (2004); Blundell, Pistaferri, and Preston (2008); Blundell and Etheridge (2010); Jappelli and Pistaferri (2010); Moffitt and Gottschalk (2011)). Meghir and Pistaferri (2004) allow the distributions of income shocks to be time varying while maintaining the assumptions (a) and (b).

Another approach models life-cycle earnings by a function of experience profiles, with household-specific but time-invariant random parameters on the intercept and the slope. A vast literature accounts for the heterogeneity in the average of lifecycle income profiles (e.g., Mincer (1974); Becker (1975); Carroll and Summers (1991); Baker (1997); Haider (2001); Guvenen (2009), etc). Guvenen (2009) labels this a Heterogeneous Income Process (HIP) model in contrast to the RIP model, and shows that the HIP for unexplained earnings fits Panel Study of Income Dynamics (PSID) data well.

Both the RIP and HIP models exclude an aggregate factor under an implicit assumption that aggregate shocks affect the cross sectional mean of income and the variance of idiosyncratic shocks among households. In other words, they exclude the scenario that some households are always more affected by aggregate shocks than other households. They also rule out the possibility that income shocks are related to past income.

Several recent empirical studies on U.S. household income data suggest that household income shocks are not iid and maybe dependent on past income. Arellano, Blundell, and Bonhomme (2017) estimate a nonparametric model of PSID household earnings and find the presence of nonlinear persistence and conditional asymmetries in earnings. Guvenen, Karahan, Ozkan, and Song (2015) analyze a large sample of life-cycle earnings of U.S. workers and find heterogeneity: in each period where most individuals experience small earnings shocks and few experience large shocks. They find that low-income workers experience large earnings shocks with low persistence, whereas high-income individuals experience shocks that are very persistent but with much lower volatility. Also, positive (negative) shocks to high-earnings workers are transitory (persistent), and the opposite is true for low-earnings workers.

Comparing the two income process models that consider homogeneous effects of income shocks on each household under the iid assumption, we consider economic uncertainty as well as household-specific uncertainty. This study explores parametric income models in which income shocks depend on latent household-specific and macro factors. We model the income process with the focus on the dependence and heterogeneity of income shocks to the macroeconomy by using a long panel of Korean household data.

The long longitudinal data allow us to learn other cross-household moments related to the income inequality beyond a simple measure of inequality (e.g., variance of cumulative income and the relationship between initial income and income growth). These moments are also used to test models. One additional aspect of households income dynamics we observe from the household panel is income mobility, which represents the change in the relative standings of households. Inequality and mobility concern different aspects of distributions of household variables. Inequality may be studied by cross-sectional or panel data. Mobility can only be studied in a panel data setting. Higher mobility is desirable but it does not necessarily associated with reduced inequality. From an economic policy perspective, inequality (dispersion) of income matters, so is mobility.

In this paper, we address the following empirical questions:

(1) What is the trend of cross sectional distributions of household earnings? Some aspects of the distributions alluded earlier include inequality, measured by cross sectional variance; correlation between initial income and future income growth; and income mobility.

(2) Do commonly employed models fit the trend of cross sectional distributions of household earnings? If not, what model does? And how does the best model explain the drivers of income inequality?

The previous studies have not paid too much attention to those issues. This paper fills this gap by answering the above questions.

Most of existing Korean studies analyze the trend of Korean household income inequality. They draw mixed conclusions on the trend in the past two decades. A number of studies (Kim and Kim (2015); Hong (2015)) find that the share of top income households has steadily increased since mid-1990s, indicating a widening inequality. On the other hand, Choi, Kim, and Park (2018) find that the Gini coefficient of household income fluctuates after the 1998 Asian Financial Crisis and trends downwards after the 2008 Global Financial Crisis. These studies suggest that the measurement of income inequality is influenced by the metric and data. While a specific metric for inequality provides a simple description, it does not by itself shed much light on the drivers of inequality.

We explore inequality in household earnings by following a cohort of households in Korean Labor and Income Panel Study (KLIPS) over 19 years, from 1997 to 2015. By tracking a fixed cohort over the life cycle, we are able to examine how earnings of each household evolve over time and relative to earnings of other households. The sample period includes large macroeconomic shocks of the Asian Financial crisis of 1998 and the Global Financial Crisis of 2008 and the subsequent recoveries.

We focus on household labor income inequality for households with a head under working age during 1997-2015. Empirical evidence suggests that the lifetime income profile may shift over time but not proportionally. We estimate alternative generalizations of the RIP and HIP that differ from the Baker and Solon (2003) model, by allowing income profile shifting over time but not proportionally. In particular, we examine a generalized RIP (G-RIP) model with a heterogeneous permanent effect of macro factor, and a generalization of the HIP model with "stochastic experience premium" (stochastic heterogeneous income profile, or SHIP model). We estimate competing models by the method of moments. For model comparison, we use the selection criteria proposed by Andrews and Lu (2001) that balance the fit and the size of the models.

We make the following contributions to the literature:

(1) We report empirical pattern of lifecycle income dynamics in KLIPS. We find that cumulative household income changes over a range of distant future are negatively correlated with the initial income. This suggests that the permanent income shocks in distant future are correlated with the current income.

(2) We estimate competing models of income process with the focus on the dependence and heterogeneity of income shocks. We consider a battery of parametric models that include and nest the RIP and HIP models. We find that the G-RIP and SHIP models are better than the RIP and HIP models for the KLIPS data based on the model selection criterion of Andrews and Lu (2001). We then decompose the components of income inequality in each model.

In Section 2 we discuss the KLIPS data set and report observed trends in household earning inequality and mobility. In Section 3 we outline our approach to estimate and select competing models. In Section 4 we report the estimate and misfit of the exiting income models, compare that with estimated alternative income models and analyze the drivers of income inequality. Section 5 concludes.

2 Empirical Patterns in the KLIPS Data

2.1 KLIPS Data

The Korean Labor and Income Panel Study (KLIPS) is comparable to the PSID in the U.S. KLIPS is a publicly available annual survey data on social and demographic information as well as economic information of individuals and households. The first wave of KLIPS started from 5,000 urban households in 1998 and 1,415 new households were added in 2008 to compensate for sample attrition.

KLIPS uses a two-stage stratified clustering method for sample and data collection. Enumeration districts (EDs) were selected and then households were selected in the EDs. Every household in the population were assigned into only one ED. KLIPS collects data via annual face-to-face surveys. Every member aged 15 or older of each household participates in annual surveys on household and individual information such as economic activities, employment status, income, expenditures, education, housing status, financial and real property wealth, debt, and so on.

The key variable is household earnings. The KLIPS surveys earnings of the previous year. We use the nineteen waves from 1998 (on 1997 income data) through 2016 (on 2015 income data) to analyze inequality and mobility and only focus on married households with observations in all 19 years. Our study focuses on household's labor income for households with a head going through the lifecycle of working age. So we follow several common practices in sample selection (e.g., Blundell, Pistaferri, and Preston (2008)). First, divorced, separated, and one-person households are dropped. Second, households with missing report on labor income are dropped. Third, split offs of the original household are out of sample. Finally, households with a heads under 25 years old or over 65 years old are dropped in all years

(1997-2015). Details of analysis on sample selection are provided in Appendix A. Table A-5 of Appendix A suggests attrition bias to the sample is not significant.

Some studies remove households experiencing large or small income changes in accord with the assumption of cross-household homogeneity in the income shocks. We do not remove income outliers because our analysis is not reply on the assumption of iid income shocks. But the empirical pattern reported later is qualitatively the same for the sample with or without outliers. The final sample after filtering work consists of the balanced panel data with 545 households (with 10,355 observations). We use 2015 as the base year to compute for real earnings.

Table 1 and 2 report household demographic variables and a decomposition of disposable income by education level, respectively. Family size declines since 2004 and a number of workers steadily increase over time, which implies that as parents are getting old, their descendants tend to have a job and make new household moving out of their parents' house. Approximately fifty percent of sample households live in metropolitan cities of more than one million people. Thirty percent of households have a head with college or higher educated and higher educated households are apt to have higher labor income and disposable income.

2.2 Detrended Income

The income process for each household i is (for $t = t_i, \cdots, T$, $i = 1, .., N$)

$$Y_{it} = \mathbf{Z_{it}} \gamma_{\mathbf{y}} + y_{it} \tag{1}$$

where real (log) labor income, Y, can be decomposed predictable and stochastic components. For the former, $\mathbf{Z}$ is a $1 \times m$ row vector of income characteristics observable (e.g., a head's education and birth year, residential district, family size, number of workers, number of children, and year dummy) and known by households at time t. $\gamma_{\mathbf{y}}$ is an m-dimensional unknown parameter vector. Boldface letters represent vectors or matrices. Defining $y_{it} =$

Table 1. Average demographic characteristics

Year	Family size	Children	Workers	Share of bigcity$^{a)}$
1997	3.983	1.725	1.420	0.554
1998	4.050	1.750	1.510	0.552
1999	4.064	1.714	1.505	0.547
2000	4.092	1.681	1.543	0.545
2001	4.110	1.628	1.628	0.541
2002	4.127	1.538	1.655	0.547
2003	4.123	1.455	1.672	0.539
2004	4.114	1.352	1.716	0.538
2005	4.097	1.255	1.758	0.538
2006	4.057	1.163	1.835	0.534
2007	4.006	1.039	1.853	0.528
2008	3.961	0.943	1.837	0.525
2009	3.901	0.837	1.868	0.527
2010	3.864	0.736	1.877	0.525
2011	3.798	0.639	1.910	0.527
2012	3.728	0.543	1.883	0.521
2013	3.695	0.428	1.850	0.523
2014	3.622	0.343	1.927	0.517
2015	3.538	0.250	1.875	0.517

Note: The table is based on the balanced panel of 545 households. $^{a)}$ represents the proportion of a household that live in metropolitan cities of more than 1 million people.

$Y_{it} - \mathbf{Z_{it}}\gamma_{\mathbf{y}}$ as the log of real labor income net of predictable individual components. Because of the presence of year dummy in the equation (1), y_{it} has mean zero in each year t.

2.3 Measuring Income Inequality and Mobility

An increase in variance affects both income inequality and income mobility. Inequality is a static statistic and mobility is a dynamic one. We use cross-sectional variance $var(y_{it})$ to measure income inequality and the correlation of the deviation from the mean over time to measure income mobility.

Denote $z_{it} = \sqrt{\frac{y_{it}}{var(y_t)}}$. By construction z_{it} has a sample mean of 0 and sample variance of 1. Is the deviation from mean persistent $(t = 1, .., T)$? We run the regression $z_{it} = \rho_t z_{it-1} + \mu_{it}$.

Table 2. Average labor, transfer, and disposable income by education level

	All			College or higher educated			High school or under educated		
Year	L.I[a)]	T.I[b)]	D.I[c)]	L.I	T.I	D.I	L.I	T.I	D.I
1997	3,471	1	3,584	4,335	0	4,477	3,105	1	3,207
1998	3,145	45	3,551	4,026	43	4,609	2,773	46	3,104
1999	3,441	23	4,008	4,314	42	4,827	3,072	15	3,661
2000	3,742	27	4,105	4,719	42	5,593	3,329	21	3,476
2001	4,291	66	4,764	5,503	140	6,498	3,779	34	4,031
2002	4,515	52	5,089	5,727	82	7,074	4,002	39	4,249
2003	4,933	61	5,262	6,410	105	7,010	4,309	42	4,523
2004	4,990	103	5,414	6,268	183	6,862	4,450	69	4,802
2005	5,268	117	5,928	6,546	171	7,917	4,728	94	5,086
2006	5,475	164	6,188	6,853	258	8,074	4,892	124	5,390
2007	5,820	110	6,303	7,279	176	7,993	5,203	83	5,588
2008	5,675	85	6,210	7,211	96	7,923	5,026	80	5,486
2009	5,765	109	6,553	7,399	197	9,149	5,074	73	5,455
2010	5,819	131	6,332	7,313	156	8,132	5,187	120	5,571
2011	5,892	152	6,409	7,347	205	8,052	5,277	129	5,714
2012	5,994	154	6,430	7,672	119	8,217	5,284	168	5,675
2013	6,102	191	6,769	7,485	196	8,330	5,516	188	6,109
2014	6,202	220	6,796	7,755	203	8,685	5,546	228	5,997
2015	6,170	304	6,969	7,619	425	9,158	5,558	253	6,044

Note: The table is based on the balanced panel of 545 households. [a)], [b)], and [c)] represent real labor income, real transfer income, and real disposable income, respectively and unit is ten thousand Korean Won in 2015.

The OLS estimate $\hat{\rho}_t = \frac{\sum_{i=1}^{n} z_{it-1} z_{it}}{\sum_{i=1}^{n} (z_{it-1})^2} = corr(y_{it-1}, y_{it})$. We use $1 - \hat{\rho}_t^2$ to measure mobility. [1]

[1]More generally, the density of z_{it} is $f_t(.)$ and the CDF is $F_t(.)$. The ranking of household i is $F_t(z_{it})$ $(0, 1)$. The statistic $M_t = 1 - corr(F_t(z_{it}), F_{t-1}(z_{it-1}))$ measures mobility. Alternatively we can use the following discrete measure: Since z_{it} is memorialized to have mean 0 and variance 1, we divide the space of z_{it} into K intervals $(-\infty, a_1), (a_1, a_2), ..., (a_K, +\infty)$. Denote the state $s_{it} = k$ if z_{it} is in the k-th interval. Then the mobility of ranking of household i is given by a $K \times K$ transition matrix $\mathbf{T_t}$ whose (k, m) element is $\mathbf{T}_t(k, m) = prob(s_{it} = m | s_{it-1} = k)$. We find the alternative measures of mobility yield similar results.

Table 3. Residual statistics

Year	Variance	Skewness	Kurtosis
1997	0.241	0.012	6.066
1998	0.243	-0.774	4.875
1999	0.201	-0.261	4.509
2000	0.199	-0.589	7.062
2001	0.193	-0.866	6.397
2002	0.193	-0.388	4.192
2003	0.210	-0.225	4.972
2004	0.213	-0.334	5.651
2005	0.195	-0.148	6.344
2006	0.197	-0.756	7.562
2007	0.182	0.049	3.782
2008	0.227	-0.949	8.827
2009	0.194	0.134	4.369
2010	0.234	-1.093	9.391
2011	0.220	-0.980	9.847
2012	0.199	-0.099	3.648
2013	0.244	-0.675	7.104
2014	0.252	-0.535	5.620
2015	0.287	-1.540	14.467

2.4 Implications of the RIP and HIP Models

Let y_{it} be unexplained earnings of household i in period t. The RIP model assumes

$$\textbf{RIP} : y_{it} = P_{it} + v_{it},$$

where permanent component P_{it} is a random walk (i.e., $P_{it} = P_{it-1} + \zeta_{it}$, where ζ is the permanent shocks) and independent of the transitory shocks v_{it}. In the literature on the dynamics of household income inequality, the variance of persistent income shocks in each period is identical for all households (but may be time-varying); so is the variance of transitory income shocks.

Suppose the head of the household has experience t_i in the initial period of the sample. We calculate the initial work experience of household head i by $t_i = initial\ age_{it} - 26$. Then with the life-cycle earning HIP model

$$\textbf{HIP} : y_{it} = \delta_i + (t + t_i)\beta_i + v_{it},$$

where the slope β_i is the experience premium.

The income models have quite different implications on household risk, inequality, and mobility. The RIP model lumps together household risk, inequality, and mobility. Assumption (b) implies that a larger variance of income shocks in a given period captures both the higher income risk faced by each household and larger income inequality across households, as well as more potential for mobility in the relative rank across households.

The HIP model separates household risk, inequality, and mobility. Tracking the lifecycle of a cohort of households with income generated by the HIP model yields transitory income risk, expanding inequality and diminishing income mobility. The combined RIP-HIP model by Baker/Solon separates but restricts the dynamics of household risk, inequality, and mobility.

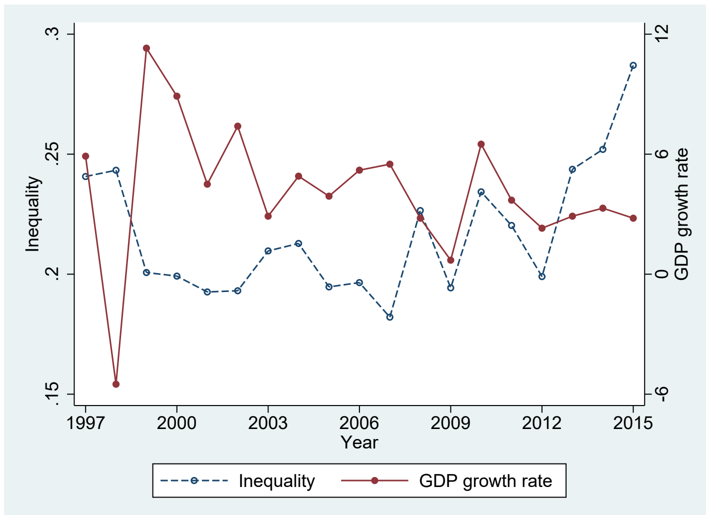
2.5 Stylized Facts and the RIP and HIP Models

Table 3 presents descriptive statistics of income, y_{it} detrended from the equation (1). We find five observations on KLIPS household earnings (y is the detrended household labor income thereafter). In the 1998 Asian Financial Crisis and the 2008 Global Financial Crisis the KLIPS households earnings are more dispersed, negatively skewed, and have with heavier tails. These also occur the end of the sample, possibly because of heterogeneity in retirement timing. We argue that neither the RIP nor HIP model explains all of the observations.

Observation 1.

The labor income inequality, var(y) is negatively correlated with GDP growth (i.e., counter cyclical), especially during recessions. Labor earning inequality increases but GDP growth rate decreases during the East Asian Crisis of 1998 and the Global Financial crisis of 2008 (See Figure 1).

Figure 1. var(y) and GDP growth rate



If we only focus on income inequality, the correlation of earnings inequality and GDP growth may be captured by a time-varying variance of income shocks ζ_{it} or v_{it}. However, the time-varying variances of iid income shocks may not match other cross-household moments of the income data.

Observation 2.

Inequality of cumulative growth income (i.e., $var(y_t - y_1)$) expands, which implies a permanent force of divergence. In the RIP model it can be used to estimate the variance of permanent component relative to the variance of transitory component. In the HIP model the divergence of $var(y_t - y_1)$ may due to accumulation of a permanent component (intercept) or the presence of a heterogeneous experience premium proportional to experience (slope β_i).

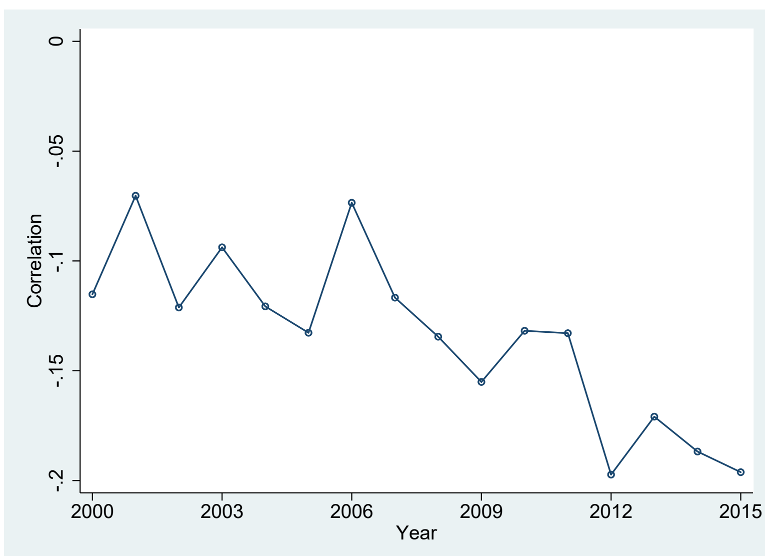
Observation 3.

The covariance $Cov(y_t - y_3, y_1) < 0$, households with high initial income y_1 experience more large negative shocks later in the lifecycle. This contradicts to the RIP model. In the HIP model the negative correlation between $y_t - y_3$ and y_1 is consistent with a negative correlation between β_i and δ_i.

Figure 2. $\mathrm{var}(\Delta y)$ and $\mathrm{var}(y_t - y_1)$

(a) Variance of growth income (b) Variance of cumulative growth income

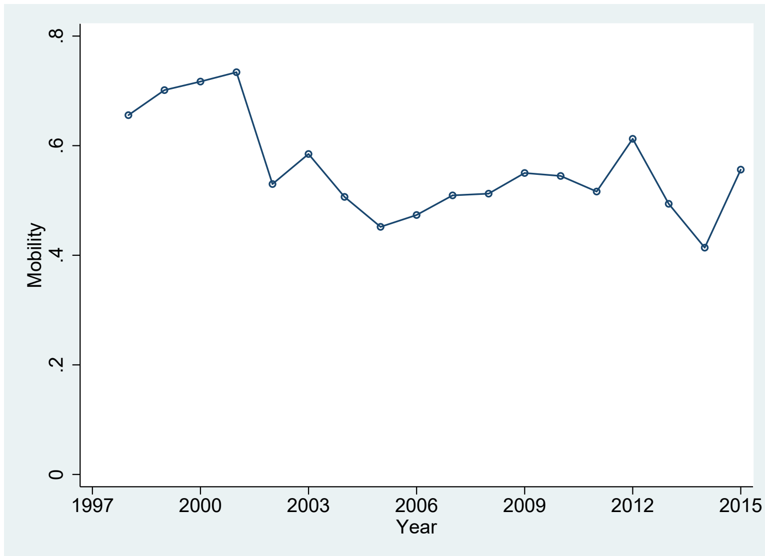
Figure 3. $\mathrm{corr}(y_t - y_3, y_1)$



Observation 4.

Labor income mobility is higher post-recession, and is relatively persistent. This suggests that in the RIP model permanent component has to be small (relative to transitory shocks), otherwise the income will diverge quickly over time and the mobility shrinks over time. In the HIP model variance of β_i is close to zero or the variance of transitory income shocks has to grow over time, otherwise income mobility should shrink to 0.

Figure 4. Mobility$(1 - \hat{\rho}_t{}^2)$

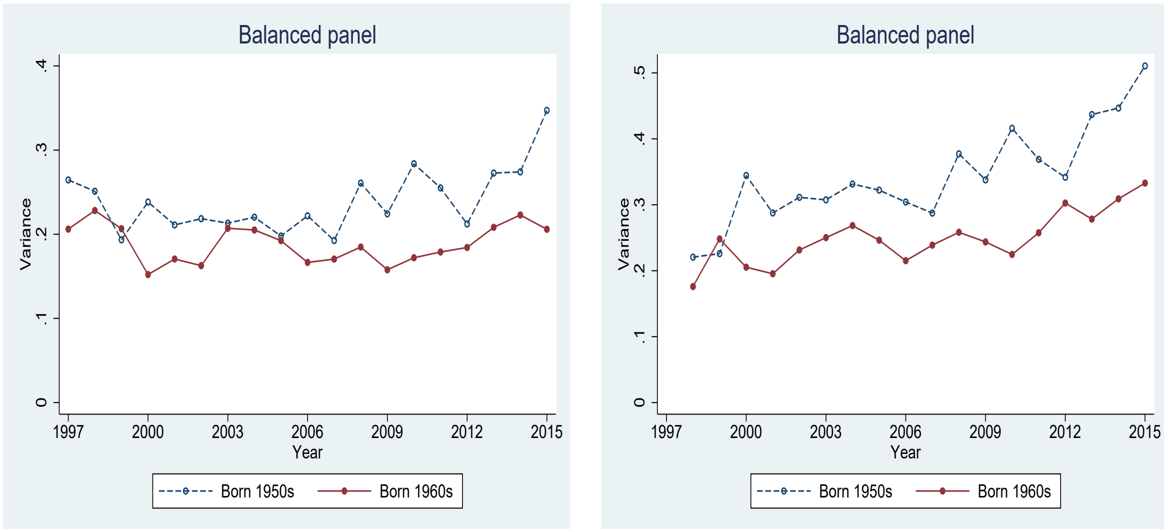


Observation 5.

The within-cohort $var(y_{it})$ is higher by cohort of born earlier. $var(y_{it} - y_{i1})$ expands for all cohorts, and expands faster for older households. One may argue that Figure 5(a) is consistent with the theory that the post 1997 permanent income shocks on different cohort are of the same variance but these shocks are small than that prior to 1997. However, the theory does not explain the Figure 5(b), which shows the variance of cumulative growth of older cohort grows faster than that of the younger one. If all cohorts are subject to the same variance of idiosyncratic permanent shocks, the variance of cumulative growth should be the same. And if the idiosyncratic permanent shocks are small as Figure 5(a), then there should not be expanding variance for the older cohort in 5(b).

Figure 5(b) is consistent with the HIP model, in which the heterogeneous income profile diverges over the life cycle (diverging experience premium=experience $\times \beta_i$ for household i). However the HIP with a constant β_i over time cannot explain the flat within-cohort inequality $var(y_{it})$ over the life cycle. The constant β_i version of the HIP also predicts that over the life cycle the households become more locked-in in the income ranking. This

Figure 5. var(y_t) and var($y_t - y_1$) by cohorts



(a) var(y_t) (b) var($y_t - y_1$)

also contradicts the persistent mobility shown in Figure 4. The data on life-cycle income inequality by cohorts suggest the presence of stochastic experience premium. The variance of the experience premium is trending lower post 1997.

We now go over the above arguments in algebra in a model that combines the RIP and HIP models. Suppose for household with initial experience t_i

$$y_{it} = (t + t_i)\beta_{it} + P_{it} + v_{it}. \tag{2}$$

If $\beta_{it} = 0$ then we have the RIP model with $var(y_{it}) \approx var(P_{i0}) + \sum_{j=1}^{t} var(\zeta_{it}) + var(v_{it})$. If all cohort experience the same $var(\zeta_{it})$ then the gap in $var(y_{it})$ across cohort has to be explained by the gap in $var(P_{i0})$, i.e., older cohorts are more diverse in 1997 because they have accumulated more years of permanent shocks than younger households up to 1997. For $var(y_{it})$ to be flat over the lifecycle there should be very small $var(\zeta_{it})$ (hence no trend in the accumulation of the $var(\zeta_{it})$). But Figure 5(b) suggests the presence of substantial permanent shocks at least for the older cohort. Hence this rejects the hypothesis $\beta_{it} = 0$. Another piece of evidence against the hypothesis $\beta_{it} = 0$ is the **Observation 3**. If $\beta_{it} = 0$

future shocks should be independent of y_1 and not negatively correlated with y_1.

Now consider the hypothesis $\beta_{it} = \beta_i$ and let $P_{it} = 0$. In the case we have the HIP model, with $var(y_{it}) \approx (t + t_i)^2 var(\beta_{it}) + var(v_{it})$. Since $(t + t_i)^2$ expands over time, for $var(y_{it})$ to stay flat over 20 years $var(\beta_{it})$ must be very small and/or $var(v_{it})$ must shrink fast over time. But if that were the case then we should not observe the growing trend in Figure 5(b) for the older cohort. Hence the HIP model does not explain all of the observations neither.

3 Estimation of Income Models

In the literature on the dynamics of household income inequality, the variance of persistent income shocks in each period is identical for all households (but may be time-varying); so is the variance of transitory income shocks. A fundamental identifying assumption is that persistent income shocks are serially and cross-sectionally uncorrelated. This assumption implies that household persistent income shocks are not directly related to shocks common to the households.

Data start on income of 1997, we track households for 19 years, to 2015. Denote the vector of unexplained income for household i as $\mathbf{y_i} = (y_{i1}, .., y_{iT})$, $T = 19$.

By design the cross-household average of $\mathbf{y_i}$ are zero. The $T \times T$ sample covariance matrices of $\mathbf{y_i}$ is

$$Var(\mathbf{y}) = \sum_{i=1}^{N} \frac{1}{N} \mathbf{y_i} \mathbf{y_i} \tag{3}$$

In matrix form we denote variables $\mathbf{Y} = (\mathbf{y_1}, \cdots, \mathbf{y_N})_{N \times T}$. Then $\frac{1}{N}\mathbf{Y}\ \mathbf{Y}$ is a $T \times T$ covariance matrix and diagonal element represents cross-sectional variance for each year.

3.1 Models of Income Process with a Macro Factor

We now introduce a macro state proxied by f_t into the model (2). Experience in the initial period of the sample of a head of a household i is t_i. We assume three sources of variation in detrended labor income for a household i, y_{it}: a direct household-specific impact of macroeconomy $\delta_i f_t$, a permanent component, and a transitory component.

A central feature is the correlation between business cycle and inequality. The correlation may stem from three sources: the direct effect, change in the variance of permanent component and change in the variance of transitory component.

We consider the following model:

Let the macro state be f_t and the permanent component of a novice worker be α_i. Similarly the initial experience premium for novice workers is β_{i0}.

$$y_{it} = \delta_i f_t + (t + t_i)\beta_{it} + P_{it} + v_{it}. \tag{4}$$

The household income residual has four components: (i) the effect of aggregate time varying factor on the households (the macro effect), $\delta_i f_t$. (ii) the effect of household's career trajectory; $\beta_{it}(t + t_i)$ captures the heterogeneity in the lifecycle trajectory (relative to other households in the sample). The slope β_{it} is a "experience premium" and a key determinant of the lifetime income. There are two types of idiosyncratic risks: (iii) The idiosyncratic permanent component $P_{it} = P_{it-1} + \zeta_{it}$, ζ_{it} has zero mean, variance $\sigma^2_{\zeta_t}$. Shocks in ζ before the sample period have a variance $\sigma^2_{\zeta_0}$. This assumption is useful in explaining the higher inequality of older cohort in 1997 and (iv) the idiosyncratic transitory component (or measurement errors) v_{it}. $v_{it} = \xi_{it} + \theta_1 \xi_{it-1}$ is MA(1), ξ_{it} has zero mean, variance $\sigma^2_{\xi_t}$.

This model nests some popular existing models. We assume

$$\begin{aligned}
&EP_{ij} = Ev_{it} = E\delta_i = E\beta_{it} = 0,\\
&cov(P_{ij}, v_{it}) = 0 \ for \ all \ j \ and \ t;\\
&cov(\delta_i, \beta_{it}) = cov(\delta_i, \beta_{i0});\\
&cov(\delta_i, P_{it}) = cov(\delta_i, \alpha_i).
\end{aligned} \tag{5}$$

$var(\alpha_i)$, $var(\beta_{i0})$, $cov(\delta_i, \alpha_i)$ and $cov(\delta_i, \beta_{i0})$ are unknown parameters.

From the assumptions of the idiosyncratic shocks,

$$var(P_{it}) = var(\alpha_i) + t_i\sigma^2_{\zeta_0} + \sum_{k=1}^{t}\sigma^2_{\zeta_k}, \tag{6}$$

$$\begin{aligned} cov(v_{it}, v_{t+j}) &= var(v_{it}) = \sigma^2_{\xi_t} + \theta^2\sigma^2_{\xi_{t-1}} \; for \; j = 0 \\ &= \theta\sigma^2_{\xi_t} \; for \; j = 1, \\ &= 0 \; for \; j > 1. \end{aligned} \tag{7}$$

The details of moment conditions of $\frac{1}{N}\mathbf{Y}\,\mathbf{Y}$ are in Appendix B.

We will estimate five models in the framework of the equation (4). The first two are the familiar RIP and HIP models.

1. The **RIP** model.

The RIP is a special case with $f_t = 0$ or $\delta_i = 0$, and $\beta_{it} = 0$.

In this case y_{it} is decomposed into a permanent component P and a mean-reverting transitory component v

$$y_{it} = P_{it} + v_{it}. \tag{8}$$

The permanent component is

$$P_{it} = P_{it-1} + \zeta_{it}, \tag{9}$$

where ζ is a permanent income shock with variance $\sigma^2_{i,t,\zeta}$, and the transitory component is $v_{it} = \sum_{j=0}^{p}\theta_j\xi_{i,t-j}$, $\xi_{i,t}$ is a transitory income shock, with variance $\sigma^2_{t,\xi}$. For this study we let $\theta_j = 1$ for $j = 0$, $\theta_j = \theta$ for $j = 1$, and $\theta_j = 0$ for $j > 1$. The permanent component P_{it} is assumed to be independent of the transitory shocks v_{it}. There are $2 + 2T$ unknown parameters $\boldsymbol{\theta} = \{\theta, \sigma^2_\alpha, \sigma^2_{\zeta_t}, \sigma^2_{\xi_t}\}$.

2. The **HIP Model**.

With $f_t = 1$, $P_{it} = 0$, and $\beta_{it} = \beta_i$ in the equation (4) we have

$$y_{it} = \delta_i + (t + t_i)\beta_i + v_{it},$$

where the slope β_i is the time-invariant experience premium and $\mathsf{E}(\beta_i) = 0$. There are $4 + T$ unknown parameters $\boldsymbol{\theta} = \{\sigma^2_\delta, \sigma^2_\beta, cov(\beta_i, \delta_i), \theta, \sigma^2_{\xi_t}\}$.

3. The **Baker/Solon Model**.

Baker and Solon (2003) combine the RIP model with the HIP, and introduce a macro factor, "p_t":

$$y_{it} = p_t(\delta_i + (t + t_i)\beta_i + P_{it}) + v_{it},$$

where $P_{it} = P_{it-1} + \zeta_{it}, v_{it} = \xi_{it} + \theta_1 \xi_{it-1}$. In the Baker/Solon model the macro factor amplifies permanent and lifetime income profile proportionally. There are $7 + 3T$ unknown parameters $\boldsymbol{\theta} = \{\sigma^2_\delta, \sigma^2_\alpha, \sigma^2_\beta, cov(\alpha_i, \beta_i), cov(\beta_i, \delta_i), cov(\alpha_i, \delta_i), \theta, \sigma^2_{\zeta_t}, \sigma^2_{\xi_t}, p_t\}$. Here the variances $\sigma^2_{\zeta_t}, \sigma^2_{\xi_t}$ can not be separately identified from p_t using moments $cov(y_{it}, y_{it+j})$ (See Appendix B).

We now introduce two new models that generalize the RIP and HIP models.

4. A **Generalized RIP (G-RIP) Model**.

The conventional identification assumptions of the RIP yields counterfactual predictions. We consider a generalized version of the RIP: We introduce a macro effect m_{it} to the permanent component: $m_{it} = f_t P_{it-1}$. If f_t is negative (positive) in a recession then high income households who have $P_{it-1} > 0$ suffer (gain) relative to low income households.

We assume $cov(P_{it}, \zeta_{it+j}) = 0$ for $j > 0$ and $cov(P_{it}, \xi_{ij}) = 0$, $cov(\xi_{it}, \zeta_{it+j}) = 0$ for all t and j and $cov(\zeta_{it}, \zeta_{it+j}) = cov(\xi_{it}, \xi_{it+j}) = 0$ for all $t = j$.

For now we let v_{it} to be iid. The permanent component is now

$$P_{it} = (1 + f_t)P_{it-1} + \zeta_{it}, \tag{10}$$

with $f_t > -1$.

The model can be identified through the following moment conditions:

$$\begin{aligned}
&cov(y_{it}, y_{it+j}) \\
&= cov(P_i, P_{it+j}) + cov(P_{it}, v_{it+j}) + cov(P_{it+j}, v_{it}) + cov(v_{it+j}, v_{it}) \\
&= var(P_{it}) \prod_{k=1}^{j} (1 + f_{t+k})
\end{aligned} \tag{11}$$

The ratio in (11) yields

$$cov(y_{it}, y_{it+j})/cov(y_{it}, y_{it+j-1}) = 1 + f_{t+j} \tag{12}$$

From (10)

$$var(P_{it}) = var(P_{i0})(\prod_{j=1}^{t}(1 + f_j))^2 + \sum_{j=1}^{t-1}(\prod_{k=1+j}^{t}(1 + f_k))^2\sigma^2_{\zeta_j} + \sigma^2_{\zeta_t}. \tag{13}$$

We also have from the identifying restrictions

$$var(y_{it}) = var(P_{it}) + var(v_{it}). \tag{14}$$

The above conditions solve for f_t, $var(v_{it})$, and $\sigma_{\zeta_t}, var(P_{i0})$.

Lastly we consider a generalization of the HIP model.

5. **Stochastic Heterogeneous Income Profile: (SHIP Model)**.

In the equation (4) β_{it} has time varying persistence with $\mathsf{E}(\beta_{it}) = 0$:

$$\beta_{it} = \rho_t\beta_{it-1} + \omega_{it}, \tag{15}$$

where ρ_t is a persistent parameter and ω_{it} is innovations to the experience premium and correlated with a permanent shock ζ_{it}. There are $7+6T$ unknown parameters $\boldsymbol{\theta} = \{\sigma^2_\delta, \sigma^2_\alpha, \sigma^2_{\beta_0}, \rho_t, cov(\delta_i, \alpha_i), cov(\delta_i, \beta_{i0}), cov(\alpha_i, \beta_{i0}), \theta, \sigma^2_{\zeta_t}, \sigma^2_{\xi_t}, \sigma^2_{\omega_t}, cov(\omega_{it}, \zeta_{it}), f_t\}$ (See Appendix B for moment conditions).

4 Estimates of Competing Income Models

We use the generalized method of moments (GMM) to estimate unknown parameters in the 5 competing income models. Table 4 presents estimation results of the competing income models including the three existing models (i.e., the RIP, HIP, and Baker/Solon models). As we discuss the RIP and HIP models based on the stylized facts in the section 2.5, the estimates for the both models are reasonable. First, permanent income shocks are small relative to transitory income shocks in the RIP model and variance of household-specific experience premium in the HIP model is small, which makes $var(y)$ flat until 2007 and labor income mobility relatively persistent. Second, in the HIP model the household-specific intercept (δ_i) is negatively correlated with the household-specific experience premium (β_i), which results in a negative correlation between income growth and initial income, $corr(y_t - y_3, y_1)$.

4.1 Income inequality

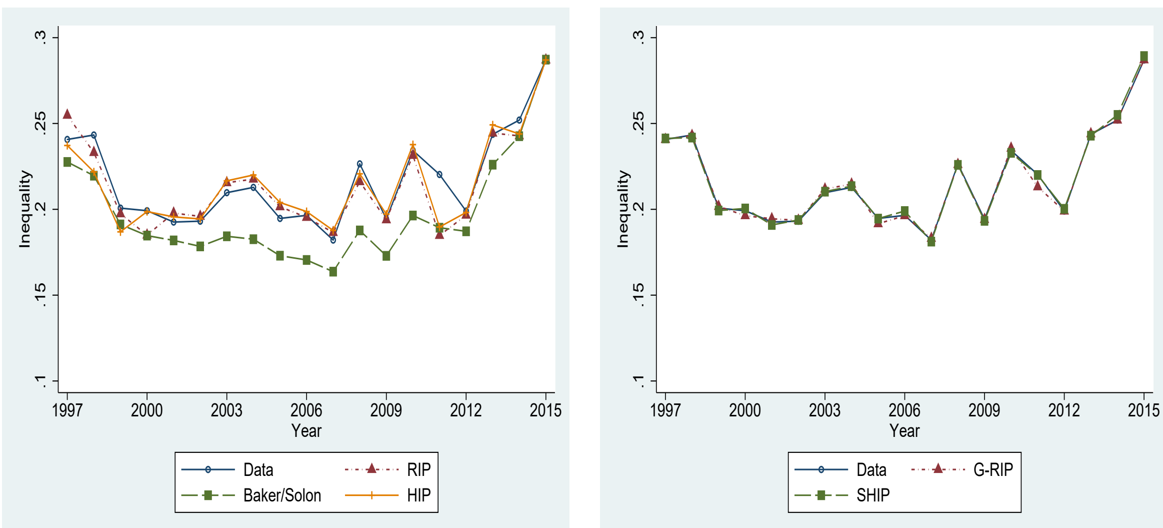
We now examine whether the income process models explain the second moments of income level. Diagonal elements of the covariance matrix, $\frac{1}{N}\mathbf{Y}\,\mathbf{Y}$ represent cross-sectional variance and Figure 6 displays the fitted inequality based on the estimates of Table 4. Comparing Figure 6(a) and 6(b), the G-RIP and SHIP models considering heterogeneous macro effects better fit than the existing models for inequality. The inequality implied by the Baker/Solon model considering the macro factor proportional to permanent component and lifetime income profile is biased downwards during the whole sample period.

Table 4. Estimation results

	RIP		HIP	Baker/Solon			G-RIP			SHIP					
Year	$\sigma^2_{\zeta_t}$	$\sigma^2_{\xi_t}$	$\sigma^2_{\xi_t}$	p_t	$\sigma^2_{\zeta_t}$	$\sigma^2_{\xi_t}$	f_t	$\sigma^2_{\zeta_t}$	$\sigma^2_{\xi_t}$	f_t	ρ	$corr(\omega,\zeta)$	σ_{ω_t}	$\sigma^2_{\zeta_t}$	$\sigma^2_{\xi_t}$
1997	0.0275	0.0915	0.0599	1.0260	0.0352	0.1018	0.0109	0.0008	0.0995	1.2811	0.1220	-0.1206	0.0203	0.0002	0.0450
1998	0.0000	0.1457	0.0992	1.0308	0.0017	0.0998	0.0998	0.0008	0.0977	1.1866	0.5297	-0.1070	0.0188	0.0001	0.0526
1999	0.0000	0.1007	0.0657	1.0092	0.0053	0.0789	-0.2256	0.0185	0.0956	1.0209	0.4844	-0.0905	0.0125	0.0000	0.0718
2000	0.0013	0.0948	0.0883	1.0097	0.0045	0.0753	-0.1350	0.0174	0.0993	0.9310	0.6498	-0.0366	0.0023	0.0134	0.1038
2001	0.0104	0.0982	0.0864	1.0107	0.0114	0.0673	0.0426	0.0139	0.0752	1.0658	0.7829	-0.0912	0.0087	0.0000	0.0646
2002	0.0000	0.0957	0.0889	1.0043	0.0062	0.0650	-0.0332	0.0058	0.0765	0.9491	1.0321	0.0599	0.0054	0.0049	0.0569
2003	0.0095	0.1062	0.1128	1.0175	0.0027	0.0706	0.0160	0.0196	0.0713	1.0401	0.7548	-0.0584	0.0134	0.0001	0.0170
2004	0.0045	0.1020	0.1134	1.0151	0.0030	0.0710	-0.0927	0.0222	0.0767	1.0615	0.5596	-0.0431	0.0098	0.0047	0.0449
2005	0.0000	0.0866	0.0969	1.0129	0.0081	0.0576	-0.0542	0.0000	0.0680	1.0165	0.7730	-0.1578	0.0086	0.0003	0.0245
2006	0.0000	0.0831	0.0923	1.0122	0.0018	0.0570	0.0153	0.0079	0.0612	1.1308	0.5950	-0.1357	0.0042	0.0001	0.0573
2007	0.0056	0.0693	0.0789	1.0099	0.0047	0.0486	-0.0619	0.0208	0.0434	1.1332	0.6814	-0.2072	0.0074	0.0024	0.0300
2008	0.0055	0.0958	0.1096	1.0069	0.0108	0.0645	-0.0862	0.0226	0.0869	0.9768	0.9535	-0.2172	0.0094	0.0071	0.0429
2009	0.0137	0.0553	0.0759	1.0041	0.0046	0.0470	-0.0462	0.0242	0.0427	1.0649	0.4728	-0.1492	0.0068	0.0089	0.0378
2010	0.0081	0.0917	0.1145	1.0054	0.0083	0.0627	-0.0602	0.0225	0.0798	1.0035	0.8234	-0.0843	0.0067	0.0004	0.0668
2011	0.0000	0.0388	0.0527	0.9947	0.0050	0.0534	-0.0716	0.0095	0.0688	0.8893	0.9220	-0.0958	0.0046	0.0000	0.0568
2012	0.0113	0.0484	0.0610	0.9812	0.0192	0.0361	-0.0589	0.0250	0.0462	0.9377	0.5267	-0.0869	0.0057	0.0023	0.0575
2013	0.0182	0.0764	0.1001	0.9850	0.0298	0.0439	0.0050	0.0344	0.0556	0.8878	1.2316	-0.0915	0.0045	0.0000	0.0582
2014	0.0000	0.0699	0.0770	0.9731	0.0310	0.0337	-0.0291	0.0104	0.0638	0.8154	1.0218	-0.0972	0.0032	0.0002	0.0572
2015	0.0476	0.0677	0.1103	0.9567	0.0591	0.0291	-0.0812	0.0540	0.0740	0.7396	0.8393	-0.1000	0.0105	0.0001	0.0188
θ	0.4150		0.3851		0.1153			0.1082				0.0240			
σ^2_α	0.0411				0.1406			0.1152				0.0008			
σ^2_β			0.0006		0.0003							0.0148			
σ^2_δ			0.3334		0.3406							0.0608			
$cov(\alpha,\beta)$					-0.0063							-0.0009			
$cov(\alpha,\delta)$					-0.1187							-0.0003			
$cov(\beta,\delta)$			-0.0129		-0.0023							-0.0060			

Under the assumptions of the RIP and HIP models, inequality tends to increase over time among the same cohort. We find income inequality depends on the state of the economy. Inequality tends to rise during a recession and to fall during a recovery. It is difficult to match this pattern by varying the variance of permanent shocks in the RIP model. The variance of permanent shocks should drop to explain the decrease in inequality during a recovery. With deterministic component driven by time-constant experience profile in the HIP model, variance of experience premium is rising over time, making it hard to explain the decrease in inequality during a recovery.

Figure 6. Fitted income inequality



(a) Existing models

(b) New income models

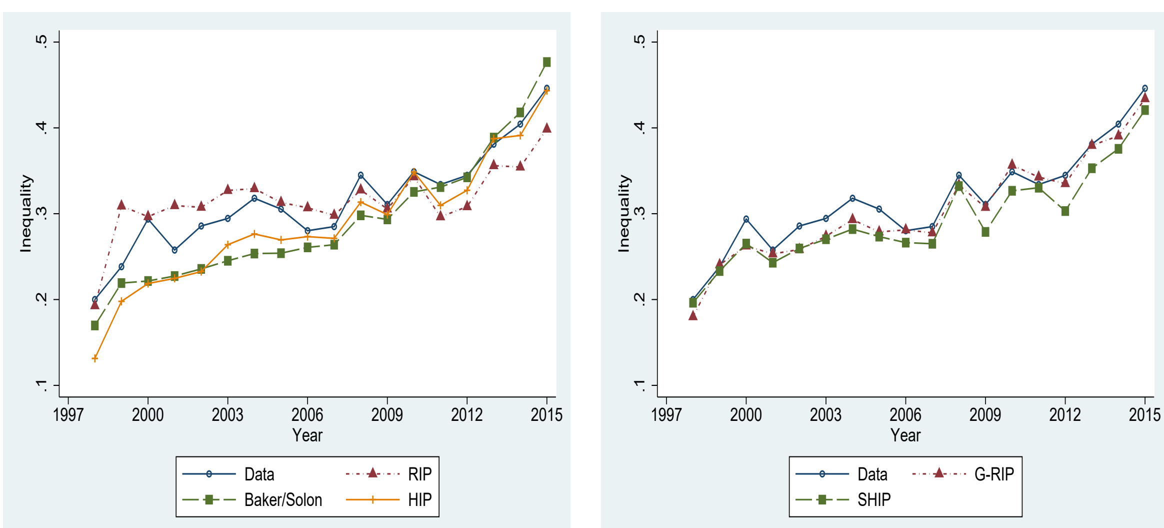
In the G-RIP and SHIP models with heterogeneous effect of macro shocks, a recession (recovery) is not necessarily associated with increase (decrease) in variance of permanent shocks. For example, in the G-RIP model, if macro effect parameter, $f_t > 0$ in a recession, low income households who have $P_{it-1} < 0$ suffer more relative to high income households, leading to an increase in income inequality. In the SHIP model, if macro effect parameter, $f_t > 1$ in a recession, a direct household-specific impact of the macroeconomy is stronger, which can also lead to an increase in income inequality. An interesting finding is that an increase in income inequality during the Asian Financial Crisis of 1998 and during the Global

Financial Crisis of 2008 appears to work through different channels. Specifically, based on the estimation results of the G-RIP model, a direct household-specific impact of macroeconomy dominates idiosyncratic permanent shock (i.e., variance of permanent shock is small) during the Asian Financial crisis of 1998. However, during the Global Financial Crisis of 2008, idiosyncratic permanent/transitory shocks dominate a direct household-specific impact of macroeconomy (i.e., variance of permanent(transitory) shocks is relatively high). We will examine which component derives change in inequality in the section 4.5.

4.2 Other moments of the covariance matrix

We analyze off-diagonal elements as well as diagonal elements by using a long panel data of KLIPS. From Figure 7, all models can explain the divergence of variance of cumulative income growth (i.e., var$(y_t - y_1)$). In the RIP model, inequality of cumulative income growth are proportional to cumulative variance of permanent shocks. In the HIP model, inequality of cumulative income growth are proportional to t squared. The G-RIP and SHIP models better fit than the existing models for inequality for cumulative income growth.

Figure 7. Fitted inequality for cumulative income growth

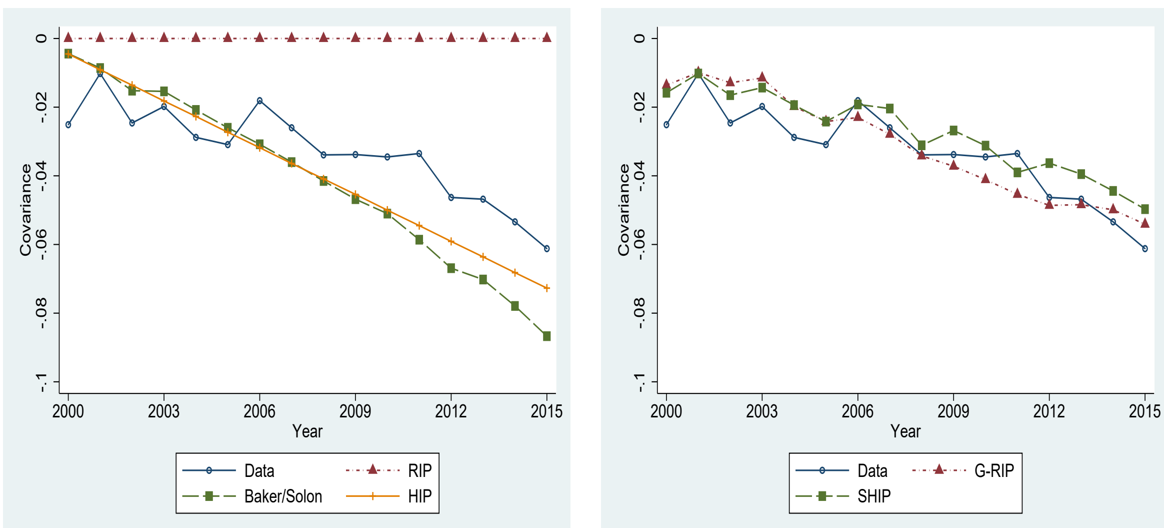


(a) Existing models

(b) New income models

The G-RIP and SHIP models better fit than the existing models as well for $\text{cov}(y_t - y_3, y_1)$ (See Figure 8). The RIP model predicts the correlation $\text{cov}(y_t - y_3, y_1)$ is 0. Under the assumption of the RIP model, as transitory component assumes an MA(1) process, income at time t is not correlated with income ahead two years. Under the HIP model, although the correlation between household-specific experience premium and household-specific intercept can be negative, the linear relationship cannot explain the real covariance. The G-RIP and SHIP models fit the negative correlation better. The future earnings-initial income correlation differs by households, the better fit by the G-RIP and SHIP models suggests that the aggregate shocks may influence this relationship.

Figure 8. Fitted $\text{cov}(y_t - y_3, y_1)$



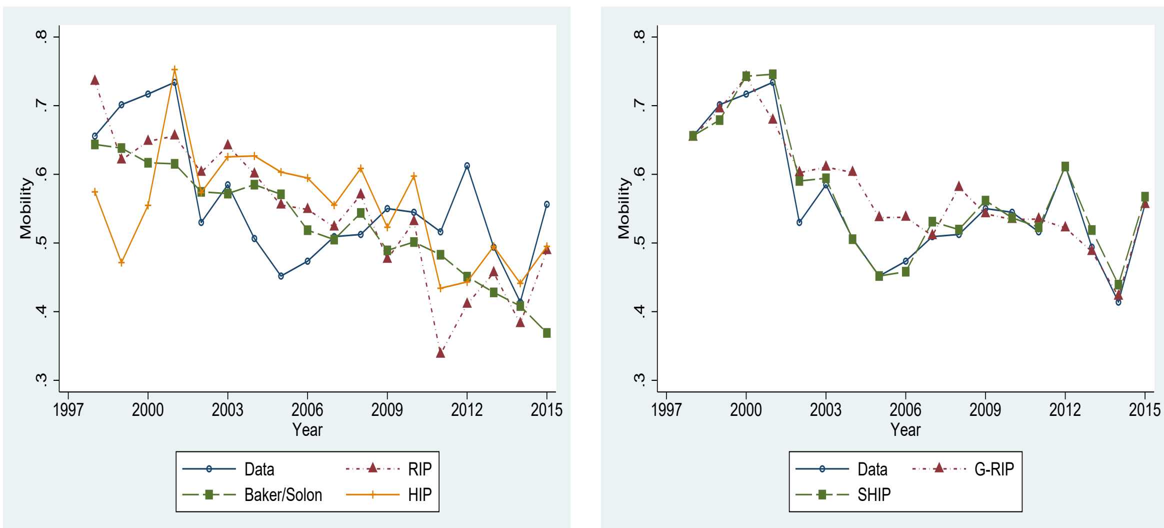
(a) Existing models

(b) New income models

4.3 Mobility

Mobility concerns how the relative standings of their own income change during the two consecutive years. Inequality and mobility may move in the same direction but it is possible that they move the in opposition direction. Under the assumption of the RIP model, inequality and mobility move in the same direction. So, inequality and mobility increase during a recession and decrease during a recovery. Under the assumption of the HIP model, mobility depend on the variance of household-specific experience premium. If the variance of household-specific experience premium is big enough, inequality increase and mobility decrease over time and both do not depend on economic state. If the variance of household-specific experience premium is negligible, mobility depends on transitory shocks. Based on the estimates of the HIP model, the variance of household-specific experience premium is small (about .0006), which results in persistent mobility.

Figure 9. Fitted mobility



(a) Existing models (b) New income models

We use $1 - \hat{\rho}^2$ to measure mobility, where $\hat{\rho}$ is the correlation between y_{t-1} and y_t. To better fit mobility, eventually, the income model should explain off-diagonal elements as well as diagonal elements of the covariance matrix, $\frac{1}{N}\mathbf{Y}\,\mathbf{Y}$ because the correlation $\text{corr}(y_{t-1}, y_t)$ is determined by $\text{var}(y_{t-1})$, $\text{var}(y_t)$, and $\text{cov}(y_{t-1}, y_t)$. Therefore, the G-RIP and SHIP models better fit than the existing models for mobility.

4.4 Model comparison

Figures 6 to 9 show that the G-RIP and SHIP models better fit than the existing models of the RIP, HIP, and Baker/Solon models for $\text{var}(y_t)$,$\text{var}(y_t - y_1)$, $\text{cov}(y_t - y_3, y_1)$, and mobility. However, by including heterogeneous macro effects the G-RIP and SHIP models contain more parameters in than the RIP and HIP models, and are expected to fit better. A comparison of the generalized models with nested models should be based on criteria that account for both model fit and the number of parameters in the model.

Andrews and Lu (2001) suggest the moment selection criteria for GMM estimation. The selection criteria, MMSC-BIC, MMSC-AIC, and MMSC-HQIC developed by Andrews and Lu (2001) are based on J statistics. J statistics can be computed by the number of parameters and of moments reflects the trade-off between the "model fit" and the "number of parameters". We apply model selection criteria, MMSC-BIC, MMSC-AIC, and MMSC-HQIC to the 5 competing income models and the model with minimum of each model criterion is deemed the best model. Table 5 shows that the G-RIP and SHIP models are better than three existing models. Specifically, the G-RIP model is the best based on MMSC-BIC and MMSC-HQIC and the SHIP is the best based on MMSC-AIC.

Table 5. Model comparison

Model	Number of parameters	MMSC-BIC	MMSC-AIC	MMSC-HQIC
RIP	40	1036.91	2123.69	1422.84
HIP	23	1110.38	2320.33	1540.05
Baker/Solon	64	-492.08	420.82	-167.89
G-RIP	59	-552.46	396.66	-215.41
SHIP	121	-242.18	257.74	-64.65

Note: MMSC-BIC(b,c)=$J_n(b,c) - (|c| - |b|)\ln n$, MMSC-AIC(b,c)=$J_n(b,c) - 2(|c| - |b|)$, and MMSC-HQIC(b,c)=$J_n(b,c) - Q(|c| - |b|)\ln(\ln n)$.Where, $J_n(b,c)$ is the J statistics developed by Andrews and Lu (See pp. 10-11, 2001), b is number of parameters, c is number of moments, n is the sample size, and Q is a number greater than 2. We select $Q = 3$ to compute MMSC-HQIC(b,c).

4.5 The compositions of earning inequality

Understanding the source of income inequality is useful for making of economic policy. If income inequality is mainly due to idiosyncratic transitory shocks then better social insurance and access to credit market for short term income shortfalls can potentially improve social welfare. If income inequality is mainly driven by permanent component then reduction of income inequality requires public policies that address long term factors that improve career trajectories, such as education and health. One question of interest is whether the main source of income inequality differs in groups of households of different education levels.

As noted in Section 3.1, unexplained earning inequality may be driven by macro effect, the effect of household's career trajectory (a deterministic career effect), permanent component, and transitory component. We decompose change in inequality into change in inequality driven by each component for the competing income models (See Table 6). The RIP and G-RIP models have permanent and transitory components. In the G-RIP model the heterogeneous macro effect is correlated with permanent shocks. The trend of inequality in the G-RIP model is driven by both permanent shocks and macro effects. The HIP model has deterministic career effect and transitory component. As the Baker/Solon and SHIP models nest the RIP model with the HIP model, both models have permanent component and deterministic career effect. However, permanent component and deterministic career effect are proportionally affected by macro effect in the Baker/Solon model. In the SHIP model, we group deterministic career effect with macro effect because household-specific intercept and experience premium in career trajectory are affected by aggregate factors.

Under the conventional assumptions without heterogeneous macro effect, the variance of permanent component and the variance driven by the deterministic career effect are increasing over time. So, the RIP and HIP models do not explain the high inequality during the East Asian Crisis of 1998, the beginning of the sample period. In the G-RIP model the high inequality during the East Asian Crisis of 1998 is mainly driven by the permanent component including the heterogeneous macro effect. In the SHIP model it is due to the transitory component. However, the increase in inequality during the Global Financial Crisis of 2008 is mainly driven by the transitory component at the G-RIP model and by both heterogeneous response to aggregate shocks and transitory component at the SHIP model. Even with similar fit of data the explanation of income inequality differs by the income model.

Table 6. The compositions of earning inequality

	RIP			HIP			Baker/Solon				G-RIP			SHIP			
Year	Δvar(y)	C3	C4	Δvar(y)	C2	C4	Δvar(y)	C2	C3	C4	Δvar(y)	C3	C4	Δvar(y)	C1	C3	C4
1998	-0.0217	0.0000	-0.0217	-0.0154	-0.0084	-0.0069	-0.0080	-0.0009	-0.0051	-0.0020	0.0023	0.0041	-0.0018	0.0006	-0.0075	0.0007	0.0074
1999	-0.0357	0.0000	-0.0357	-0.0349	-0.0072	-0.0277	-0.0284	-0.0078	0.0003	-0.0209	-0.0413	-0.0392	-0.0021	-0.0425	-0.0623	0.0006	0.0192
2000	-0.0124	0.0013	-0.0137	0.0117	-0.0059	0.0176	-0.0065	-0.0007	-0.0018	-0.0039	-0.0053	-0.0090	0.0037	0.0012	-0.0443	0.0135	0.0320
2001	0.0128	0.0104	0.0024	-0.0031	-0.0046	0.0015	-0.0028	0.0000	0.0052	-0.0080	-0.0018	0.0222	-0.0241	-0.0097	0.0295	0.0000	-0.0392
2002	-0.0019	0.0000	-0.0019	-0.0012	-0.0033	0.0022	-0.0035	-0.0014	0.0003	-0.0024	-0.0009	-0.0019	0.0010	0.0029	0.0053	0.0053	-0.0077
2003	0.0196	0.0095	0.0101	0.0222	-0.0021	0.0243	0.0059	0.0048	-0.0044	0.0056	0.0182	0.0234	-0.0052	0.0165	0.0564	0.0000	-0.0399
2004	0.0021	0.0045	-0.0024	0.0034	-0.0008	0.0042	-0.0017	0.0011	-0.0032	0.0005	0.0029	-0.0025	0.0053	0.0032	-0.0290	0.0043	0.0279
2005	-0.0161	0.0000	-0.0161	-0.0160	0.0005	-0.0164	-0.0096	0.0018	0.0020	-0.0134	-0.0231	-0.0145	-0.0086	-0.0189	0.0016	0.0000	-0.0204
2006	-0.0062	0.0000	-0.0062	-0.0053	0.0018	-0.0071	-0.0025	0.0028	-0.0045	-0.0008	0.0048	0.0117	-0.0069	0.0044	-0.0288	0.0004	0.0328
2007	-0.0088	0.0056	-0.0144	-0.0110	0.0030	-0.0141	-0.0069	0.0030	-0.0014	-0.0084	-0.0132	0.0047	-0.0179	-0.0178	0.0086	0.0009	-0.0273
2008	0.0296	0.0055	0.0241	0.0330	0.0043	0.0287	0.0240	0.0034	0.0048	0.0158	0.0430	-0.0003	0.0433	0.0447	0.0287	0.0030	0.0129
2009	-0.0222	0.0137	-0.0359	-0.0235	0.0056	-0.0291	-0.0147	0.0041	-0.0015	-0.0173	-0.0320	0.0117	-0.0437	-0.0326	-0.0372	0.0098	-0.0051
2010	0.0375	0.0081	0.0294	0.0405	0.0069	0.0336	0.0234	0.0060	0.0019	0.0155	0.0415	0.0049	0.0366	0.0395	0.0096	0.0009	0.0290
2011	-0.0466	0.0000	-0.0466	-0.0480	0.0081	-0.0561	-0.0071	0.0026	-0.0005	-0.0091	-0.0225	-0.0119	-0.0106	-0.0128	-0.0030	0.0002	-0.0100
2012	0.0118	0.0113	0.0005	0.0085	0.0094	-0.0009	-0.0021	0.0020	0.0134	-0.0174	-0.0141	0.0086	-0.0227	-0.0197	-0.0243	0.0039	0.0007
2013	0.0479	0.0182	0.0297	0.0510	0.0107	0.0403	0.0388	0.0086	0.0227	0.0076	0.0451	0.0359	0.0091	0.0426	0.0428	-0.0009	0.0007
2014	-0.0017	0.0000	-0.0017	-0.0053	0.0120	-0.0172	0.0166	0.0033	0.0234	-0.0101	0.0079	-0.0004	0.0083	0.0120	0.0132	-0.0002	-0.0010
2015	0.0443	0.0476	-0.0033	0.0431	0.0132	0.0299	0.0446	0.0019	0.0475	-0.0047	0.0351	0.0248	0.0103	0.0343	0.0723	0.0004	-0.0384

Note: C1, C2, C3, and C4 represent change in inequality driven by macro effect, deterministic career effect, permanent component, and transitory component, respectively.

Tables C-1 to C-5 in the Appendix C present the composition of earning inequality by education level. Interestingly, inequality of the group with low education increased but inequality of the high education group decreased during the East Asian Crisis of 1998. However, inequality of both groups increased and inequality of the high educated group increased more than the low educated group during the Global Financial Crisis of 2008. Also, inequality of both groups increased and inequality of the high education group increased more than the low education group at the end of the sample period. The trend of inequality for the low education households is mainly driven by the macro effect or the permanent component with the heterogeneous response to the macroeconomy since 2013. The trend of inequality for the high education group is mainly driven by both the macro effect (or permanent component with the heterogeneous macro effect) and the transitory component in 2013 and 2015. The main reason why inequality of both groups increases at the end of the sample period is likely retirement by some head of households.

5 Concluding Remarks

This paper contributes to the literature of household income inequality by presenting new empirical observations and new models of income process that explain the empirical observations.

Using a long panel of Korean households of KLIPS we examine the trend of inequality and mobility of household earnings post the Asian Financial Crisis. We find income inequality is correlated with the macroeconomy and households with initial low income have higher future income. We also find persistently high income mobility.

We analyze the performance of commonly employed models of income process, the RIP and HIP, and a combination of the RIP and HIP model by Baker/Solon. We also introduce new generalizations of the RIP and HIP, the generalized RIP (G-RIP) and stochastic heterogeneous income profile (SHIP). Our G-RIP and SHIP models fit the KLIPS data better according to the model selection criteria of Andrews and Lu (2001). The G-RIP and SHIP models shed new light on the drivers of income equality. We find that a significant source of income inequality is heterogeneous responses by households to aggregate shocks. This means that a large recession not only reduces the average income but also widens inequality because some households do persistently worse over business cycles than the population in general. We also find that the sources of income inequality differ by the education level of the head of households. Further analysis of the KLIPS data will help reevaluating public policies on income stabilization and social insurance implemented after the Asian Financial Crisis.

A Appendix A: Sample Selection

Table A-1. Response and nonresponse rates for original 5,000 households in KLIPS

Year	Remaning in sample	Attritors	Cumulative rate	In from attritors
1997	5,000			
1998	4,378	622	12.44%	
1999	4,044	483	19.12%	149
2000	3,866	430	22.68%	252
2001	3,798	290	24.04%	222
2002	3,862	271	22.76%	335
2003	3,862	223	22.76%	223
2004	3,822	218	23.56%	178
2005	3,820	185	23.60%	183
2006	3,775	163	24.50%	118
2007	3,710	181	25.80%	116
2008	3,658	193	26.84%	141
2009	3,607	191	27.86%	140
2010	3,528	167	29.44%	88
2011	3,517	98	29.66%	87
2012	3,472	118	30.56%	73
2013	3,451	68	30.98%	47
2014	3,421	81	31.58%	51
2015	3,393	68	32.14%	40

Table A-2. Drop-out sample

Sample selection method	Number of households
One-person household	560
No-head or no-spouse household	675
Separate household	351
Head age (age < 25 or age > 65)	736
Labor income (zero or missing report)	650
Not participation on whole survey	1483
Total	4455

A-1 Test for bias in sample attrition

Let an interesting model assume the linear parametric model:

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t,\ y_t \text{ observed if } A_t = 0 \tag{16}$$

where x_t is an independent variable (a vector is possible), ϵ_t is a random variable with zero mean, and A_t is attrition indicator (a.g., $A_t = 1$ if an observation, y_t is missing at time t because of attrition). As the density of primary interest is $f(y_t|x_t)$, if $f(y_t|x_t) = f(y_t|x_t, A_t = 0)$ which is conditional on x_t and $A_t = 0$, attrition bias occurs. Let the attrition function be

$$A_t^* = \delta_0 + \delta_1 x_t + \delta_2 z_t + v_t. \tag{17}$$

$$\begin{aligned} A_t &= 1 \text{ if } A_t^* \geq 0 \\ &= 0 \text{ if } A_t^* < 0, \end{aligned} \tag{18}$$

where A_t^* is a latent index and z_t is an auxiliary variable which is not included in explanatory variable, x_t. Attrition bias is attributed by i) selection on observables ($Pr(A_t = 0|y_t, x_t, z_t) = Pr(A_t = 0|x_t, z_t)$) and ii) selection on unobservables ($Pr(A_t = 0|y_t, x_t, z_t) = Pr(A_t = 0|x_t, z_t)$). According to Fitzgerald, Gottschalk, and Moffitt (1998), the BGLW test developed by Becketti, Gould, Lillard, and Welch (1988) can be used for attrition bias caused by selection on observables and another data set with much less attrition and comparable to KLIPS are used for test on attrition bias caused by selection on unobservables. Here, we test for attrition bias due to selection on observables.

An auxiliary variable, z_t should be observed for both attritors as well as non-attritors and do not affect attrition, A_t (Fitzgerald, Gottschalk, and Moffitt (1998), Alderman, Behrman, Kohler, Maluccio, and Watkins (2001)). Fitzgerald, Gottschalk, and Moffitt (1998) suggest labor income of initial wave as z_t for the test on attrition bias and we use labor income in 1997 for z_t. To find the effect of household's characteristics on attrition rate, we compare simple statistics about household's characteristics between "Always in" to represent non-attritors group and "Ever out except death" to represent attritors group in Table A-3. According to Table A-3, "Always in" and "Ever out except death" groups have statistical differences in characteristics such as age, education, number of children, and residential city under the 10% of significant level. High income households tend to be attritors but not statistically significant at the 10% level.

Table A-3. 1997 characteristics by attrition: Age 25-64

	Always in	Ever out	Ever out(died)	Ever out(not died)
Age	45.6	42.1	53.2	41.1
Male heads	88.6%	87.9%	73.0%	89.4%
Education:				
<12	41.6%	27.3%	63.5%	23.8%
12	39.0%	37.9%	29.7%	38.7%
13-16	16.8%	30.6%	4.1%	33.2%
>16	2.5%	4.2%	2.7%	4.3%
Own home	60.3%	46.2%	56.8%	45.1%
Annual household's disposable income for those w/ income > 0:				
Mean	3,088	3,107	1,518	3,253
10 percentile	859	619	25	999
50 percentile	2,797	2,597	1,199	2,797
90 percentile	5,694	5,994	3,996	5,994
Var. of log income	2.517	4.144	10.270	3.069
Annual household's consumption for those w/ consumption > 0:				
Mean	2,079	2,083	1,400	2,149
10 percentile	799	599	300	799
50 percentile	1,998	1,798	999	1,998
90 percentile	3,796	3,996	2,997	3,996
Var. of log consumption	1.092	1.355	2.157	1.192
Num. of children	1.1	1.1	0.3	1.2
Bigcity	55.8%	60.9%	48.6%	62.1%
Sample size	1,931	836	74	762

Note: represents significant difference from"Always in" at 10% level and the unit of household's disposable income and consumption is ten thousands Korean Won.

Table A-4 represents whether labor income of initial wave affect an attrition. Based on the results estimated by the probit model, household labor income in 1997 does not statistically affect an attrition regardless of conditioning on other observable variables such as age, education, family size, residential city and so on. This implies that labor income in 1997 is suitable for an auxiliary variable.

Table A-4. Ever-Out attrition probits for household labor income

	Model 1		Model 2		Model 3	
	Coeff.	$\partial p/\partial x$	Coeff.	$\partial p/\partial x$	Coeff.	$\partial p/\partial x$
Intercept	-0.897		-0.897		-0.857	
Income$^{a)}$	0.014	0.004	0.018	0.005	-0.019	-0.005
No income	-0.094	-0.024	-0.103	-0.026	-0.043	-0.011
Squared income$^{b)}$			-0.022	-0.006	0.070	0.017
Age					0.016	0.004
Squared age$^{a)}$					-0.342	-0.085
Male head					-0.029	-0.007
Own house					-0.139	-0.035
Education<12					-0.114	-0.028
Education>13					0.270	0.067
Bigcity					0.025	0.006
Pseudo R^2	0.0006		0.0006		0.0323	
Log Like.	-2024.27		-2024.23		-1960.10	

Note: Ever-out attritors due to death are excluded and represents significance at 10% level. Also, $^{a)}$ and $^{b)}$ mean the coefficient and marginal effect are multiplied by 10^3 and 10^8, respectively.

We implement BGWL test on whether effects of respondent's characteristics on auxiliary variable between total sample and nonattriting sample are different. If the test result is statistically insignificant, attrition bias caused by selection on observables is ignorable (Alderman, Behrman, Kohler, Maluccio, and Watkins (2001)). Comparing the effects of respondent's characteristics on an auxiliary variable, variables related to age only are significantly different at the 10% level between two groups. However, the magnitude of differences are not large, which means that the attrition on observables is not ignorable but the effects of attrition on the coefficients, β are not huge. In the case of the appearance of attrition bias, the residual, ϵ_t is larger than the residual under the unbiased estimators of β and the inequality of the residuals extracted from the equation (1) would be slightly overstated because of attrition bias.

Table A-5. 1997 log household labor income for BGLW test.

	Total	Always in	Difference
Intercept	5.422	4.927	0.495
Age	0.090	0.111	-0.021
Squared age$^{a)}$	-0.099	-0.122	0.023
Male head	0.388	0.453	-0.065
Own house	0.112	0.102	0.010
edu<12	-0.272	-0.291	0.019
edu>13	0.334	0.296	0.038
Bigcity	0.115	0.109	0.006
Sample size	3,857	1,729	
Adj. R^2	0.2135	0.2065	

Note: Ever-out attritors due to death are excluded and represents significance at 10% level. Also, $^{a)}$ means the coefficient is multiplied by 100.

B Appendix B: Moment Conditions

The elements of $\frac{1}{N}\mathbf{Y}\,\mathbf{Y}$ in the equation (4) can be computed by a function of unknown parameters:

$$\begin{aligned}
var(y_{it}) &= var(\delta_i f_t + (t+t_i)\beta_{it} + P_{it} + v_{it}) \\
&= f_t^2 var(\delta_i) + (t+t_i)^2 var(\beta_{it}) + var(P_{it}) + var(v_{it}) + 2f_t cov(\alpha_i, \delta_i) \\
&+ 2f_t(t+t_i)cov(\delta_i, \beta_{i0}) + 2(t+t_i)cov(\beta_{it}, P_{it}),
\end{aligned} \tag{19}$$

$$\begin{aligned}
&cov(y_{it}, y_{it+j}) \\
&= cov(\delta_i f_t + (t+t_i)\beta_{it} + P_{it} + v_{it},\ \delta_i f_{t+j} + (t+j+t_i)\beta_{it+j} + P_{it+j} + v_{it+j}) \\
&= \sigma_\delta^2 f_t f_{t+j} + (t+t_i)(t+j+t_i)cov(\beta_{it}, \beta_{it+j}) + var(P_{it}) + cov(v_{it}, v_{t+j}) \\
&+ cov(\delta_i, \beta_{i0})[f_t(t+j+t_i) + f_{t+j}(t+t_i)] + cov(\delta_i, \alpha_i)(f_t + f_{t+j}) \\
&+ (t+t_i)cov(\beta_{it}, P_{it+j}) + (t+j+t_i)cov(\beta_{it+j}, P_{it}).
\end{aligned} \tag{20}$$

B-1 Moment conditions of $\frac{1}{N}\mathbf{Y}\,\mathbf{Y}$

1. RIP Model

$$var(y_{it}) = var(P_{it} + v_{it}) = var(P_{it}) + var(v_{it}), \tag{21}$$

$$\begin{aligned}
cov(y_{it}, y_{it+j}) &= cov(P_{it} + v_{it}, P_{it+j} + v_{it+j}) \\
&= cov(P_{it}, P_{it+j}) + cov(v_{it}, v_{it+j}) \\
&= var(P_{it}) + cov(v_{it}, v_{it+j}).
\end{aligned} \tag{22}$$

2. HIP Model

$$\begin{aligned}
var(y_{it}) &= var(\delta_i + (t+t_i)\beta_i + v_{it}) \\
&= \sigma_\delta^2 + (t+t_i)^2\sigma_\beta^2 + 2(t+t_i)cov(\beta_i, \delta_i) + var(v_{it}),
\end{aligned} \tag{23}$$

$$
\begin{aligned}
&cov(y_{it}, y_{it+j}) \\
&= cov(\delta_i + (t+t_i)\beta_i + v_{it}, \delta_i + (t+j+t_i)\beta_i + v_{it+j}) \\
&= \sigma_\delta^2 + (t+t_i)(t+j+t_i)\sigma_\beta^2 + (2t+j+2t_i)cov(\beta_i, \delta_i) + cov(v_{it}, v_{it+j}).
\end{aligned} \tag{24}
$$

3. Baker/Solon Model

$$
\begin{aligned}
var(y_{it}) &= var(p_t(\delta_i + (t+t_i)\beta_i + P_{it}) + v_{it}) \\
&= p_t^2(\sigma_\delta^2 + (t+t_i)^2\sigma_\beta^2 + var(P_{it}) + 2(t+t_i)cov(\beta_i, \delta_i) + 2cov(\alpha_i, \delta_i) \\
&+ 2(t+t_i)cov(\alpha_i, \beta_i) + var(v_{it}),
\end{aligned} \tag{25}
$$

$$
\begin{aligned}
&cov(y_{it}, y_{it+j}) \\
&= cov(p_t(\delta_i + (t+t_i)\beta_i + P_{it}) + v_{it}, p_{t+j}(\delta_i + (t+j+t_i)\beta_i + P_{it+j}) + v_{it+j}) \\
&= p_t p_{t+j}[\sigma_\delta^2 + (t+t_i)(t+j+t_i)\sigma_\beta^2 + (2t+j+2t_i)cov(\beta_i, \delta_i) \\
&+ (2t+j+2t_i)cov(\alpha_i, \beta_i) + 2cov(\alpha_i, \delta_i) + var(P_{it})] + cov(v_{it}, v_{it+j}).
\end{aligned} \tag{26}
$$

4. G-RIP Model

$$
\begin{aligned}
var(y_{it}) &= var(P_{it} + v_{it}) = var(P_{it}) + var(v_{it}) \\
&= var(P_{i0})(\prod_{j=1}^{t}(1+f_j))^2 + \sum_{j=1}^{t-1}(\prod_{k=1+j}^{t}(1+f_k))^2\sigma_{\zeta_j}^2 + \sigma_{\zeta_t}^2 + var(v_{it}),
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
&cov(y_{it}, y_{it+j}) \\
&= cov(P_{it} + v_{it}, P_{it+1} + v_{it+j}) = cov(P_i, P_{it+j}) + cov(v_{it}, v_{it+j}) \\
&= var(P_{it})\prod_{k=1}^{j}(1+f_{t+k}) + cov(v_{it}, v_{it+j}).
\end{aligned} \tag{28}
$$

5. SHIP Model

$$
\begin{aligned}
var(y_{it}) &= var(\delta_i f_t + (t+t_i)\beta_{it} + P_{it} + v_{it}) \\
&= f_t^2 var(\delta_i) + (t+t_i)^2 var(\beta_{it}) + var(P_{it}) + var(v_{it}) + 2f_t(t+t_i)cov(\delta_i, \beta_{it}) \\
&+ 2f_t cov(\delta_i, P_{it}) + 2(t+t_i)cov(\beta_{it}, P_{it}) \\
&= f_t^2\sigma_\delta^2 + (t+t_i)^2((\prod_{k=1}^{t}\rho_k)^2\sigma_\beta^2 + \sum_{k=1}^{t-1}(\prod_{j=k+1}^{t}\rho_j)^2\sigma_{\omega_k}^2 + \sigma_{\omega_t}^2) + var(P_{it}) \\
&+ var(v_{it}) + 2f_t(t+t_i)\prod_{k=1}^{t}\rho_k cov(\delta_i, \beta_{i0}) + 2f_t cov(\delta_i, \alpha_i) \\
&+ 2(t+t_i)(\prod_{k=1}^{t}\rho_k cov(\beta_{i0}, \alpha_i) + \sum_{k=1}^{t-1}(\prod_{j=k+1}^{t}\rho_j cov(\omega_{ik}, \zeta_{ik})) + cov(\omega_{it}, \zeta_{it})),
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
&cov(y_{it}, y_{it+j}) \\
&= cov(f_t\delta_i + (t+t_i)\beta_{it} + P_{it} + v_{it}, f_{t+j}\delta_i + (t+j+t_i)\beta_{it+j} + P_{it+j} + v_{it+j}) \\
&= f_t f_{t+j}\sigma_\delta^2 + f_t(t+j+t_i)cov(\delta_i, \beta_{it+j}) + f_t cov(\delta_i, P_{it+j}) + f_{t+j}(t+t_i)cov(\delta_i, \beta_{it}) \\
&+ (t+t_i)(t+j+t_i)cov(\beta_{it}, \beta_{it+j}) + (t+t_i)cov(\beta_{it}, P_{it+j}) + f_{t+j}cov(\delta_i, P_{it}) \\
&+ (t+j+t_i)cov(\beta_{it+j}, P_{it}) + cov(P_{it}, P_{it+j}) + cov(v_{it}, v_{it+j}) \\
&= f_t f_{t+j}\sigma_\delta^2 + f_t(t+j+t_i)\prod_{k=1}^{t+j}\rho_k cov(\delta_i, \beta_{i0}) + f_t cov(\delta_i, \alpha_i) \\
&+ f_{t+j}(t+t_i)\prod_{k=1}^{t}\rho_k cov(\delta_i, \beta_{i0}) + (t+t_i)(t+j+t_i)\prod_{k=t+1}^{t+j}\rho_k var(\beta_{it}) \\
&+ (t+t_i)(\prod_{k=1}^{t}\rho_k cov(\beta_{i0}, \alpha_i) + \sum_{k=1}^{t-1}(\prod_{j=k+1}^{t}\rho_j cov(\omega_{ik}, \zeta_{ik}) + cov(\omega_{it}, \zeta_{it})) \\
&+ f_{t+j}cov(\delta_i, \alpha_i) + (t+j+t_i)(\prod_{k=1}^{t+j}\rho_k cov(\beta_{i0}, \alpha_i) + \sum_{k=1}^{t}(\prod_{j=k+1}^{t+j}\rho_j cov(\omega_{ik}, \zeta_{ik})) \\
&+ var(P_{it}) + cov(v_{it}, v_{it+j}).
\end{aligned} \tag{30}
$$

C Appendix C: The Compositions of Earning Inequality by Education

Table C-1. The compositions of earning inequality for the RIP

	e=1			e=2		
Year	Δvar(y)	permanent	transitory	Δvar(y)	permanent	transitory
1998	-0.0233	0.0001	-0.0234	-0.0199	0.0000	-0.0199
1999	-0.0290	0.0000	-0.0290	-0.0498	0.0000	-0.0498
2000	-0.0327	0.0007	-0.0334	0.0288	0.0021	0.0267
2001	0.0206	0.0044	0.0162	0.0037	0.0224	-0.0187
2002	0.0029	0.0022	0.0007	-0.0168	0.0000	-0.0168
2003	0.0010	0.0090	-0.0080	0.0614	0.0097	0.0517
2004	-0.0021	0.0010	-0.0031	0.0136	0.0071	0.0065
2005	-0.0046	0.0012	-0.0058	-0.0403	0.0000	-0.0403
2006	-0.0169	0.0000	-0.0169	0.0106	0.0049	0.0057
2007	-0.0070	0.0084	-0.0154	-0.0093	0.0000	-0.0093
2008	0.0142	0.0008	0.0134	0.0587	0.0038	0.0549
2009	-0.0019	0.0243	-0.0262	-0.0588	0.0000	-0.0588
2010	0.0629	0.0125	0.0504	-0.0227	0.0000	-0.0227
2011	-0.0595	0.0000	-0.0595	-0.0254	0.0000	-0.0254
2012	0.0068	0.0105	-0.0037	0.0313	0.0093	0.0220
2013	0.0192	0.0142	0.0050	0.1208	0.0235	0.0973
2014	0.0115	0.0000	0.0115	-0.0366	0.0000	-0.0366
2015	0.0199	0.0302	-0.0103	0.1003	0.0897	0.0106

Note: e=1 if head's education is high school or lower and e=2 if head's education is college or higher.

Table C-2. The compositions of earning inequality for the HIP

	e=1			e=2		
Year	Δvar(y)	deterministic	transitory	Δvar(y)	deterministic	transitory
1998	-0.0163	-0.0083	-0.0080	-0.0170	-0.0087	-0.0083
1999	-0.0286	-0.0070	-0.0216	-0.0466	-0.0074	-0.0391
2000	-0.0055	-0.0058	0.0003	0.0541	-0.0061	0.0603
2001	-0.0006	-0.0045	0.0039	-0.0082	-0.0048	-0.0034
2002	0.0045	-0.0033	0.0078	-0.0159	-0.0035	-0.0124
2003	0.0034	-0.0020	0.0054	0.0614	-0.0022	0.0636
2004	-0.0017	-0.0008	-0.0009	0.0163	-0.0009	0.0171
2005	-0.0036	0.0005	-0.0041	-0.0381	0.0004	-0.0386
2006	-0.0174	0.0017	-0.0191	0.0092	0.0017	0.0074
2007	-0.0069	0.0029	-0.0099	-0.0076	0.0030	-0.0107
2008	0.0148	0.0042	0.0106	0.0528	0.0044	0.0485
2009	-0.0008	0.0054	-0.0062	-0.0536	0.0057	-0.0592
2010	0.0694	0.0067	0.0627	-0.0276	0.0070	-0.0346
2011	-0.0544	0.0079	-0.0623	-0.0272	0.0083	-0.0355
2012	-0.0048	0.0092	-0.0139	0.0270	0.0096	0.0174
2013	0.0201	0.0104	0.0096	0.1220	0.0109	0.1111
2014	0.0107	0.0117	-0.0010	-0.0393	0.0122	-0.0515
2015	0.0165	0.0129	0.0036	0.1099	0.0135	0.0963

Note: e=1 if head's education is high school or lower and e=2 if head's education is college or higher.

Table C-3. The compositions of earning inequality for the Baker/Solon

Year	e=1				e=2			
	Δvar(y)	deterministic	permanent	transitory	Δvar(y)	deterministic	permanent	transitory
1998	-0.0062	-0.0002	-0.0063	0.0002	-0.0108	-0.0022	-0.0017	-0.0069
1999	-0.0251	-0.0058	-0.0007	-0.0186	-0.0397	-0.0115	-0.0052	-0.0230
2000	-0.0145	-0.0011	-0.0031	-0.0103	0.0111	-0.0020	-0.0009	0.0140
2001	-0.0055	-0.0007	0.0001	-0.0049	0.0035	0.0000	0.0183	-0.0147
2002	0.0016	0.0002	-0.0004	0.0017	-0.0187	-0.0057	-0.0050	-0.0080
2003	-0.0024	0.0024	-0.0012	-0.0036	0.0304	0.0034	0.0067	0.0203
2004	-0.0037	-0.0006	-0.0011	-0.0020	0.0052	0.0004	0.0012	0.0036
2005	-0.0059	0.0037	-0.0036	-0.0061	-0.0282	-0.0019	-0.0058	-0.0205
2006	-0.0048	-0.0008	-0.0014	-0.0026	0.0052	0.0061	-0.0016	0.0006
2007	-0.0022	0.0013	0.0101	-0.0136	-0.0105	-0.0052	0.0038	-0.0092
2008	0.0199	0.0007	0.0070	0.0121	0.0295	0.0088	-0.0047	0.0254
2009	-0.0056	0.0050	0.0057	-0.0163	-0.0319	-0.0081	0.0029	-0.0267
2010	0.0382	0.0104	0.0021	0.0257	-0.0106	-0.0010	-0.0028	-0.0067
2011	-0.0030	0.0014	0.0046	-0.0091	-0.0148	-0.0022	0.0019	-0.0145
2012	-0.0157	-0.0007	0.0112	-0.0261	0.0148	-0.0016	0.0048	0.0116
2013	0.0184	0.0066	0.0094	0.0024	0.0774	0.0100	0.0434	0.0240
2014	0.0190	0.0045	0.0179	-0.0035	0.0072	-0.0020	0.0323	-0.0231
2015	0.0285	0.0028	0.0380	-0.0122	0.0776	-0.0058	0.0647	0.0187

Note: e=1 if head's education is high school or lower and e=2 if head's education is college or higher.

Table C-4. The compositions of earning inequality for the G-RIP

Year	e=1			e=2		
	Δvar(y)	permanent	transitory	Δvar(y)	permanent	transitory
1998	0.0064	0.0276	-0.0213	-0.0069	-0.0786	0.0717
1999	-0.0358	-0.0441	0.0083	-0.0545	-0.0356	-0.0188
2000	-0.0222	-0.0087	-0.0135	0.0331	-0.0132	0.0463
2001	-0.0006	0.0077	-0.0083	-0.0025	0.0532	-0.0558
2002	0.0071	0.0145	-0.0074	-0.0209	-0.0356	0.0146
2003	0.0003	0.0184	-0.0181	0.0592	0.0380	0.0213
2004	-0.0023	-0.0171	0.0148	0.0164	0.0293	-0.0130
2005	-0.0134	-0.0050	-0.0084	-0.0480	-0.0360	-0.0121
2006	0.0002	-0.0024	0.0026	0.0175	0.0422	-0.0248
2007	-0.0110	0.0183	-0.0293	-0.0172	-0.0244	0.0072
2008	0.0303	-0.0195	0.0499	0.0700	0.0406	0.0294
2009	-0.0188	0.0328	-0.0516	-0.0613	-0.0381	-0.0232
2010	0.0643	0.0161	0.0482	-0.0125	-0.0170	0.0045
2011	-0.0248	-0.0123	-0.0125	-0.0263	-0.0123	-0.0140
2012	-0.0273	0.0126	-0.0399	0.0264	0.0028	0.0235
2013	0.0188	0.0137	0.0051	0.1084	0.0793	0.0291
2014	0.0161	-0.0004	0.0165	-0.0127	0.0033	-0.0160
2015	0.0135	0.0209	-0.0074	0.0863	0.0349	0.0514

Note: e=1 if head's education is high school or lower and e=2 if head's education is college or higher.

Table C-5. The compositions of earning inequality for the SHIP

	e=1				e=2			
Year	Δvar(y)	macro	permanent	transitory	Δvar(y)	macro	permanent	transitory
1998	0.0085	0.0208	0.0024	-0.0147	-0.0049	-0.0476	0.0004	0.0424
1999	-0.0389	-0.0606	0.0065	0.0153	-0.0587	-0.0463	0.0006	-0.0131
2000	-0.0177	-0.0359	0.0023	0.0158	0.0197	-0.0108	0.0026	0.0279
2001	-0.0069	0.0147	0.0002	-0.0218	0.0184	0.0459	-0.0008	-0.0266
2002	0.0088	0.0207	-0.0004	-0.0114	-0.0303	-0.0241	0.0011	-0.0073
2003	-0.0005	0.0669	0.0004	-0.0678	0.0657	0.0326	-0.0006	0.0337
2004	-0.0011	-0.0381	0.0020	0.0349	0.0161	0.0294	-0.0004	-0.0129
2005	-0.0078	-0.0024	0.0018	-0.0072	-0.0481	-0.0284	0.0007	-0.0203
2006	-0.0016	-0.0473	0.0014	0.0444	0.0076	0.0328	-0.0006	-0.0246
2007	-0.0149	0.0220	0.0009	-0.0378	-0.0146	-0.0168	0.0006	0.0016
2008	0.0366	-0.0081	0.0116	0.0331	0.0700	0.0404	-0.0006	0.0302
2009	-0.0207	-0.0108	0.0105	-0.0204	-0.0706	-0.0372	0.0008	-0.0342
2010	0.0602	0.0442	-0.0004	0.0165	-0.0082	-0.0246	0.0003	0.0161
2011	-0.0081	-0.0185	0.0016	0.0087	-0.0071	-0.0219	0.0007	0.0141
2012	-0.0429	-0.0196	0.0026	-0.0259	0.0102	0.0103	0.0004	-0.0005
2013	0.0189	0.0175	0.0003	0.0011	0.1368	0.0711	-0.0014	0.0671
2014	0.0172	0.0084	0.0003	0.0085	-0.0132	-0.0084	0.0004	-0.0052
2015	0.0117	0.0087	0.0005	0.0026	0.0630	0.1412	0.0013	-0.0795

Note: e=1 if head's education is high school or lower and e=2 if head's education is college or higher.

References

Alderman, H., J. R. Behrman, H.-P. Kohler, J. A. Maluccio, and S. C. Watkins (2001) "Attrition in longitudinal household survey data: some tests for three developing-country samples," *Demographic research*, 5, 79–124.

Andrews, D. W., and B. Lu (2001) "Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models," *Journal of econometrics*, 101(1), 123–164.

Arellano, M., R. Blundell, and S. Bonhomme (2017) "Earnings and consumption dynamics: a nonlinear panel data framework," *Econometrica*, 85(3), 693–734.

Baker, M. (1997) "Growth-rate heterogeneity and the covariance structure of life-cycle earnings," *Journal of labor Economics*, 15(2), 338–375.

Baker, M., and G. Solon (2003) "Earnings dynamics and inequality among Canadian men, 1976–1992: Evidence from longitudinal income tax records," *Journal of Labor Economics*, 21(2), 289–321.

Becker, G. S. (1975) "Front matter, human capital: a theoretical and empirical analysis, with special reference to education," in *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, Second Edition*, pp. 22–0. NBER.

Becketti, S., W. Gould, L. Lillard, and F. Welch (1988) "The panel study of income dynamics after fourteen years: An evaluation," *Journal of Labor Economics*, 6(4), 472–492.

Blundell, R., and B. Etheridge (2010) "Consumption, income and earnings inequality in Britain," *Review of Economic Dynamics*, 13(1), 76–102.

Blundell, R., L. Pistaferri, and I. Preston (2008) "Consumption inequality and partial insurance," *American Economic Review*, 98(5), 1887–1921.

Carroll, C. D., and L. H. Summers (1991) "Consumption growth parallels income growth: some new evidence," in *National saving and economic performance*, pp. 305–348. University of Chicago Press.

Choi, J., S. Kim, and S. Park (2018) "Income Inequality in Korea in the Post Global Financial Crisis Period," *The Korean Journal of Economic Studies*, 66(1), 115–142.

Fitzgerald, J., P. Gottschalk, and R. Moffitt (1998) "An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics," Discussion paper, National Bureau of Economic Research.

Guvenen, F. (2009) "An empirical investigation of labor income processes," *Review of Economic Dynamics*, 12(1), 58–79.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015) "What do data on millions of US workers reveal about life-cycle earnings risk?," Discussion paper, National Bureau of Economic Research.

Haider, S. J. (2001) "Earnings instability and earnings inequality of males in the United States: 1967–1991," *Journal of labor Economics*, 19(4), 799–836.

Hong, M.-K. (2015) "Top Wage Shares in the Long Run of Korea (1958-2013)," *Korean Journal of Labor Studies*, 21(1), 191–220.

Jappelli, T., and L. Pistaferri (2010) "Does consumption inequality track income inequality in Italy?," *Review of Economic Dynamics*, 13(1), 133–153.

Kim, N. N., and J. Kim (2015) "Top incomes in Korea, 1933-2010: Evidence from income tax statistics," *Hitotsubashi Journal of Economics*, pp. 1–19.

Meghir, C., and L. Pistaferri (2004) "Income variance dynamics and heterogeneity," *Econometrica*, 72(1), 1–32.

Mincer, J. (1974) "Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.," .

Moffitt, R. A., and P. Gottschalk (2011) "Trends in the covariance structure of earnings in the US: 1969–1987," *The Journal of Economic Inequality*, 9(3), 439–459.