Working Paper 2015-09

The Effects of an Aging Population on Personal Income Tax Revenue in Korea

- Evidence from a dynamic microsimulation

Jayee Ko · Hyukjin Kwon

KIHASA Korea Institute for Health and Social Affairs

The Effects of an Aging Population on Personal Income Tax Revenue in Korea: evidence from a dynamic microsimulation

Jayee Ko, Associate Research Fellow
Hyukjin Kwon, Gyeongnam National Univ. of Science and Technology

Korea Institute for Health and Social Affairs
Building D, 370 Sicheong-daero, Sejong city
30147 KOREA
http://www.kihasa.re.kr
ISBN: 978-89-6827-307-0 93330

Contents

CHAPTER 1
Introduction ········· 1

CHAPTER 2
Literature Review and model development ········· 7

1. Established literature and its limits ········· 9
2. MMESP Overview ········· 15
3. K.Sim: Basic structure and simulation process ········· 29

CHAPTER 3
Validation test of resulting estimates ········· 37

1. Validation test of income distribution estimates ········· 41
2. Testing the validity of income tax estimates ········· 53

CHAPTER 4
Long-term forecasts on income taxes: population aging and increasing pension income ········· 53

1. Baseline forecasts ········· 55
2. Income tax revenue forecasts based on adjustments to tax brackets ········· 59
3. Rising National Pension contribution rate and the income tax revenue forecasts ········· 65

CHAPTER 5
Conclusion 71

References 79

Appendix 83

List of Tables

〈Table 1〉 Life Events and Parameters for Simulation ··· 18
〈Table 2〉 Income Panel Analysis (GLS Random Effect Model with an AR (1) Disturbance) ··· 23
〈Table 3〉 Tax Revenue and Estimates, 2010-2012 (Nominal) ··· 50
〈Table 4〉 Average Effective Tax Rates (AETRs) and Tax Concentrations, 2010-2012 (Nominal) ··· 51
〈Table 5〉 Baseline Forecasts, 2014-2100 ··· 57
〈Table 6〉 Case 1: Tax Brackets Readjusted, 2014-2100 ··· 60
〈Table 7〉 Case 2: Tax Brackets Readjusted, 2014-2100 ··· 61
〈Table 8〉 Social Security-Related Tax Deductions ··· 65
〈Table 9〉 Raising National Pension Contribution Rates and Changes in Tax Structure, 2014-2100 ··· 68

〈Appendix Table 1〉 Composite Income Tax Deductions by Year ··· 83
〈Appendix Table 2〉 Income Tax Rates by Year ··· 84
〈Appendix Table 3〉 Income Tax Deductions ··· 85
〈Appendix Table 4〉 EITC and CTC Eligibility Criteria ··· 86

List of Figures

〈Figure 1〉 MMESP and K.Sim ··· 20
〈Figure 2〉 Content and Scope of K.Sim ··· 30
〈Figure 3〉 Comparison to Other Income Distribution Estimates ··· 42
〈Figure 4〉 Wage and Business Income Distribution by Age ··· 44
〈Figure 5〉 Wage and Business Income Profile by Birth Cohort ··· 44
〈Figure 6〉 Changes in the Wage and Business Income Distributions ··· 45
〈Figure 7〉 Changes in the Pension Income Distribution ··· 46

〈Figure 8〉 Changes in the Total Income Distribution 47
〈Figure 9〉 EITC Distribution. 2010-2012 (Nominal) 52
〈Figure 10〉 Economic Variables and Hypotheses 56
〈Figure 11〉 Gaps in Income Tax Revenue as a Share of Nominal GDP: Pension Income Excluded, 2014-2100 64
〈Figure 12〉 Trend in the Income Tax Revenue: Based on Increases in Contribution Rates, 2014-2100 67

〈Appendix Figure 1〉 Wage and Business Income Distributions (Converted into Current Price Using CPI) 88
〈Appendix Figure 2〉 Pension Income Distributions (Converted into Current Price Using CPI) 88
〈Appendix Figure 3〉 Total Income Distributions (Converted into Current Price Using CPI) 88

1 Introduction

1 Introduction

As Korea is fast becoming a post-aged society, characterized by a shrinking workforce and a concomitantly shrinking economy, Korean policymakers are under mounting pressure to address and solve the issues likely to affect future spending programs, particularly those involving social security benefits. As social security benefits in Korea are distributed to members of diverse income levels, spending patterns, and lifecycle development stages, it is crucial to accurately diagnose the current and future statuses of tax revenue and aggregate social spending and find appropriate policy solutions. It has become even more crucial and urgent in the light of anticipated demographic changes associated with Korea's population aging.

Korea has undergone dramatic economic growth and its social security system has been evolving around social insurance spending. As of 2014, the Korean government provides 22 percent of the total budget required for social insurance programs. The rapid increase in recent years in the amount of the tax revenue spent directly on social security benefits has raised concerns, as the government will soon need to adjust tax rates just to sustain the current social security structure, while chances of its raising contributions to social insurance schemes remain slim. In other words, the Korean government will be compelled to fund inevitable increases in social security spending by altering major tax item rates.

Given current tax rates and in the interest of the efficiency of the overall economy, the government may be able to raise the personal income tax (PIT) and value-added tax (VAT) rates slightly. Raising tax rates, however, usually attracts controversy, and it necessarily affects other taxes collected and transferred to other spending programs by altering the existing system of incentives within the current economy. Therefore, Korean policymakers will need to adjust tax rates in ways that are compatible with the overarching framework of the taxes and benefits system that exists today. Before trying to find ways to ensure a more equitable distribution of tax burdens across different generations during rapid demographic change, policymakers must first consider and analyze the distribution of burdens and benefits under the current system. Given this necessity, it is no wonder that a growing number of researchers and policymakers are focusing on the long-term prospects and stability of the Korean government's fiscal management policy.

The majority of existing studies on the long-term prospects of government revenue and spending in Korea take aggregate or macroscopic approaches. However, the worldwide trend today, particularly in advanced economies, is to analyze the short- and long-term social and economic effects of tax and fiscal policies using microsimulation models that take into account the interactions of decision-makers on the micro level, including individuals and businesses. The Congressional Budget Office (CBO) of the United States, for instance, has been using its own microsimulation model, known as CBOLT, since 2001 in order to analyze and forecast the fiscal and distributive effects of various tax and social security programs over long periods, i.e., 75 years or longer (CBO, 2009). In contrast to macro-simulation models, microsimulation models provide particularly useful ways to identify the winners and losers of specific changes in various policies and programs in both the short and long term.

This study therefore starts by developing a tax-benefit model, K.Sim, to be used for the design and analysis of changes in tax, fiscal and benefit programs in the intermediate to long term. In an effort to minimize the material, human, and temporal resources required in the development of this model, we based it on the Microsimulation Model for Economic and Social Policy (MMESP) in Korea (Kwon et al., 2013) that was recently developed through collaboration among several researchers. As a dynamic microsimulation model, MMESP comprises various modules on such factors as birth rates, education, marital status, employment, income levels, the National Pension Scheme, and death rates. Our model, K.Sim, therefore uses the MMESP to forecast and analyze how population aging, as measured by the long-term changes in these personal and household factors, will affect income tax revenue.

This study is organized as follows. In the first section of Chapter 2, we survey and discuss the limits of the established literature in forecasting likely changes in income tax revenue over the long term and explain the need for a new micro-simulation model such as ours. In the second section, we describe the characteristics of MMESP in detail. In the third section, we explain the overall structure, composition, and methodology of K.Sim, our new tax-benefit model. In Chapter 3, we review the validity of our model by comparing our simulation results to the indicators of the actual economy. In Chapter 4, we forecast and analyze the long-term effects of population aging on income tax revenue and delineate major policy implications. In Chapter 5, we briefly discuss the shortcomings of this study and the tasks for future research.

2 Literature Review and model development

1. Established literature and its limits
2. MMESP Overview
3. K.Sim: Basic structure and simulation process

2 Literature Review and model development

1. Established literature and its limits

A number of studies have been published in recent years in Korea providing long-term forecasts on how population aging will affect income tax revenue. Park and Hong (2011), for instance, provides an empirical analysis of how population aging will reduce the proportion of wage income in disposable household income over time. These authors predict that as the wage or wage income level begins to fall along with the number of working-age households over the long term, the source of wage income tax revenue will gradually shrink in Korea. Nevertheless, the authors also forecast that given the stable structure of wage income distribution and its relative immunity to demographic change, the amount of wage income tax revenue will mostly remain intact.

Park et al. (2012), in making long-term forecasts on income tax revenue, links analysis of likely changes in income tax revenue resulting from changes in the tax base to the findings of Park and Hong (2011). According to Park et al. (2012), the amount of tax revenue will increase in a scenario of increasing income inequality, given the progressive structure of income taxes in Korea. However, decreases in the labor income share and in business income will contract the source of the tax revenue, dragging the size of the tax revenue as a share of GDP down from 2.2 percent in 2011 to 1.97 percent by 2030.

Seong (2012) and the National Assembly Budget Office (NABO, 2012), on the other hand, take micro-level approaches to long-term forecasts on income tax revenue. Both studies, extending to the year 2060, base their forecasts on the average income and tax burden projections by age group made in the Household Income and Expenditure Survey (HIES) and on the population and household projections made by Statistics Korea. NABO (2012) concludes that the amount of wage and composite income taxes as a share of the GDP will rise very slowly, from 3.43 percent in 2012 to 3.95 percent in 2036, and from then on will begin to fall gradually to reach 3.77 percent by 2060. However, Seong (2012) predicts that with the inflation effect controlled, the amount of wage and composite income taxes as a share of the GDP will keep rising from 2.29 percent in 2011 to 5.76 percent in 2050. Seong (2012) concludes that the amount of these taxes in the light of the nominal GDP growth rate will peak at 2.45 percent by 2022 but will fall rapidly after to reach 1.85 percent by 2060.

The two studies differ significantly in their conclusions due to differences in the premises and methods of analysis they use. NABO (2012) forecasts the changing distribution of the tax-paying population using the relative tax burden profiles of different income-paying age groups and population aging estimates. The chosen method of analysis is relatively simple, as it involves multiplying the average per capita tax burden on each age group by the size of the population of each group and adding up the sums of all age groups to estimate the resulting amount of income tax. Seong (2012), on the other hand, employs a relatively more advanced method, estimating changes in the distributions of household income and spending until 2060, using Statistics Korea's population and household projections and concomitantly adjusting the weights of household distribution by age group. Seong (2012), in particular, analyzes how the abolition of certain deductible items, such as the costs of education, would affect the amount of tax revenue, thus achieving a feat that would have been unimaginable with simpler methodologies, such as that used by NABO (2012). Nevertheless, the accuracy and utility of Seong's advanced model is somewhat offset by its complexity.

While both NABO (2012) and Seong (2012) adopt, like this study, microsimulation approaches forecasting how population aging will affect the tax revenue, this study is set apart from them and the established literature for the following reasons.

First, we take into account more specific features and factors of the income tax structure in Korea, particularly focusing on the assumption that the number of pension beneficiaries and the amount of pension benefits will multiply rapidly in the coming decades as the public pension schemes mature. According to the National Pension Projections Commission (NPPC, 2013), the proportion of pension beneficiaries aged 65 or older will multiply dramatically from 32.8 percent in 2013 to 80.6 percent by 2050 and to 91.3 percent by 2060. The amount of pension benefits received by these beneficiaries will similarly increase, while the proportion of the elderly population earning less than the minimum income, and thus qualifying for basic deductions (currently KRW 1 million or less), will grow.[1)] Also, the amount of pension income received by persons who joined the National Pension Scheme in and after 2002 will be taxable. The declining birth rate will further affect the income tax revenue structure by radically shrinking the amounts of basic deductible items associated with childcare and education.

Second, as in the case of researchers in other advanced economies, we use a microsimulation model to forecast how certain policy decisions affect individuals and households.[2)]

---

1) Under the current tax law, Koreans cannot deduct from their income the donations made by their parents who are not persons eligible for basic deductions, the amounts of money they spent with their credit cards, or the liability insurance premiums they pay. However, Koreans of all income levels who pay for part of the living expenses of their parents, and thereby incur certain medical costs, may deduct those medical costs from their income. Accordingly, the increase in the proportion of the elderly population with income more than KRW 1 million each will serve to contract the amount of deductions in the income tax structure.

2) Other studies published in Korea that explain the current status, applications, and methodologies of microsimulation models and apply them to their own analyses include Seong et al. (2008), Kwon and Han (2009), and

While the two other studies discussed above from 2012 also adopt micro-level approaches, they differ somewhat from the dynamic microsimulation approach that is in vogue around the world today. The term "dynamic" indicates that the model changes the quantities of the social and economic factors affecting individuals and households from year to year according to predefined simulation rules. The individual in a given dynamic microsimulation model will experience changes in his or her status over time, such as being married, being employed, giving birth or being out of job, according to certain idiosyncratic, economic, and social factors. Seong (2012), in contrast, keeps the characteristics of individuals and households constant, only changing the weights attached to these actors according to forecast household distributions. His is therefore a static model.

We can distinguish between dynamic and static microsimulation models depending on whether they allow for variations or changes in the characteristics of the individual subject

---

Kim et al. (2010). According to these studies, it was Orcutt (1957) that first introduced the methodology of microsimulation. Thanks to the accumulation of micro-data, the development of increasingly advanced software and hardware, and the increasing policy need and support for long-term forecasts on the effects of social and economic policies (pensions, education, healthcare, etc.), microsimulation methods became more popular—noticeably in the early 1990s and particularly in advanced economies. Over 40 models are currently being used in a dozen nations or so worldwide today. The emergence of the International Microsimulation Association in October 2005 and the founding of the International Journal of Microsimulation in 2007, have promoted to the continuing spread and progress of microsimulation models.

in the analysis over time. Seong (2012), for instance, draws upon Statistics Korea's projections of population distribution by age group in determining the weight to attach to each age group's population. He then performs regression analyses on certain factors, such as the number of household members, in order to forecast future population distributions. Such a static model, however, hardly has any room for deciding the weights to attach to more specific economic and social factors, such as the educational attainments and employment types of individuals. Korea, like many other advanced countries worldwide, does not provide official projections on future population distributions by factors other than sex and age. Moreover, static models are incapable of predicting phenomena that have not taken place or that have been rare until now but are likely to occur with increasing frequency in the future. For example, as our National Pension structure matures, the number of pensioners and their income levels will naturally increase. So, with a static model, we cannot make reliable forecasts about the likely changes resulting from these factors in the elderly population well into the future.

Dynamic models overcome this static-model deficiency by allowing for changes in individual characteristics over time according to predefined rules. However, dynamic models are far more technically complex than static ones. Researchers are still struggling to find ways to solve various technical difficulties that arise, such as the instability and dependency of various combinations and marginal distributions, created according to predefined rules and identification errors plaguing the rules themselves. Researchers therefore increasingly resort to alignment mechanisms in order to offset the difficulties in dynamic models.

Static and dynamic models have their respective advantages and weaknesses. Dekkers and Belloni (2009) and Li and O'Donoghue (2013) nonetheless recommend dynamic models in the analysis of population aging, as these models have room for greater ranges of specific phenomena and characteristics.[3] This study therefore uses a dynamic microsimulation model to determine how the declining birth rate and aging of population in Korea—taking into account the growing number of pensioners and the rise in the pension income—is likely to affect the income tax revenue in the country over the long term.

2. MMESP Overview

In order to make long-term forecasts on the tax income revenue of a given country, we need data and forecasts on the changing income distribution over time. These data and forecasts, in turn, must reflect the age and cohort effects on the

3) As for the current status, purposes, decision-makers, and methodological characteristics of microsimulation models worldwide today, see McCraken and Gupta (2009) and Li and O'Donoghue (2013).

personal income level, in addition to the effects of interactions among individual actors. Some researchers may opt to assume and understand the overall rise in the wage level as a factor of the long-term expansion of tax bases over time. This is because the basic age and income profile structure may remain relatively unchanged, notwithstanding dramatic changes in the by-occupation composition of the workforce or in certain occupations and industries.

Nevertheless, the changes in tax bases are not solely reliable upon the age and income profile structure. Tax bases change over time due to changes in the composition of working people and pensioners, other taxable income earners, and the positions of economic actors on the labor market. Also, the income tax regime in Korea, involves diverse deductions that vary in light of household composition, spending, and saving patterns. Forecasts on the income tax revenue therefore require forecasts on the changing economic, social, and income characteristics of individuals and households. In order to forecast and analyze possible long-term changes in the public pension schemes in Korea, we decided to use MMESP, a microsimulation model that had already been developed by the time we began this study.[4] Of the diverse microsimulation models,

4) For a more detailed explanation of the model, it is recommended that readers consult Kwon et al. (2013c). These authors provide a thorough description of the basic characteristics and outcomes of this model, thus eliminating the need for a duplicated explanation in the present study. Our study therefore provides only a brief overview of the model and the changes

MMESP can be grouped with other dynamic population with cross-sectional aging models. MMESP is dynamic in the sense that it assumes annual changes to the social and economic status of individual decision-makers given the probability of life events (being employed, being out of job, being married, giving birth, dying, etc.).

The data and methods for deciding the microscopic life events and determining their probabilities using MMESP are listed in Table 1.[5] The scope of analysis using MMESP extends to the entire population and not just to any specific age or cohort groups, as the model involves using a base dataset that consists of a one-percent sample (i.e., 445,000 individuals) drawn from the Population and Housing Census (2005).[6] MMESP is cross-sectional in its aging model as its aging process involves simulating the economic and social status of all the target individuals at time t, and then it simulates likely changes in that status across the section as the focus of our analysis moves to t+1 and so forth. The longitudinal aging process of simulating likely changes in the social and economic status of

made to it since 2013. We the authors of this study are always open to personal inquiries and comments regarding the model.

5) For the specific hypotheses, forecasting methodologies, and outcomes of these different modules, see Kwon et al. (2013).

6) The modifications we made to the original data set include the conversion of households into families and the specification of the school enrollment rates by the legal enrollment age. Additional imputations we have made include the distinction between full-time and part-time wage earners, income levels, and the status of contributions to and benefits from the National Pension. For more specific discussions on these, see Kwon et al. (2013c).

each individual throughout his or her lifetime, or between time t and time t+N, may have its technical and theoretical merits. The cross-sectional aging process, however, allows for a more realistic simulation of interactions among diverse individuals.7)

〈Table 1〉 Life Events and Parameters for Simulation

Module	Life event	Alignment set	Parameters
Population, birth	Birth	Population projections in Third Pension Spending Forecasts (i.e., number of babies born each year).	Korean Labor and Income Panel Studies (KLIPSs) I-XI and logit model (i.e., age, square of age, employment status, number of children, waiting time until birth and its square).
Household composition	Marital status (first- and second-time marriage)	Number of marriages per 1,000 in Statistics Korea's population trend surveys.	HIEP's and Pereses's (2002) "MateMatching" algorithms (i.e., men's age, six age difference dummies for married couples, educational attainment difference dummy).
	Divorce	Number of divorces per 1,000 in Statistics Korea's population trend surveys.	Random.
Education	Enrollment in high schools	Enrollment and employment rates by sex, as researched by KEDI's Education Statistics Research Center.	Random.
	College enrollment		KLIPSs III-XI and multinomial logit model (i.e., household income and parents' educational attainments).
	Military enlistment	Data on enlistment candidates by grade on physical examinations (2006-2011, National Assembly Auditing, 2011).	Random.
Labor market 1	Changes in employment status	Economically Active Population Surveys (distribution of persons by	KLIPSs III-XI and multinomial logit model (i.e., age, square of age, marital status dummies,

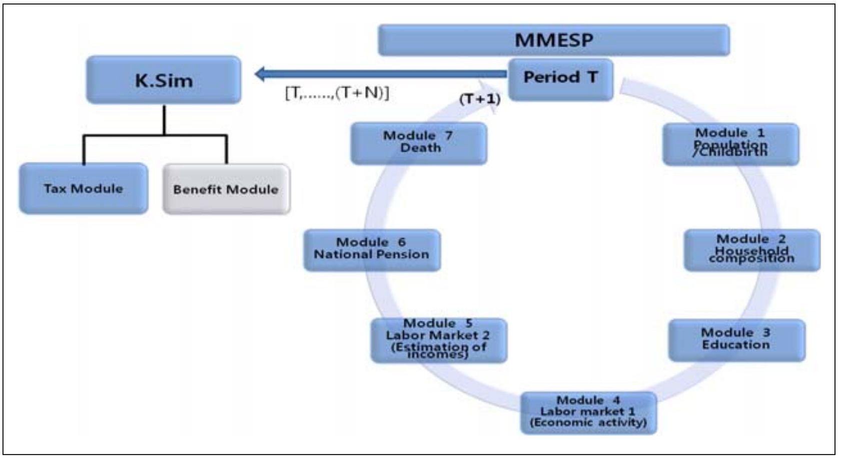
7) Dekkers and Belloni (2009) describes the dynamic population with cross-sectional aging model as the most useful when it comes to simulating and analyzing the equalization of pension income levels before and after retirement, poverty among the elderly, and population aging. For the specific reasons, see Dekkers and Belloni (2009).

Module	Life event	Alignment set	Parameters
		sex, age (every five years), employment status, and position at workplace) and basic macro variables in Third Pension Spending Forecasts.	number of schooling years, number of children under age five, economic/employment status previous term [five dummies]).
Labor market 2	Wage income	Basic macro variables in Third Pension Spending Forecasts.	KLIPSs III-XI and panel probability effect model (AR(1)-GLS) (i.e., age, sex, number of household members, house ownership, marital status, educational attainments, employment history).
National Pension	Contributions and benefits	Macro and institutional variables in Third Pension Spending Forecasts.	Additional Economically Active Population Surveys (2007-2011, every August), and logistic regression model (i.e., sex, age, square of age, marital status, four dummies for educational attainments, three dummies for employment status).
Death	Death	Population projections in Third Pension Spending Forecasts (i.e., death rate by sex and age).	Random.

Note: Based on Table 5 in Kwon et al. (2013) but partially modified and supplemented.

Figure 1 illustrates the process or order in which MMESP simulates the anticipated variations across the entire period from 2006 to 2100. The program generates a dataset for each year, containing the simulation results based on variations in the income, economic, and social status of individuals and households. K.Sim, the model we have developed to forecast long-term changes in income tax revenue, crucially relies on these year-by-year datasets generated by MMESP to forecast the income tax burdens on individuals. This allows us to save time and effort in our forecasting task. Since Kwon et al. (2013), MMESP has been modified in the following ways.[8)]

〈Figure 1〉 MMESP and K.Sim



First, the basic variables concerning the population, macro-economics, and the National Pension were replaced with the projections included in the Third National Pension Spending Forecasts (NPSF3). Used in other long-term forecast studies in Korea, these projections facilitate the comparison between this and other studies.

Second, the model also has been modified to specify and re-fine the income estimation process. The income estimation module is the most important of all microsimulation modules for the purposes of this study. The creators of MMESP esti-mated income by first dividing individuals into groups accord-ing to employment status (which can change over time). This

8) MMESP as created and used by Kwon et al. (2013) is Version 1.0. This study uses MMESP Version 1.5.

study, by contrast, divides people by the time-immune factor of sex and estimates changing income levels for each sex. The objective is to determine the effects of cross-sectional heterogeneity and the time-variant idiosyncrasies of individuals on individual income levels. Income is therefore estimated using a random effect model that assumes the first-order autocorrelation.[9)]

$$\ln(w_{it}) = \alpha + X \cdot \beta + \mu_i + \varepsilon_{it} \quad \cdots\cdots \quad (1)$$

$$\varepsilon_{it} = \rho\varepsilon_{it-1} + \nu_{it} \quad \cdots\cdots \quad (2)$$

$$(i = 1, \ldots, N;\ t = 1, \ldots. t)$$

One advantage of using a random effect model is that it allows us to estimate not only how cross-sectional variables at fixed points in time affect the income level, but also how the variations in those variables over time affect the income level. Numerous microsimulation studies have used the income estimates, based on cross-sectional analysis, to forecast future income levels. This method, however, assumes that income levels at fixed points in time are affected by identical cross-sectional characteristics only, and so it fails to take into account how factors from the past (e.g., past income levels) could affect future income levels.

9) This study relies on the income estimation model used in Zaidi et al. (2009), but allows for variations in certain independent variables. In the random effect model assuming a first-order autocorrelation of the error term, we could not obtain efficient estimates, unless we used the generalized least square (GLS) method as proposed by Baltagi and Wu (1999).

A random effect model, on the other hand, allows us to take into account both time-variant and time-invariant factors in forecasting income levels in the following three ways:

(1) By determining how time-invariant factors would affect income, i.e., how the observed attributes, including constants, educational attainments, and sexes, and unobserved attributes specific to individuals (μ_i) together would affect income;

(2) By determining how the time-variant error term, ε_{it}, would affect income, first by determining the initial using the formula $\varepsilon \sim N(0, \sigma_\varepsilon^2)$, and then using Equation (2) to update it annually.[10] Here, ν_{it} follows the normal distribution, $N(0, \sigma_\nu^2)$, and σ_ν is obtained using theformula $\sqrt{(1-\rho^2)} \cdot \sigma_\varepsilon$.

(3) By determining how time-variant factors, such as employment history, age, the square of age, and the economic activities of the spouse, would affect income. For our study, we converted employment history into three dummy variables, depending on whether each individual concerned was unemployed prior to time t but employed at time t, employed at both times, or economically non-active prior to time t but employed at time t (i.e., "ee," "ue," and "ne").

10) In Equation (2), we assumed that , $|\rho| < 1$, $\nu \sim iid(0, \sigma_\nu)$was satisfied.

〈Table 2〉 Income Panel Analysis (GLS Random Effect Model with an AR (1) Disturbance)

	Independent variable	Male	Female
General	Age	0.12506***	0.07381***
	Square of age	-0.00143***	-0.00094***
	Household head (= 1)	0.07334***	0.04246***
	Spouse (= 1)	0.12382***	0.07320***
	Spouse employed (= 1)	-0.05211***	-0.03690***
	Number of children aged 8 or younger	-	-0.06587***
Educational attainments	Less than middle school (= 1)	-0.31751***	-0.34045***
	High school (= 1)	-0.06588***	-0.14901***
	Four-year university (= 1)	0.18889***	0.21515***
	Graduate school (= 1)	0.37117***	0.35358***
Employment status	Full-time worker (= 1)	0.22019***	0.30743***
	Unpaid worker (= 1)	0.23870***	0.32743***
Employment history	ue (= 1)	-0.04655***	-
	ne (= 1)	-0.03382***	-0.08167***
Year dummies	2001	0.04951***	0.05659***
	2002	0.12737***	0.11805***
	2003	0.15128***	0.15765***
	2004	0.18301***	0.18673***
	2005	0.18677***	0.18895***
	2006	0.22126***	0.22293***
	2007	0.25029***	0.23915***
	2008	0.25189***	0.25425***
	Constant	4.74609***	5.56262***
rho_ar		0.33123075	0.36529865
sigma_u		0.37206519	0.40041666
sigma_e		0.36062317	0.35430091
rho_fov		0.51561269	0.56087604
N		31876	17981
R2	within	0.1180	0.1340
	between	0.4807	0.3732
	overall	0.3707	0.3337

Notes: 1) The dependent variables were obtained by taking natural logarithms of the average monthly wage (KRW) multiplied by 12 months.

2) The base variables for the dummies are the two-year college diploma for educational attainments and being employed for two consecutive points in time for employment history.

3) The asterisks, ***, **, and * stand for significance levels of p 〈 one percent, p 〈 five percent, and p 〈 10 percent, respectively.

The data we used for our income estimates are drawn from KLIPSs III through XI. We also turned the monthly average wage into annual income by multiplying it times 12. Then, we adjusted the annual income level using the inflation rate from 2010 and taking natural logarithms of it. Table 8 lists the results of our GLS random effect model with an AR (1) disturbance.[11] Based on the distribution of disturbance, we took into account the stochastic element and estimated per capita income for each year. The nominal wage growth rate of each year was converted using the current price of each year based on the projections in the NPSF3.

Third, for the alignments in the simulation of National Pension participation, we tried to appropriate the alignments used in the NPSF3 as much as possible. Earlier studies assumed that the data observed in 2011 would remain intact into the future. However, as the NPSF3 assumes that the participation rate in the National Pension Scheme will rise at a certain pac e[12], we likewise assume that the number of participants in the

11) Final results, obtained after excluding statistically insignificant variables from the regression equation, are presented here.

12) The NPSF3 first estimates the size of the economically active population in the total population, and then it applies the National Pension participation rate to the economically active population aged 18 to 59 to arrive at the total number of pension participants. The study then applies the residence-based participation rate to this number to estimate the numbers of workplace-based and residence-based participants. In light of the upward trend in the participation rate, the NPSF3 assumes that the National Pension participation rate increases in the form of a log-function (to 87.4 percent in 2011 and to 90 percent by 2015, to remain intact afterward). The total participation rate for each year is based on the

scheme and the number of pensioners will grow at a certain pace.

MMESP estimates the number of National Pension participants for each year according to the probability of participation, predefined on the basis of individuals' social and economic attributes as well as the alignment of pension types, sexes, and age groups given for each year. However, the NPSF3 bases its probabilities on the distribution of participants by sex and the participation rates of each five-year age group as of 2011, the five-year forecasts on the distribution of residence-based participants by sex and age group for the period of 2011 to 2050, the ratio of contribution-exempt residence-based participants by sex and age group as of 2010, and the 10-year forecasts on the distribution of participants by sex and the economic activity of each group for the period of 2011 to 2083.

Pursuant to the pension law, persons who work for less than 15 hours per week and persons who are not required by law to join the National Pension Scheme because they can participate in other public pension schemes (e.g., for government employees, people in military service, and private school teachers) are also excluded from the process of estimating participation

participation rates by sex and age group, estimated on the basis of the actual participation rates by sex and age group as of 2011. The relative composition of the workplace- and residence-based participants is also forecast on the basis of the observed data from 2011 (55.9 percent to 44.1 percent) and is expected to reach 70 percent to 30 percent by 2050. The proportion of contribution-exempt participants among residence-based participants is also expected to decrease to 30 percent by 2050 and remain intact afterward, on the basis of the observed data from 2011.

rates. The NPSF3 includes both persons who pay premiums on their pension insurances and persons who are exempt from making such contributions in its estimates of the overall economically active population. MMESP, on the other hand, estimates the pension participation rate only among the premium-paying members of each pension type, i.e., whether workplace- or residence-based.

MMESP thus estimates the pension participation rate by pension type, sex, and age group for each year in the following process. First, using the economic activity rate estimated in the NPSF3, the model estimates the number of participation-eligible persons in the working-age population (aged 18 to 59) by sex and five-year age group. Second, assuming that the participation rate, whether for each sex or age group, will rise as a log-function until 2015 and will remain the same in the years afterward, the model forecasts the likely number of pension participants by sex and age group. Third, based on the 10-year forecasts on the participation rates of sexes and different age groups of residence-based participants for the years 2011 to 2050, the model estimates the number of participants of each pension type. For the years after 2050, as in the NPSF3, the model assumes that the rates estimated for the year 2050 will remain intact. Fourth, as in the NPSF3, the model assumes—using the technique of linear interpolation—that the ratio of contribution-exempt participants will steadily drop from 56.5 percent in 2011 to 30 percent by 2050 and remain the same afterward. Based on this assumption, MMESP estimates the index of the contribution-exempt participation ratio by sex and age group for each year.

In this process, MMESP estimates (a) the number of workplace-based National Pension participants, (b) the number of residence-based participants, (c) the number of contribution-exempt participants, and (d) the number of economically active persons who are not required by law to participate in the National Pension for each sex and age group. Throughout the simulation period starting after 2011, the model estimates the participation rates among workplace-based participants using the formula (a)/((a)+(d))%, and among residence-based participants using the formula (b)/((b)+(c))%.

Next, in categorizing full-time, part-time, and daily workers as workplace-based pension participants, on the one hand, and unpaid workers and the unemployed as residence-based ones, on the other, MMESP simulates National Pension participation according to participation probabilities based on the participation rates estimated for each year, sex, age group, and individual. Then MMESP deduces the distribution of pension insurance premiums or contributions. Here the MMESP also readjusts the floor and ceiling on the minimum monthly income, which has been the basis for determining the amount of National Pension benefits since 2011, using the rates of increase in the A-value. MMESP assumes that full-time paid workers participate in the pension scheme for 12 months out of each year, but simulates different numbers of months of participation for part-time, daily, and unpaid workers. It does this by first estimating the ratios of different types of workers by sex, age group, and participation months using the data from the Korean Retirement and Income Studies (KReIS) of the National Pension Service (NPS) and then applying a random sampling method based on a uniform probability distribution.

The process of estimating the number of National Pension beneficiaries and the amount of pension benefits they receive is much simpler by comparison. Since we have already generated yearly records on pension participation using our simulation model, we can now apply these records along with the data on pension benefits and eligibility criteria. However, to estimate the number of beneficiaries of disability, survivor, and early old-age pension benefits, we need a prior estimate on the probabilities of disability and of receiving old-age pension benefits. Thus, we borrow the forecasts made in the NPSF3. But to estimate the number of survivor pension beneficiaries, we used the pension participation histories of the deceased and their survivors instead of the survivor ratio forecast in the NPSF3.[13] For estimating the amounts of all pension benefits,

13) Therefore, we simulate the survivor pension scenarios using the death module, and we simulate the loss of survivor pension benefits due to the survivor's marriage to a new partner, and other such reasons, using the

we used the individual pension histories accumulated each year, with consideration for the changes introduced into the pension scheme as part of the Pension Reform of 2007.[14)]

Our simulation also differs from that of the original MMESP in that we have added the variable of each individual's parental ID. As a result, we extended the scope of basic deductions (based on certain income and age-related criteria) to children whose parents are eligible for such deductions, but who do not live with those parents. For households with two or more children each, the deductions were confined to one child for each household only given the dependency, age, and income criteria.

3. K.Sim: Basic structure and simulation process

Figure 2 illustrates the overarching structure and scope of K.Sim, a microsimulation model we have specifically developed for designing and evaluating mid- to long-term changes in Korea's taxation, fiscal, and social welfare policies. Given the constraints of time and the limits on human and material re-

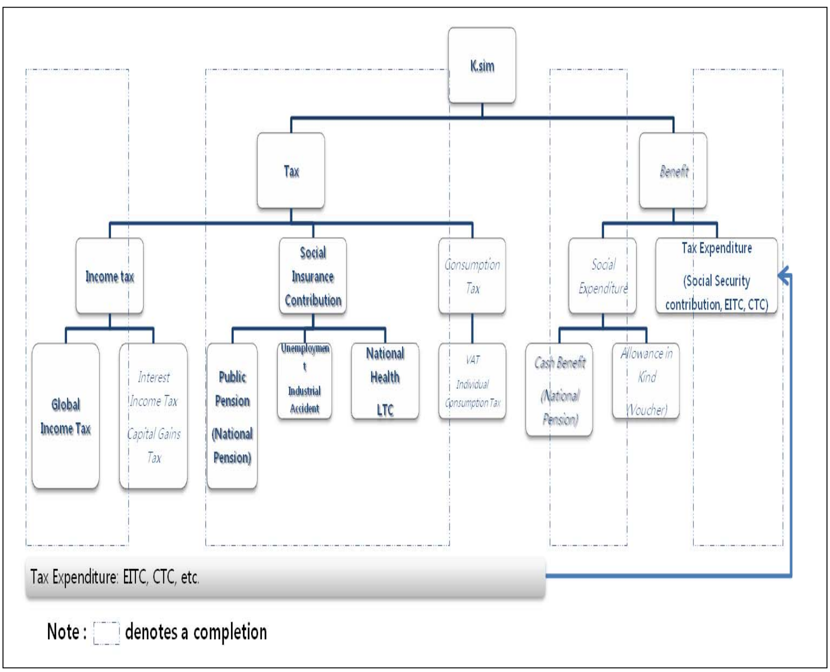
family composition module.

14) Our microsimulation model rests upon relatively simple assumptions concerning pension benefits, and therefore it offers the advantage of forgoing additional assumptions about the number of pension beneficiaries as required by the NPSF3. On the other hand, for a comparison of MMESP and the NPSF3 on the number of pension beneficiaries and its trend, and for greater detail regarding the types of National Pension benefits, their formulae, and other such matters, see Kwon et al. (2013c).

sources, however, we had to restrict ourselves to developing modules for forecasting long-term changes in wages and business income, on the one hand, and in the income tax revenue and fiscal spending on pension benefits, on the other.

In other words, the main task of our study is to estimate the individual tax burdens on wages, business income, and pension income that individuals earn and thereby forecast the income tax revenue for each year, given the likely changes in personal income and household makeup over time.

〈Figure 2〉 Content and Scope of K.Sim



The scope of this study is therefore narrower than Figure 2 suggests. Nevertheless, the tax revenue from wage and pension income makes up 69 percent of total income revenue (KRW 47.8 trillion as of 2013). Moreover, the income tax revenue also plays a decisive role in determining the amounts of revenue from other tax items dependent on disposable household income.[15)]

The simulation process of K.Sim involves the following. The most important factor we need to consider in estimating personal income taxes in Korea is the extent to which diverse income deductions affect the resulting tax revenue. Since the datasets or recipient files generated by MMESP only offer information on personal attributes, income levels, and deduction-eligible household compositions, we need additional data from other sources to determine the effects of special tax exemptions and deductions applicable to individual or household consumption.

This study draws such data from the donor files of HIES using

---

15) The tax revenue, which includes indirect taxes (e.g., VATs and excise taxes), national taxes (e.g., stock trading and capital gain taxes), and real estate-related local taxes (e.g., property and acquisition taxes), rests on certain assumptions regarding the changing patterns of spending and saving (investment) throughout an individual lifespan. The changing patterns of individuals' spending and saving, in turn, reflect different levels of disposable income determined by the given income tax regime. This study's K.Sim model may not take into account the entire structure and amount of government revenue, but it is significant as a first step toward the development and completion of a more refined microsimulation model necessary for research on taxation and fiscal policies.

the technique of propensity score statistical matching (PSSM). A type of statistical matching,[16] PSSM is used to create a control group similar to the participating group in situations where experimentation is impossible. By matching the two groups using the single variable of the propensity score, PSSM can help researchers avoid the problem of dimensions often associated with multivariate analyses. The more popular technique of predictive mean matching (PMM) similarly helps researchers avoid the problem of dimensions, but PSSM is the more practical tool to use.[17]

16) The basic framework for statistical matching is as follows (Rässer, 2002). First, let us suppose two datasets, A and B, are two different samples extracted from an identical population. Let us label the group of attributes found in both datasets Z. Let us suppose that information X pertains to dataset A only, and information Y, to dataset B only. What we need is a new dataset providing both types of information, X and Y, as well as Z. If we do not have such a dataset, we need to create a new dataset, C, by integrating the datasets A and B. Now, in our case, let us equate A and B with the recipient files and the donor files, respectively, so we are able to transfer information Y on the donor files to the recipient files. In this process, we need to use a distance function that allows us to measure similarity between the two types of files. Here the distance function, based on Z, is used to find donor file records that are identical or most similar to a given recipient's file records. Once our integrated dataset, C, is completed through this matching process, we can treat C as a single representative sample of the given population and proceed with our statistical analysis. This statistical matching process rests upon the following assumptions. First, X, Y, and Z are multivariate probability variables in the density function fXYZ, and there is no single dataset encompassing all the three. Second, recipient and donor file records will be extracted at random, independently of each other, from the same given population. In other words, for a common variable, Z = z, records from the two datasets, X and Y, can be combined only when the two datasets are contingently independent of each other. This is called the contingent independency assumption (CIA).

17) Gu and Rosenbaum (1993) demonstrated that propensity score matching can

More specifically, a propensity score can be defined as the contingent probability of the ith record from the recipient files (T = 1) with the covariate vector at Z = z (Equation (3)). The propensity score estimate can be obtained using Equation (4).

$$e(z_i) = P(T=1|Z=z_i) = g(z_i{}'\beta). \quad \text{(3)}$$

$$\hat{e}(z_i) = g(z_i{}'\hat{\beta}) = \frac{1}{1+e^{-z_i{}'\beta}}. \quad \text{(4)}$$

The matching process used in this study can be summarized as follows. First, the information we need to estimate the amount of income taxes—such as per-household liability insurance premiums, medical expenses, religious donations, household members' age distribution, occupations, educational attainments—is drawn from the donor files, so that we can estimate the per capita education cost for different age groups, including preschool children, elementary and secondary-school students, college students, graduate students, and adults. Second, the household structure in the recipient files—centered on the current household heads—is converted into a structure of taxable households. Taxable households are a new technical category of households we have created to find households with parents, regardless of whether the household heads are

create more balanced matching samples than either when the Mahalanobis distance functions were used or when the Mahalanobis caliper propensity scores were used. For more on PMM, see Robaly, Carsso and Saleem (2005).

living with their parents or not, who meet the income and age criteria of eligibility for basic deductions. Third, we bring together the recipient files (T = 1) and the donor files (T = 0) and create common variables for estimating propensity scores. These common variables include relationship to household heads (i.e., household head themselves, spouses, direct descendants, and direct ascendants); sexes; schooling years; age and the square of age; personal wage income; personal business income; and household income. Fourth and finally, we perform a logistic regression analysis using these common variables as explanatory variables and proceed with the matching process as follows.[18)]

$$\|P_i - P_j\| < \varepsilon, \quad j \in I_0, \ \varepsilon \leq 0.25\sigma_P$$

Where P_i : propensity score of ith record in the recipient file,
P_j : propensity score of jth record in the donor file,
I_0 : a group of donor file records, ε: caliper, σ_p : propensity score SD.

Here we match the recipient files from the period 2006 to 2013 with the corresponding donor files of the same period. For the period of 2014 and beyond, we match the amounts of income and consumption in the donor files from 2013 with those of the recipient files on the basis of the NPSF3's forecasts

---

18) We used psmatch2.ado of STATA for our matching process, according to the rule of k-nearest neighbors matching with $\varepsilon = 0.01$. For more on this process, see Leuven and Sianesi (2003).

on the nominal wage growth rate and the inflation rate.[19)]

Once the matching process is completed for each year, we estimate the amount of personal income for that year using the following process. First, we estimate the wage income (for wage earners) by applying and subtracting income deductions from the amount of wage income. For business income, we assume that the resulting estimate itself is business income from which all necessary expenses have been subtracted.[20)] Next, for pension income, we refer to the individual histories of recipients participating in, and receiving benefits from, the National Pension Scheme to estimate the non-taxable pension income, such as survivor and disability pensions, pension income upon premiums paid prior to January 1, 2002, and old-age pension benefits accrued prior to 2009. Then we subtract this amount of non-taxable pension income from the total amount of public pension income and apply the applicable pension income deductions. Next, from the amount of composite income (combining all the three types of income), we deduct personal (basic, additional, and multiple children) and special (for wage-earners only) income deductions on the basis of the

19) Given the significant differences in the observed values of the recipient and donor files (with the observed values of the donor files significantly smaller), we had to resort to sampling with replacements for our matching process.

20) The KLIPS panel provides data on the wage income/wage and sales revenue of unpaid workers. Here the former is equivalent to the net income (from which all business expenses have been removed). Because this study bases its income equations on the KLIPS panel, this study treats the income of unpaid workers as business income.

household makeup and spending patterns,[21] thereby estimating the tax bases to which different tax and deduction rates are to be applied in the predefined order. Having estimated the amount of per capita income tax, we then estimate the amounts of EITC and CTC (effective after 2014, with first benefits provided in 2015) that will be paid in the year following the year of imputation—pursuant to the Special Act on Tax Spending—according to the eligibility criteria of each given year.[22]

21) The deductions that were not included in the analysis of this study include the personal deductions for persons on the National Basic Living Security Program, deductions on interests on the old-age housing pension, special deductions on monthly rents for wage-earners, deductions on pension savings, and deductions on credit card spending.

22) For the years after 2014, we estimated the amounts of taxes according to the anticipated changes under the new tax law. For the specifics of the tax law applicable to our analysis period, see Appendix Tables 1 and 4.

3 Validation test of resulting estimates

1. Validation test of income distribution estimates
2. Testing the validity of income tax estimates

3 Validation test of resulting estimates

Given the nature of simulation tests, the resulting values and figures can differ from simulation to simulation, so we need to test and ascertain the validity of estimates resulting from our simulations. Validation tests also allow us to identify endogenous and exogenous errors that might exist, which we can then correct and thereby enhance the validity of our analysis. Based on the advice of Citro and Hanushek (1997) to researchers and developers of various simulation models, we arrived at several guidelines as follows. First, we should provide accurate estimates of policy outcomes that are not divergent from the reality. Second, we should provide the uncertainty and hypothesis-sensitivity of our resulting estimates as outcomes of our models. And third, we should stay up to date with the best knowledge and information concerning our models.[23] This is why, for instance, macrosimulation models are required to test the sensitivity of their results against changes in major hypotheses concerning decision-makers at levels higher than the micro-level.

Testing the validity of microsimulation results, however, can

23) For more on this topic, see Kwon and Han (2009).

be far more complex and difficult. A microsimulation model like ours is composed of different, yet interrelated, modules. The sensitivity of our results to changes in certain hypotheses or assumptions therefore may not mean much.[24] Those conducting microsimulation studies, therefore, test the validity of their results by comparing them to the observed values from reality or the results of other studies with similar objectives.[25]

In an effort to test and ascertain the validity of our model, we compare our results to the actually observed distribution of wage and business income from 2006 to 2012, as reported in Statistical Korea's HIESs and also in the KLIPSs of the Korea Labor Institute (KLI). Then we review the stability of the distribution of estimated income for the period of 2013 to 2100. Finally, we compare our income tax estimates to the actually observed values reported by the National Tax Service (NTS).[26]

---

24) For more on the realistic alternatives to the validation tests of microsimulation results, see O'Donoghue (2001) and Toder et al. (2002).

25) For instance, Caldwell and Morrison (2000) compare the results of DYNACAN of Canada and CORSIM of the United States with the results of other fiscal projection models. The Urban Institute of the United States also compares the results of its life event simulation model—named MINT and developed upon request from the US Social Security Administration (SSA)—with outside data. Of course, some studies test the sensitivity of their results from certain modules. The majority of microsimulation researchers and authors, however, believe that such comparisons are the most general method for testing the validity of simulation results.

26) Comparing the fiscal projection results illustrated in Figure 11 of this study and the results illustrated in Figure 5 of Kwon et al. (2013) is a sensitivity test of sorts. The former illustrates the forecasts of the NPSF3 based on macroeconomic and demographic variables, while the latter illustrates population projections based on forecasts made by Park et al. (2013) and Shin et al. (2012). Both figures also illustrate differences resulting from the

1. Validation test of income distribution estimates

In this section, we shall test the validity of our income distribution estimates by comparing them to the actually observed distributions of personal wage and business income as reported in Statistics Korea's HIESs and KLI's KLIPSs.[27] Before we proceed, however, the following needs to be pointed out.

First, the HIESs investigate and report the personal wage and business income of household heads and their spouses separately, but report the income of other household members as a sum. For our comparison, we estimate the income of each household member other than the household head or his/her spouse by dividing the sum of other household members' income by the number of other employed household members.

Second, our comparison does not include single-person households in the scope of analysis, which the HIESs have done since 2006. Also, we do not reflect possible differences that may result from the number of major changes that have been taking place in technicalities and methodologies of the HIESs,

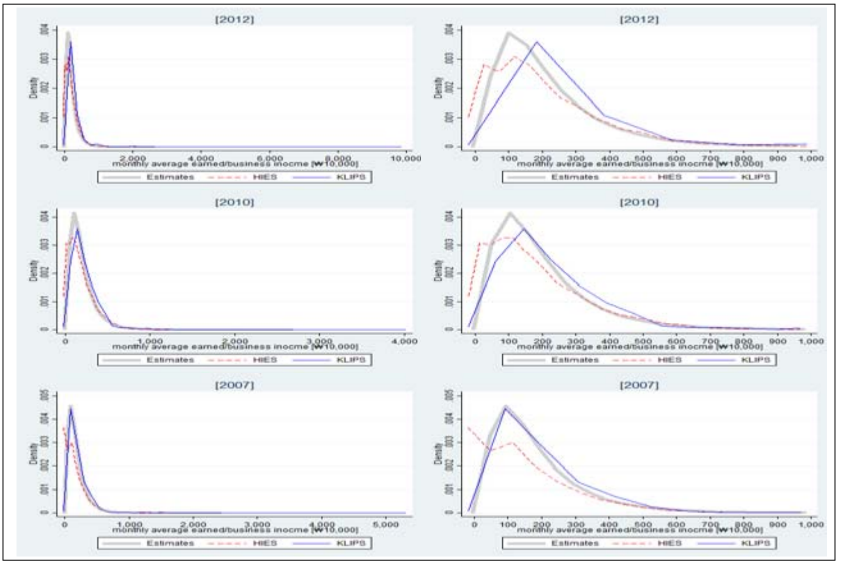
---

errors that have been identified and corrected. Compare the pension fiscal deficit and the forecast year of the depletion of the pension fund. This study predicts that the fiscal balance involving the pension will turn into deficit by 2038 and the pension fund will be depleted by 2052, while Kwon et al. (2013) predicts these events to occur in the years 2043 and 2058, respectively. The NPSF3 also predicts the years 2044 and 2060.

27) The average monthly income, reported by the KLIPSs, involves variables p [Survey Number] 1642 for wage earners and p [Survey Number] 1672 for unpaid workers.

such as the reform of multipurpose connected samples in 2007, the reform of the sorting system in 2009, and the change of the annual data-collecting system in 2011.[28)]

〈Figure 3〉 Comparison to Other Income Distribution Estimates



Notes: 1) Average monthly wage and business income, at current price (in KRW 10,000).
2) Income values reported in the HIESs and the KLIPSs are cross-sectional values that take into account the given weights.
3) The averages were estimated by excluding persons whose wage or business income is zero.
4) The income distributions are kernel density functions smoothed using STATA's Epanechnikov.

Sources: 1) Statistics Korea, Household Income and Expenditure Surveys, each year; KLI, Korean Labor and Income Panel Studies, each year.
2) Estimation results.

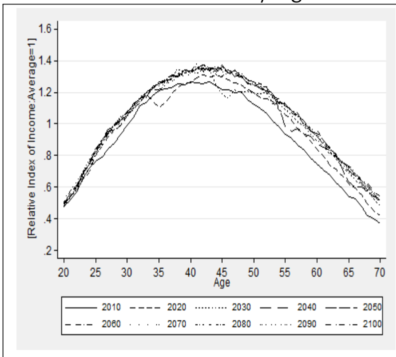
28) A thoroughgoing comparison would have to take into account all the relevant attributes of the given data, but it can be a costly and cumbersome task. This study therefore leaves such a task for future research and mitigates possible shortfalls by including the KLIPSs' labor panel in the scope of comparison.

Figure 3 compares the income distributions of 2007, 2010, and 2012 as reported or estimated in the three sources. The graphs on the left side show the overall income distribution, with the typical long tail drawing toward the right. The right tail is the longest in the KLIPSs' graphs and the shortest in the HIESs'. The density of low income levels is relatively greater in the HIESs' graphs than in other sources. The graphs on the right side capture the density of income levels at KRW 10 million a year or below with greater specificity. Our estimates approximate the KLIPSs' graphs, which tend to move to the right of our estimates in the recent years. However, we ought not to overlook the relative brevity of the time series subject to our comparison.

Next, we review the stability of the long-term future income distributions. Figure 4 charts the distribution of income by age for every 10 years, as measured against the overall average income distribution. The graphs show that the older one gets, the higher one's income level becomes. The income level reaches its relative peak at age 44-45, and begins to decline at age 50 and over. This is the typical hump-shaped age-and-income profile.

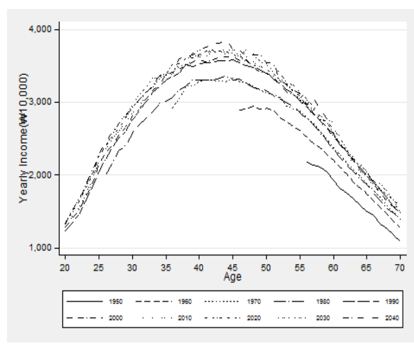
Figure 5, on the other hand, charts the distribution of average annual wage and business income for persons aged 20 to 70, for every 10 birth years between 1950 and 2040. These birth cohort graphs also show the income level to reach its peak at age 44-45. The reason the income level of persons born prior to 1980 stays lower than that of persons born in and after 2000 may have to do with the increasing education that later generations receive. The decrease in the size of the school-age population raises the college enrollment rate significantly, but the current income estimation process does not reflect endogenous differences among college students. This may explain the relatively higher income levels of later generations. [29]

〈Figure 4〉 Wage and Business Income Distribution by Age



Note: The averages were estimated by excluding persons whose wage or business income is zero

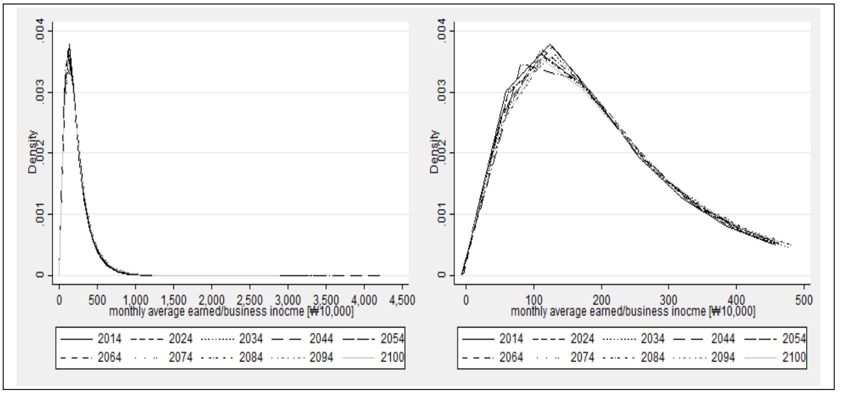
〈Figure 5〉 Wage and Business Income Profile by Birth Cohort



Notes: 1) Based on the constant price of 2012, converted using the nominal wage growth rate.
2) The averages were estimated by excluding persons whose wage or business income is zero

29) This is a matter that warrants further research and analysis with the future improvements to be made to our model.

〈Figure 6〉 Changes in the Wage and Business Income Distributions



Notes: 1) Average monthly wage and business income, converted into the current price as of 2012 using the nominal wage growth rate.
2) The averages were estimated by excluding persons whose wage or business income is zero.
3) The income distributions are kernel density functions smoothed using STATA's Epanechnikov.

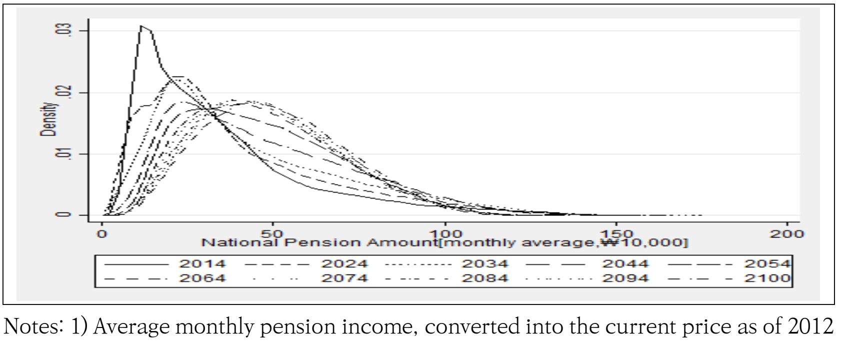
Figures 6 and 7 chart the long-term changes in the distribution of different types of income for every 10 years of the simulation period.30) Figure 6 shows the distribution of wage-earners' wage income and unpaid workers' business income, estimated in our simulations and converted into the current price for 2012 using the nominal wage growth rate. As the

30) To estimate long-term changes in income distribution, we need to convert our estimates into a current price. Yet the pattern of change in the distribution varies depending on which variable—the nominal wage growth rate or the inflation rate—we use as the discount variable for our conversion. As changes in income levels are closely related to economic growth and inflation, using the inflation rate can allow us to capture changes in income as results of economic growth. These results, however, are difficult to chart in long-term graphs. In this study, therefore, we use the nominal wage growth rate to convert our estimates into a current price. For estimates converted using the inflation rate, see Appendix Figures 1 through 3.

overall income distribution (left) and the distribution of income at KRW 5 million or less (right) both show, our simulations cause little change to the income distributions. However, there are small year-by-year variations observed in the lengths of the tails of the graphs toward the right, which likely reflect the effects of error terms on income in our simulations.

Figure 7 shows the distributions of National Pension benefits of all types (i.e., old-age, disability, and survivor benefits) in our simulations. Note that although we converted our estimates into the current price using the nominal wage growth, the center of distribution moves only gradually to the right in the early years, and it remains more or less the same from around 2040 onwards. In other words, as the National Pension system matures, individuals participate in it for longer periods of time and earn greater pension income over time.

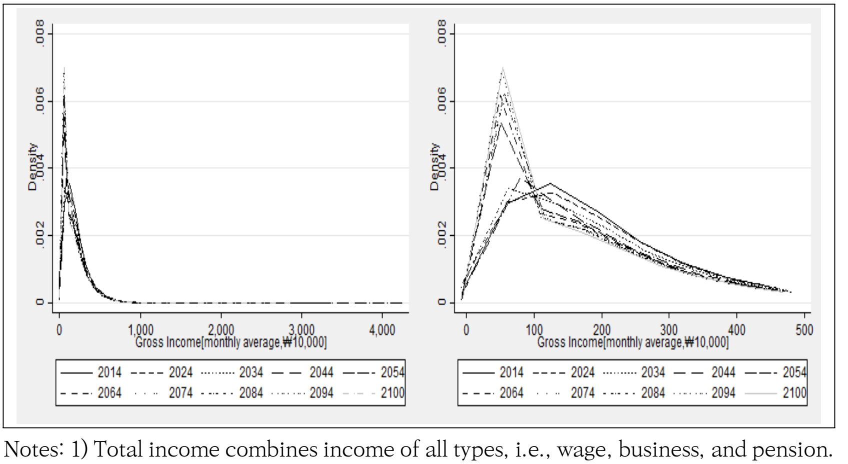
〈Figure 7〉 Changes in the Pension Income Distribution



Notes: 1) Average monthly pension income, converted into the current price as of 2012 using the nominal wage growth rate.
2) The averages were estimated by excluding persons whose pension income is zero.
3) The income distributions are kernel density functions smoothed using STATA's Epanechnikov.

Figure 8 shows the distribution of total income, combining all the three types—wage, business, and pension—of income. In general, earners of wages and business income are in the working-age groups, while pension income earners are in the post-retirement groups. Our microsimulation model, however, also shows income estimates for old-age pension earners with wage or business income and for recipients of disability and/or survivor pension business with wage or business income. For our purposes, we thus define total personal income as a sum of income from both pension and non-pension sources.

〈Figure 8〉 Changes in the Total Income Distribution



Notes: 1) Total income combines income of all types, i.e., wage, business, and pension.
2) Total income for each year was converted into the current price as of 2012 using the nominal wage growth rate.
3) The averages were estimated by excluding persons whose total income is zero.
4) The income distributions are kernel density functions smoothed using STATA's Epanechnikov.

Figure 7 shows that the center of the total income distribution tends to move toward the left over time, indicating an increase in the density of low-income earners. This is because the number of pension earners will increase over time, with the maturity of the National Pension Scheme, whose average income will be relatively lower than other types of income. The amounts of pension benefits, after the initial benefits, are arrived at by only using the inflation rate (minus the economic growth effect). Figure 8, however, shows the changes more noticeably, as the amount of pension income was converted into the current price using the nominal wage growth rate.[31)]

2. Testing the validity of income tax estimates

Here we shall review the validity of the income tax estimates our microsimulation model has produced.[32)] Table 3 shows the

31) Converting total income into a current price using the Consumer Price Index (i.e., the inflation rate) shows that the overall income distribution moves slowly toward the right over time (Appendix Figure 3).

32) We omitted the period of 2006 to 2010 from our income tax estimation process for a number of reasons, including the great volatility of the market attendant upon the onset of the global financial crisis, and, most importantly, the great problems we have found with the raw data for the HIESs of 2006 through 2009 in terms of the sampling process, weights, personal attribute coding, and so forth. The HIESs began to include single-person households in their scope in 2006, and they underwent significant changes afterward, including the replacement of the continued samples in 2007, the reform of the sorting system in 2009, and the change of the income and spending data collection method in 2011. In the meantime, the raw data for the years prior to 2009 were rife with absurd errors, such as household heads in their 30s whose grandparents were aged zero to 10 and seniors in their 90s that were categorized as “descendants” of household heads in their 30s. Our consultation with Statistics Korea

income tax estimates based on the actual tax revenue reported in the years 2011 to 2013 and on the estimates in the HIESs and K.Sim.[33] The estimates were given year-by-year simulation tests, but do not reflect all the decisive factors, such as the amounts and rates of income deductions under the Special Tax Reduction and Exemption Control Act (STREC) and penalty taxes. Table 3 shows that the estimates of both the HIESs and K.Sim are mostly within a range of three-percent from the observed values in reality. While different individuals earn, report, and pay taxes on their income at different periods of time, the presence of such arrangements as interim prepayments and withholding taxes helps to mitigate the effect of these differences. The estimated tax amount is similar to the balance of the determined tax amount, from which the amount of prepaid tax has been subtracted, according to the guidelines of the Yearbooks of National Tax Statistics. The amounts of determined taxes tend to be higher than the actual amounts of taxes to be paid due to various deductions and exemptions. Our estimates hover slightly above the amounts of labor and composite income taxes levied and slightly below the sums of all the related taxes.

---

revealed that there were technical issues in the process of entering and cleaning raw data. We therefore decided to omit the HIESs for the years prior to 2010 from the scope of our study.

33) The estimates reported in the HIESs reflect income from rents and other sources and the personal deductions for persons on the National Basic Living Security Program, all of which were omitted from K.Sim.

〈Table 3〉 Tax Revenue and Estimates, 2010-2012 (Nominal)

(Unit: KRW 1 trillion)

<table>
<tr><th colspan="3"></th><th>2011</th><th>2012</th><th>2013</th></tr>
<tr><td rowspan="9">Actual collection</td><td colspan="2">Total national tax revenue</td><td>192.4</td><td>203</td><td>201.9</td></tr>
<tr><td colspan="2">Income tax revenue</td><td>42.3</td><td>45.8</td><td>47.8</td></tr>
<tr><td rowspan="5">Taxes on wage, business and pension income</td><td>Wage income tax (a)</td><td>18.4</td><td>19.6</td><td>21.9</td></tr>
<tr><td>Composite income tax (b)</td><td>8.3</td><td>9.9</td><td>10.9</td></tr>
<tr><td>Business income tax (c)</td><td>1.7</td><td>1.7</td><td>..</td></tr>
<tr><td>Other income tax (d)</td><td>0.908</td><td>0.981</td><td>..</td></tr>
<tr><td>Pension income tax (e)</td><td>0.002</td><td>0.006</td><td>..</td></tr>
<tr><td colspan="2">a+b</td><td>26.7</td><td>29.5</td><td>32.8</td></tr>
<tr><td colspan="2">a+b+c+d+e</td><td>29.3</td><td>32.2</td><td>32.8</td></tr>
<tr><td rowspan="2">Estimates</td><td colspan="2">HIES</td><td>27.1</td><td>30.4</td><td>32.7</td></tr>
<tr><td colspan="2">K.sim</td><td>27.3</td><td>29.7</td><td>33.3</td></tr>
</table>

Sources: NTS, Yearbooks of National Tax Statistics, each year; Ministry of Strategy and Finance (MSF, 2014a), Report on the Settlement of National Accounting Year 2013; K.Sim estimates.

NTS's Yearbooks, however, report only the simple processed sums of each category of taxes and do not show how the various factors not considered in our study are offset against one another. In order to test the validity of our estimates, we need to compare them to the actual distributions of global income tax burdens that include wage, business, and pension income taxes. Since no authority in Korea has made available official distributions on income tax distributions based on per-capita tax records, we approximated income tax distributions using data on the 100percentiles of wage and composite income taxpayers that the NTS recently disclosed.[34)]

34) We use the data that NTS submitted to the office of National Assemblyman Jonghak Hong on the global income, tax brackets, and calculated and determined tax amounts of the 100 percentiles of wage and composite income taxpayers for the years 2007 to 2012.

〈Table 4〉 Average Effective Tax Rates (AETRs) and Tax Concentrations, 2010–2012 (Nominal)

Year	DATA	ATER	Calculated tax concentration	Determined/estimated tax concentration
2010	100percentiles	7.0%	0.7938	0.8232
	HIESs	7.5%	0.7312	0.7620
	K.Sim	7.7%	0.7030	0.7382
2011	100percentiles	7.2%	0.7957	0.8245
	HIESs	7.9%	0.7392	0.7709
	K.Sim	8.0%	0.7050	0.7401
2012	100percentiles	7.3%	0.7915	0.8211
	HIESs	8.2%	0.7289	0.7604
	K.Sim	8.2%	0.6944	0.7291

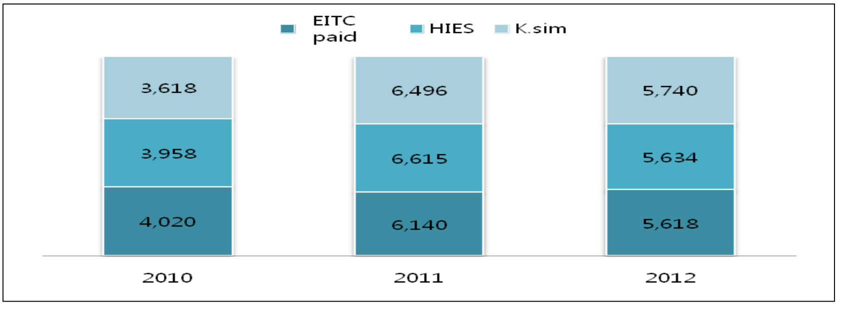
Sources: NTS, Combined Income Data on 100 Percentiles (2010-2012); K.Sim estimates.

Table 4 shows the distributions of calculated, determined, and estimated tax amount concentrations, as well as the AETRs (i.e., ratios of calculated tax amounts to income amounts) for the years 2010 to 2012. While different studies define the concept and scope of income differently, our simulations show estimated rates to be relatively higher and tax concentrations to be lower than the actually reported data. This reflects the inherent limits to the micro-survey data used in our simulations. The lack of certain types of information has led to the under-application of possible deductions, while the omission of high-income groups from the samples of the micro-survey data has downsized the estimated total tax amount. As Figure 9 shows, however, our estimates are quite close to the actually observed values on wage income tax credits (EITCs).

In sum, the lack of available data limits the extent to which our estimates reflect our institutional and economic reality. But it is not enough to render our study and its long-term scope invalid.

〈Figure 9〉 EITC Distribution. 2010-2012 (Nominal)

(Unit: KRW 100 million)



4 Long-term forecasts on income taxes: population aging and increasing pension income

1. Baseline forecasts
2. Income tax revenue forecasts based on adjustments to tax brackets
3. Rising National Pension contribution rate and the income tax revenue forecasts

4 Long-term forecasts on income taxes: population aging and increasing pension income

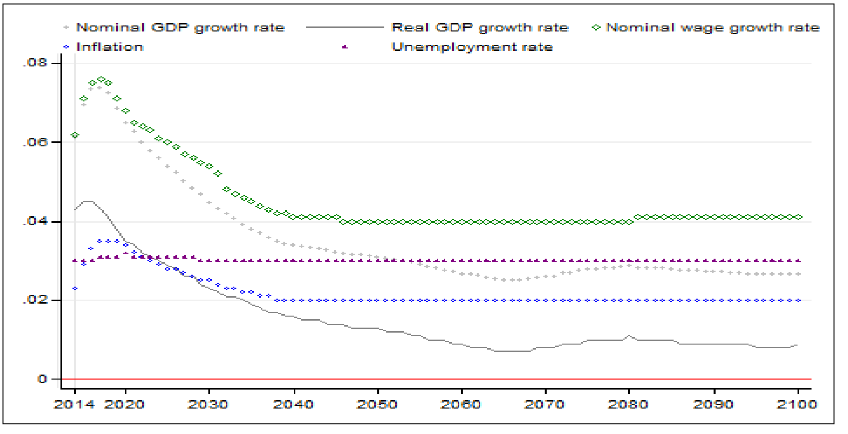
1. Baseline forecasts

Now we need to review the results of our income simulation results concerning the period spanning from 2014 to 2100. Based on the income tax law effective as of 2014, we run simulations on wage, business and pension income tax distributions and compare the tax threshold adjustment methods used. Then we analyze how the rises in the pension contributions would affect the tax revenue.

The macroeconomic forecasts and other such estimates used in our simulations are taken from the NPSF3 (Figure 10). Note that we assume the indivisibility of labor supply, i.e., that labor supply is so inelastic a factor with respect to changes in tax burdens that we can simply disregard it in our simulation tests.35)

35) The traditional supply-centered view holds that the rise in after-tax income induces increases in working hours and the workforce participation rate. Other theoretical studies have yet to reach a consensus on this point, but they generally show that the effect of labor supply depends on the sizes of income and substitution effects. In static labor-leisure choice models, labor supply is expressed as a curve, without the income effect and drawing upward toward the right. Supply-centered economists therefore argue that the government ought to lower taxes on labor in order to increase incentives for individuals to work harder. However, such drops in the income tax rates can be effective only when we posit a labor supply curve whose substitution effect overwhelms the income effect and which draws upward toward the right. Contrary to this textbook expectation, numerous

〈Figure 10〉 Economic Variables and Hypotheses



Source: NPPC (2013).

Our baseline forecasts are based on the assumption that no change by way of conscious policy decisions will take place from 2014 to 2100. Table 5 thus shows that the AETR, which we define as the estimated tax amount divided by before-tax income, will rise from 3.6 percent in 2014 to 34.2 percent by 2100. With income earners of all classes thus paying income taxes at increasingly high rates, the progressivity of the Korean

empirical studies have produced labor supply curves that either draw downward toward the right or that are refracted toward the right (Dickinson, 1992; Chung-Lin, 2003). Given the indivisibility of labor supply, the majority of households benefitting from lowered income tax rates are indeed those of low-skilled and lowly-paid workers who are comparatively more bound to the employment contracts they have signed than more skilled workers. High-income earners are mostly self-employed or involved in professions. They have a greater ability to lower their effective tax burdens either by incorporating their businesses or readjusting their income portfolios. Given this reality, labor supply may be very inelastic to the changes in the tax burdens, either in the form of direct changes to the income tax system or rises in the real tax bases.

income tax system declines drastically to below 0.1 percent by 2060 and to almost zero by 2100.[36] In particular, the ratio of tax-exempt persons will plummet from 29.5 percent in 2014 to zero percent in 2060 and after. The net redistribution effect of income taxes will continue to increase until 2040 or so, but will plummet drastically after, nearing almost zero by 2090.

〈Table 5〉 Baseline Forecasts, 2014-2100

Year	Before-tax income Gini	Tax progressivity	Net redistribution index	AETR	Ratio of tax-exempt persons	Share of nominal GDP
2014	0.3793	0.3069	0.0192	3.6%	29.5%	2.7%
2020	0.3764	0.2672	0.0273	6.1%	14.7%	4.5%
2030	0.3760	0.2090	0.0369	10.9%	4.8%	7.7%
2040	0.3828	0.1611	0.0401	15.6%	1.8%	10.8%
2050	0.3842	0.1195	0.0374	20.0%	0.6%	13.7%
2060	0.3843	0.0795	0.0296	24.2%	0.2%	16.7%
2070	0.3852	0.0497	0.0207	27.7%	0.0%	19.2%
2080	0.3842	0.0290	0.0129	30.6%	0.0%	21.3%
2090	0.3821	0.0148	0.0066	32.7%	0.0%	22.9%
2100	0.3855	0.0054	0.0020	34.2%	0.0%	24.0%

Notes: 1) Confined to persons with wage, business, or pension income.
2) Here "tax progressivity" is defined as the Kakwani progressivity index, i.e., before-tax income Gini coefficient, subtracted by the tax concentration index.
3) The net redistribution index is obtained by subtracting the after-tax income Gini coefficient from the before-tax income Gini coefficient, and therefore, itequals the balance of subtracting the ranking change index from the vertical equality index.

Source: K.Sim estimates.

These forecasts are based on the rather extreme assumption that the current nominal taxation system will remain intact in

36) The distribution indices used in our study are all based on the S-Gini coefficient, assuming the ethical weight parameter of $v=2$. For the specifics of the equation, see Ko(2013).

all its details into the future, with all deduction amounts and rates, limits, and tax brackets fixed to their current values and taxes imposed on nominal and not real income. Under this hypothesis, the forecasts listed above represent the maximum margins by which the income revenue will rise in the future, notwithstanding the aging of the Korean population, due to the nominal bracket creep or the nominal fiscal drag.

It is quite unlikely for these baseline forecasts to materialize in the future. NABO (2012) and Seong (2012) therefore either hypothesize that the per-capita tax burden of each age or sex group will remain constant or adjust the tax brackets in relation to the GDP growth before estimating income revenue, thereby inhibiting the extreme bracket creep effect shown in Table 5. Nevertheless, we include our baseline forecasts in this study to distinguish between effects of such factors as the increase in the number of pension earners, the increase in the amount of pension income, and the decrease in the number of dependent children, on the one hand, and the effects of continuous tax brackets, on the other. Moreover, we assume that these factors will also affect per capita tax rate. Therefore, we refrain from assuming little change in the effective per capita tax rate, focusing instead on the greater likelihood that the declining birth rate and population aging will significantly change the effective per capita tax rate.

Historically, however, the Korean government has actively intervened before these effects materialized, by either raising the deduction ceilings or readjusting the sizes of tax brackets. It would be economically and politically foolish if not impossible to retain the current nominal taxation system, which, if unaltered, will radically increase tax burdens on individuals. So policymakers will have to make some changes and adjustments on a regular basis. Thus, we need to review our income revenue estimates by taking into account two possible policy alternatives.

2. Income tax revenue forecasts based on adjustments to tax brackets

We envision two different ways of readjusting tax brackets in reviewing our income tax revenue estimates. In Case 1, we assume that the tax brackets will be resized according to the CPI. Here we apply the consumer inflation rate to all our nominal variables to convert our estimates into the current price of 2014, and then we re-convert the tax amounts we have estimated, by applying the income tax law, back to the prices of the given years. This allows us to control the nominal fiscal drag resulting from net increases in consumer prices, and we can also to capture increases in the tax burden due to the growth of real income.

〈Table 6〉 Case 1: Tax Brackets Readjusted, 2014-2100

Year	Tax progressivity	Net redistribution index	AETR	Ratio of tax-exempt persons	Share of nominal GDP
2014	0.3069	0.0192	3.6%	29.5%	2.7%
2020	0.2820	0.0237	5.0%	19.0%	3.8%
2030	0.2468	0.0284	7.1%	10.4%	5.3%
2040	0.2245	0.0326	8.9%	6.2%	6.8%
2050	0.2095	0.0364	10.7%	3.6%	8.4%
2060	0.1903	0.0390	12.7%	2.1%	10.3%
2070	0.1719	0.0406	14.7%	1.3%	12.2%
2080	0.1522	0.0409	16.8%	0.7%	14.1%
2090	0.1299	0.0390	19.0%	0.4%	15.8%
2100	0.1089	0.0360	21.1%	0.2%	17.4%

Note: Confined to persons with wage, business, or pension income.
Source: K.Sim estimates.

Table 6 shows an increase in the tax rates that is significantly slower than the one we have seen in our baseline forecasts. While the overall tax structure, in terms of tax progressivity and the net redistribution effect, changes rather rapidly in the early years, the pace of change slows down in the later years to lag behind the baseline forecasts by four or five decades. As a result, the tax burden on individuals who earn KRW 16 million or less, at the current price of 2014, will increase only slightly, from 3.6 percent in 2050 to 5.7 percent in 2100. Nevertheless, the amount of income tax revenue as a share of the nominal GDP will consistently increase, from 2.7 percent in 2014 to 17.4 percent by 2100 or at 5.7 percent each year on average.

In Case 2, we readjust the tax brackets by linking the increases in tax burdens to the increases in the real income, or according to the nominal income growth rate (Table 7).

〈Table 7〉 Case 2: Tax Brackets Readjusted, 2014-2100

Year	Tax progressivity	Net redistribution index	AETR	Ratio of tax-exempt persons	Share of nominal GDP
2014	0.3069	0.0192	3.6%	29.5%	2.7%
2020	0.3029	0.0195	3.8%	27.9%	3.0%
2030	0.2937	0.0187	3.8%	27.5%	3.1%
2040	0.2917	0.0187	3.8%	27.5%	3.3%
2050	0.2962	0.0193	3.8%	27.5%	3.5%
2060	0.2921	0.0193	3.9%	26.6%	3.8%
2070	0.2934	0.0194	3.9%	26.9%	4.0%
2080	0.2947	0.0197	3.9%	26.8%	4.2%
2090	0.2927	0.0193	3.9%	26.6%	4.2%
2100	0.2924	0.0190	3.8%	27.4%	4.2%

Note: Confined to persons with wage, business, or pension income.
Source: K.Sim estimates.

Table 7 shows little change in the ratio of tax-exempt persons for the next 86 years, indicating that the overall tax structure will remain more or less constant throughout our simulation period. The assumption underlying Case 2 is the same as the assumption of constant tax burdens for age cohorts explicitly or implicitly endorsed in numerous other studies. However, in contrast to NABO (2012), Park et al. (2012), and Seong (2012) that show the amount of income tax revenue as a share of the nominal GDP declining over time, our forecasts show the share as steadily rising in the future.

We need to carefully interpret this difference as being the result of complex interactions among multiple factors. For instance, the amended tax law of 2014 that we use in our simulation is a subject of increasing controversy in Korea because of the increases in tax burdens it is likely to cause.

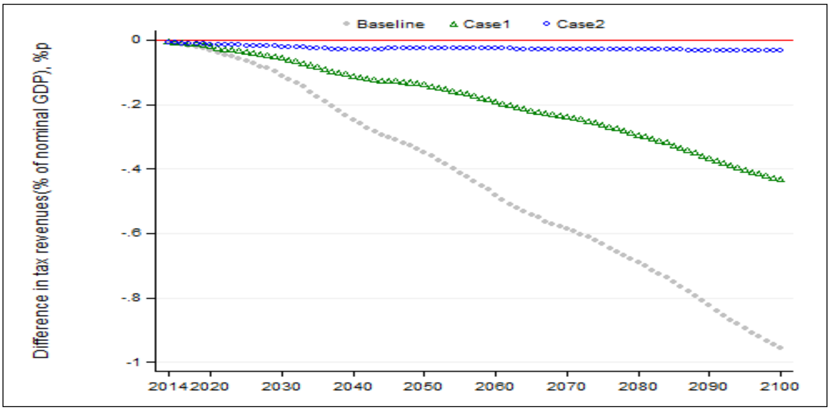
In the meantime, the increase in the number of seniors receiving pension benefits, the rise in their income levels, and the decrease in the number of dependent children due to the declining birth rate—all of which are the other key assumptions underlying this study—also serve to decrease the amounts of income deductions, thus boosting the tax-increasing effect of the amended tax law of 2014 even more.

As we expected, the Case 1 trend falls somewhere between the baseline scenario and Case 2, but it shows a significant divergence between when pension income is included and when it is not. The divergence is not as significant in Case 2. The main reason can be found in the different methods that are used to determine and adjust the amounts of National Pension benefits to be paid. Initial amounts of pension benefits and later amounts are determined according to different methods. The initial amounts reflect the inflation and the growth in the real income over the period of one's participation in the pension scheme. The later amounts, however, reflect the inflation rate only.

In Case 1, therefore, a great number of pension income earners will remain taxpayers at the point of receiving initial benefits and in subsequent years. That is why excluding pension income produces such a significant difference in forecasts. In Case 2, however, only pension earners who earn certain levels of income in pension benefits will become taxpayers, and they will gradually join the group of tax-exempt persons. This is because in Case 2, the tax brackets are linked to the nominal wage growth rate, while the amounts of pension benefits these persons receive will be readjusted according to the inflation rate, thus dropping in relative value.

How, then, can the amount of income tax revenue as a share of the nominal GDP keep growing, as shown in Table 7, despite the long-term decrease in the number of taxpayers as forecast in Case 2? See Figure 12. The declining birth rate and population aging in Korea will also decrease the number of deduction-eligible persons or dependents per taxpayer in the long term. Case 1 is especially noteworthy when the pension income is included, as it shows with the greatest level of clarity that the two demographic changes and the resulting increase in the number of pension earners will ultimately contract the amounts of deductions. Most other states worldwide that use the price indexation technique also use their respective CPIs as the basis for forecasts in this regard. The decrease in the number of deduction-eligible dependents per taxpayer will lead to a decrease in the amounts of deductions provided, which will ultimately raise the tax bases and rates for taxpayers. The overall income tax revenue, of course, will be subject to far more diverse factors aside from these, such as the distribution of taxpayers' income. Figure 11 therefore illustrates the margins by which the amount of the income tax revenue as a share of the nominal GDP will decrease, if we exclude pension income from our baseline scenario and Cases 1 and 2. The margins will differ in size depending on which policy choice is made, but can widen to as much as one percentage point between the baseline scenario and Case 1.

〈Figure 11〉 Gaps in Income Tax Revenue as a Share of Nominal GDP: Pension Income Excluded, 2014-2100



Source: K.Sim estimates.

In other words, the declining birth rate and population aging may mean a decrease in the number of taxpayers over the long term, but they also mean decreases in the number of deduction-eligible family members per taxpayer, and, as a consequence, decreases in the amounts of deductions (on education and medical costs, for example). The declining birth rate and population aging therefore will affect the income tax revenue both ways.

3. Rising National Pension contribution rate and the income tax revenue forecasts

Given the fact that the inevitable increases in fiscal spending on social security benefits will become major fiscal pressures on the Korean government, the NPSF3 and Shin et al. (2013) argue for increases in the rates of contributions to the National Pension Scheme.[37) However, we need to keep in mind that the entire amount of pension contributions is deductible from taxable income. According to the Tax Expenditure Budget Report of 2015 (Table 8), the aggregate amount of deductible pension contributions in Korea amounted to KRW 1.36 trillion or 0.7 percent of the total national tax revenue in 2013.

〈Table 8〉 Social Security-Related Tax Deductions

(Unit: KRW 100 million)

Deduction type	2013 (actual)	2014 (estimate)	2015 (budget)
Deductions on National Pension contributions	13,632	14,468	12,425
Special deductions on social insurance contributions	21,569	23,580	19,917
Deductions on out-of-pocket expenses for NHI, etc.	8,648	9,401	10,727
Deductions on employers' social insurance contributions due to increases in the number of hired employees	301	366	382

37) According to KIHASA's long-term fiscal projections concerning spending on social security (SOCX), budget spending as a share of the GDP will rise from 3.5 percent in 2013 to 5.7 percent by 2060, while social security spending will multiply from 6.3 percent to 23.2 percent.

Deduction type	2013 (actual)	2014 (estimate)	2015 (budget)
Total (A)	44,150	47,815	43,451
Total national tax revenue (B)	2,019,065	2,164,529	-
A/B	2.19%	2.21%	-

Notes: 1) Special deductions on social insurance contributions: Up to KRW 1 million per year on the National Health Insurance (NHI), the Long-term Elderly Recovery Insurance, the entire amount of the Unemployment Insurance, and other social insurances.

2) The figure given for the amount of deductions on employers' social insurance contributions (applicable in the case of small and medium businesses only) combines the deductions on both the income tax and the corporate tax.

Source: MSF (2014b), Tax Expenditure Budget Report 2015.

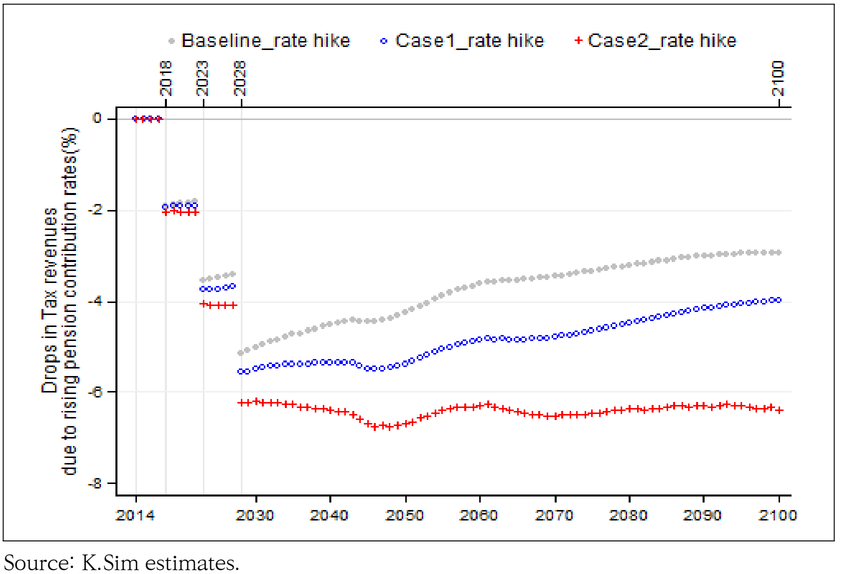
Now, we need to forecast how the rise in the National Pension contribution rate will affect income tax revenue. We thus borrow Scenario 3 on the rise in the social insurance contribution rate from Baek (2014), which assumes a rise that is sufficient to ensure the National Pension Scheme's payment capacity until 2100.[38]

However, raising the rate of contributions to the National Pension may not only increase the amounts of income deductions, but also lower the tax bases for individual taxpayers and thereby cut the applicable tax rates. As Figure 12 shows, raises in contribution rates can decrease the income tax revenue by up to 7.2 percent (after 2028, in Case 2). In the baseline scenario and Case 1, the amount of income tax revenue gradually returns to its normal size thanks to the increases in

38) More specifically, Scenario 3 assumes a contribution rate of nine percent for years 2013 to 2017, 11 percent for years 2018 to 2022, 13 percent for years 2023 to 2027, and 15 percent for years 2028 and beyond.

real income and tax burdens, notwithstanding increases in the contribution rates. In Case 2, however, the natural increases in tax burdens are inhibited by keeping the current tax burden rate constant. In this scenario, therefore, the income tax revenue keeps decreasing over time.

〈Figure 12〉 Trend in the Income Tax Revenue: Based on Increases in Contribution Rates, 2014-2100



Source: K.Sim estimates.

Furthermore, raising the rate of contributions to the National Pension can worsen the taxation structure in Korea. Table 9 shows that higher National Pension contribution rates ultimately lower the AETRs, thereby compromising the redistributive function of income taxes.[39] In the meantime, tax progressivity tends to decrease in the baseline scenario, but increases by contrast in Case 2. This reflects the regressive nature of the National Pension system, which places ceilings on the applicable income levels and imposes fixed rates of contributions, and the year-to-year changes in income distributions. This phenomenon will require deeper analysis in the future.

In sum, raising the contribution rate may boost the fiscal sustainability of the National Pension Scheme, but its overall effect may not be so welcome, as it could decrease the overall income tax revenue for the Korean government. These simulation results emphasize the fact that policymakers ought to consider all the related institutions and laws before making any conscious policy interventions.

〈Table 9〉 Raising National Pension Contribution Rates and Changes in Tax Structure, 2014-2100

(Unit: %p)

Year	Tax progressivity			Net redistribution index			AETR		
	Baseline	Case1	Case2	Baseline	Case1	Case2	Baseline	Case1	Case2
2020	0.03	0.05	0.05	-0.06	-0.05	-0.04	-0.12	-0.10	-0.08
2030	-0.02	0.01	0.10	-0.23	-0.17	-0.12	-0.54	-0.38	-0.24
2040	-0.07	-0.07	0.08	-0.25	-0.21	-0.12	-0.67	-0.46	-0.24
2050	-0.04	0.01	0.22	-0.23	-0.23	-0.13	-0.81	-0.56	-0.26
2060	-0.13	-0.02	0.16	-0.21	-0.23	-0.12	-0.81	-0.60	-0.25

39) Pfähler (1990) decomposes the redistributive effect of income taxes as the effects of average tax rates and tax progressivity, using the formula, average tax rate/(1-average tax rate)*(progressivity of PIT burden).

Year	Tax progressivity			Net redistribution index			AETR		
	Baseline	Case1	Case2	Baseline	Case1	Case2	Baseline	Case1	Case2
2070	-0.19	0.00	0.20	-0.20	-0.25	-0.12	-0.86	-0.68	-0.26
2080	-0.24	0.01	0.20	-0.19	-0.24	-0.12	-0.85	-0.73	-0.26
2090	-0.29	-0.03	0.16	-0.19	-0.23	-0.12	-0.83	-0.76	-0.25
2100	-0.32	-0.07	0.16	-0.19	-0.23	-0.12	-0.84	-0.79	-0.25

Note: Confined to persons with wage, business, or pension income.
Source: K.Sim estimates.

5 Conclusion

5 Conclusion

This study is the first in Korea to develop and apply a dynamic microsimulation model to the analysis and forecast of how population aging in the country will affect the amount of income tax revenue over the long term. In particular, the study focuses on the effects of the household makeup change (resulting from the declining birth rate and population aging) and the increase in the amount of per capita pension income (due to the maturity of the National Pension system) on future amounts of individual and household income and tax deductions. Our findings can be summarized as follows. First, assuming the current tax law (as of 2014) to remain constant without any policy changes made to tax brackets, the ratio of tax-exempt persons will plummet from the current 29.5 percent to almost zero percent by 2060. As a result, the amount of the income tax revenue as a share of the nominal GDP will multiply from 2.7 percent in 2014 to 16.7 percent in 2060 and to 24 percent by 2100. The extreme results of this simulation scenario illustrate the nature of policy interventions that need to be made. The question is how and to what extent.

Second, assuming that the upward pressure on the tax burden under the nominal taxation system will be inhibited by means of full price indexation, the amount of income tax revenue as a share of the GDP will increase to 10.3 percent by 2060 and to 17.4 percent by 2100, lagging about four or five decades behind our earlier baseline scenario. Next, adjusting the tax brackets according to nominal wage growth rate instead of the inflation rate will ensure the level of tax progressivity and the ratio of tax-exempt persons remain more or less constant over the next 86 years, with the amount of income tax revenue as a share of the nominal GDP rising only slightly to 4.2 percent by 2100. The simulations of these different scenarios reveal the importance of adjusting tax brackets in consideration of all the interrelated factors, including the amount of the income tax revenue to be obtained and the tax burdens to be imposed on individuals.

Third and most importantly, this study illustrates the necessity of basing future policy decisions regarding National Pension revenue on a number of inevitable demographic changes, including the declining birth rate and population aging, and on how these larger phenomena will change the number of persons eligible for basic deductions per household. The amount of per capita pension income may not increase much, but the foregoing factors will all serve to contract the amounts of income deductions and raise the tax bases, effectively increasing the income tax revenue in the long term. For instance, if we readjust the tax brackets according to the nominal wage growth rate, the decrease in the taxpayer population caused by the low birth rate and population aging may appear, at first sight, to decrease the amount of the income tax revenue before 2060. However, given the diversity of income tax deductions in Korea, the decrease in the taxpayer population will also serve to shrink the amounts of income deductions and thereby increase the resulting income tax revenue to an unexpected extent. Raising the National Pension contribution rate also warrants further debates and analyses, as it has significant implications concerning the decrease in income tax revenue and the redistributive effect of the income tax system in Korea.

The limits to the methodology used in this study directly reflect the shortfalls of the basic datasets used for our simulation model.[40)]

The datasets used include Statistics Korea's Population and Housing Census, KLI's KLIPSs, and Statistics Korea's HIESs. The KLIPSs and the HIESs are representative microsurveys in Korea that provide a comprehensive range of information on the income and spending patterns of Korean households using a 10-percent sample from the Population and Housing Census. The Census in Korea, however, still relies almost exclusively on outdated traditional survey methods that are rife with risks of data omission, which is why it is held in low regard as a framework for survey samples. The presence of significant survey er-

40) The microsimulation model used in this study is not free from the problems and issues often plaguing similar models. For more detailed discussions of these issues, see Kwon et al. (2013) and Kwon and Han (2009).

rors in this incomplete framework, such as the omission of high-income earners from the sample, therefore complicates the task of estimating the individual and household income and income tax burden distributions with great accuracy. The established literature in this regard converges on the conclusion that we may overcome sampling errors by enlarging the number of samples and adjusting the weights assigned accordingly. However, there are no known statistical solutions to non-sampling errors of the kind often associated with an incomplete sampling framework or inadequate survey methods. We will never be able to solve these errors unless the Korean authorities fully disclose micro-data for public administration, such as all individual taxpayer records.

The one-percent sample of the Population and Housing Census that Statistics Korea discloses to the public is also known to over-represent or under-represent certain cohorts. The undersampling problem is especially prominent with the age cohort of 18 to 22. The one- and two-percent samples from 2010 carry the same problem, as the Censuses conducted in these years exclude members of households currently absent due to academic or work-related reasons or military service. We may therefore be able to mitigate the effect of this problem by adding more samples of this age cohort to the currently published sample materials. We also need to additionally survey and disclose information on the sex and age distribution of the members of each age cohort, given the increasing importance of research on the long-term social and economic impact of population aging.

Notwithstanding the great processes and recent strides it has made, the dynamic microsimulation model is still plagued with accuracy issues and shortcomings. We need more research to overcome the inherent limits of this methodology. We hope this study will provide a meaningful first step toward that end.

References

NPPC, Fiscal Calculations for the National Pension, 2013: Long-term Fiscal Projections Concerning the National Pension, 2013.

NTS, Yearbooks of National Tax Statistics, each year.

NABO, Long-term Fiscal Projections and Analysis, 2012-2060, 2012.

Kwon, H., W. Shin, J. Han, D. Jeong, C. Park, J. Ryu and G. Kim, "Prospects and Evaluation of the Appropriateness of the Public Pension Reform of 2007: A Long-term Analysis Based on Dynamic Microsimulations," Journal of Social Security, 29(4), 2013, pp. 1-54.

Kwon, H. and J. Han, Research and Development of a Microsimulation Model for the National Pension, National Pension Research Institute (NPRI) Policy Report, 2009.

MSF, Report on the Settlement of National Accounting Year 2013, 2014(a).

-----------, Tax Expenditure Budget Report 2015, 2014(b).

Kim, H., E. Kim, and S. Jeon, Redistribution Effect of the Social Security System on Lifetime Income, 2010, KLI.

Park, Y., S. Jang, and S. Kim, "An Error-Correcting LC Model for Death Rate Forecasts," Journal of Surveys, 14(2), 2013, pp. 19-47.

Park, H., M. Park, H. Kim, et al., Analysis and Forecasts of Factors of Mid- to Long-term Changes in the Tax Revenue, 2012, KIPF.

Baek, H., "Analysis of Long-term Fiscal Prospects for the National Pension Using Different Scenarios on the Contribution Rate Increases," Korean Association of Public Finance Autumn Conference 2014, 2014, KAPF.

Seong, M., Long-term Forecasts on the Income and Consumption Tax Revenue: Based on Forecasts on the Future Distribution of Sources of Taxes, 2012, KIPF.

Seong, M., B. Jeon, and B. Jeon, Taxation and Fiscal Spending Simulation Model: Development of KIPFSIM08, 2008, KIPF.

Shin, G. and H. Kwon, Middle-Aged and Elderly Retirees and the Rate of Early Retirement Pension Benefits, NPRI, 2013.

Shin, H., J. Won, J. Heo et al., Research and Development of a Fiscal Projection Model for Social Security Programs, commissioned policy report, 2013, Ministry of Health and Welfare and KIHASA.

CBO, CBO's Long-Term Model: An overview, Background Paper, No.3235, 2009.

Chung-Cheng Lin, "A backward-bending labor supply curve without an income effect," Oxford Economic Papers, 55, 2003, pp.336-343.

Citro, C. F., & Hanushek, E. A. (Eds.), Assessing for Data, Research, and Models: Needs for Data, Research, and Models, 1997, National Academy Press.

Dekkers, G. and Belloni, M., "Micro simulation, pension adequacy and the dynamic model MIDAS: an introduction," Project AIM-Deliverable, 4, 2009.

Dickinson, D.,"Productivity shocks and real business cycles," Journal of Monetary Economics, 29, 1992, pp. 191-208.

Gu, X.S. and P.R. Rosenbaum, "Comparison of multivariate matching methods: Structure, distances and algorithms," Journal of Computational and Graphical Statistics, 2, 1993, pp. 405-420.

Harding, A., and Gupta, A. (Eds.), Modelling Our Future : Population

Ageing, Social Security and Taxation, 2007, Elsevier.

Ko, Jayee, "Inflation, Taxation, and Income Redistribution in Korea, 2005-2007," 9th APEA Conference, 2013, Osaka, Japan, Asia-Pacific Economic Association.

Leuven, E. and B. Sianesi, "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing," 2003, http://ideas.repec.org/c/boc/bocode/s432001.html.

Li, Jinjing. and Cathal O'Donoghue, "A survey of dynamic microsimulation models: uses, model structure and methodology," International Journal of Microsimulation, 6(2), 2013, pp. 3-55.

McCracken, M. and Anil Gupta, "The Search for Ideas among Existing Microsimulation Models," Chapter 17 in Harding, A., & Gupta, A. (Eds.), Modelling Our Future: Population Ageing, Social Security and Taxation, 2007, Elsevier.

O'Donoghue, C., "Introduction to the Special Issue on Dynamic Microsimulation Modelling," Brazilian Electronic Journal of Economics, 4(2), 2001.

O'Donoghue, C.. "Dynamic Microsimulation: A Methodological Survey," Brazilian Electronic Journal of Economics, 4(2), 2001.

Pfähler, W., "Redistributive Effect of Income Taxation: Decomposing Tax Base and Tax Rates Effect," Bulletin of Economic Research, 42(2), 1990, pp. 121-129.

Rässler, S. Statistical Matching. A Frequentist Theory, Practical Applications, and Alternative Baysian Approaches(Lecture Notes in Statistics), 2002, Springer.

Rohaly, J., Carasso, A., and Saleem M.A.,"The Urban Brookings Tax

Policy Center Microsimulation Model: Documentation and Methodology for Version0304," 2005, Tax Policy Center at http://www.taxpolicycenter.org/UploadedPDF/411136_documentation.pdf.

Zaidi, Ashghar., Harding, Ann. and Williamson, Paul. (Eds), New Frontiers in Microsimulation Modelling, 2009, Ashgate.

Appendix

〈Appendix Table 1〉 Composite Income Tax Deductions by Year

(Unit: KRW 10,000)

Deduction type			`06	`07	`08	`09	`11	`12	`13	`14
Personal deductions	Basic deductions	Self	100			150				
		Spouse	100			150				
		Dependents	100			150				
	Additional deductions	Elderly	100 (65 or older) 150 (75 or older			100 (70 or older)				
		Disability	200							
		Housewife	50							
		Single-parent	N/A						100	
		Small family	100 (single-person household), 50 (two-person household), repealed in 2007.							
		Young children at age 6 or younger	100							Repealed
		Birth/adoption	N/A		200					Repealed
		Multiple children	N/A	50+(number of children-2)*50			100+(number of children-2)*200			Repealed
Pension contributions	National Pension	Self	Entire amount							
Special income deductions for wage earners	Insurance premiums	National Health Insurance	Entire amount							
		Long-term Elderly Recovery Insurance	N/A		Entire amount					
		Unemployment Insurance	Entire amount							
		Other insurances	100							Repealed
	Medical expenses	Self	Entire amount at 3% of wage or above							Repealed
		Dependents	Up to 500 at 3% of wage or above			Up to 700 at 3% of wage or above				
	Education expenses	Self	Entire amount							Repealed
		Preschool children	Up to 200			Up to 300				
		Primary and secondary schools	Up to 200			Up to 300				

Deduction type			'06	'07	'08	'09	'11	'12	'13	'14
		Colleges/universities	Up to 700			Up to 900				
	Donations	Religious donations	Ceiling on deductible religious donations = amount of composite income – total amount of donations = amount of income*10/100+amount of income*5/100 or total amount of donations, whichever is smaller. Ceiling on other deductible donations = amount of income*15/100.							Repealed
Standard deductions	Wage earners		100							Repealed
	Income earned other than wage		60							

〈Appendix Table 2〉 Income Tax Rates by Year

Tax base	Until 2008	2009	2010-2011	2012-2013	2014 -
Up to KRW 12 million	8	6	6	6	6
Up to KRW 46 million	17	16	15	15	15
Up to KRW 88 million	26	25	24	24	24
Up to KRW 150 million	35	35	35	35	35
Above KRW 150 million					38
Above KRW 300 million				38	

〈Appendix Table 3〉 Income Tax Deductions

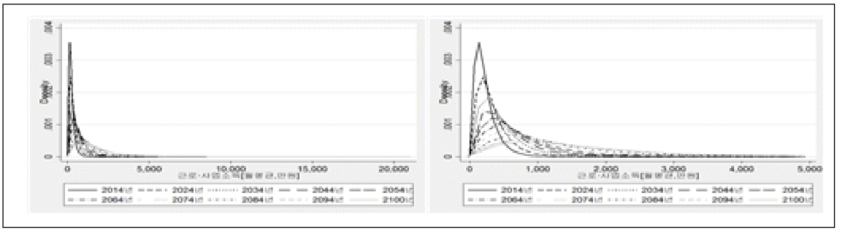
<table>
<tr><th></th><th colspan="2"></th><th colspan="2">2006-2013</th><th colspan="2">2014 and beyond</th></tr>
<tr><td rowspan="5">Wage income deductions</td><td colspan="2">Calculated tax amount</td><td>Deduction rate</td><td>Limit</td><td>Deduction rate</td><td>Limit</td></tr>
<tr><td colspan="2">Up to KRW 500,000</td><td>55</td><td>N/A</td><td>55</td><td>N/A</td></tr>
<tr><td rowspan="3">Above KRW 500,000</td><td>Up to KRW 55 million in total income</td><td rowspan="3">30</td><td rowspan="3">KRW 500,000</td><td rowspan="3">30</td><td>KRW 660,000</td></tr>
<tr><td>Up to KRW 70 million in total income</td><td>KRW 630,000</td></tr>
<tr><td>Above KRW 70 million in total income</td><td>KRW 500,000</td></tr>
<tr><td colspan="3">Dependent children deductions</td><td colspan="2">N/A</td><td colspan="2">KRW 150,000 per year per child + additional KRW 200,000 per year for every child after the first two.</td></tr>
<tr><td colspan="3" rowspan="4">Special deductions</td><td colspan="2" rowspan="4">N/A</td><td colspan="2">Insurance premiums:
Limit: KRW 1 million, deduction rate: 12/100.</td></tr>
<tr><td colspan="2">Medical expenses: above 3% of total wage
Limit: KRW 7 million, deduction rate: 15/100.</td></tr>
<tr><td colspan="2">Education expenses:
Limit: KRW 9 million per year per college/university student, KRW 3 million per year for each student, entire amount of expenses on special education.
Deduction rate: 15/100.</td></tr>
<tr><td colspan="2">Religious donations: with limits, up to 15/100 of donations.</td></tr>
<tr><td colspan="3" rowspan="2">Standard deductions</td><td colspan="2" rowspan="2">N/A</td><td colspan="2">KRW 120,000 per wage earner.</td></tr>
<tr><td colspan="2">KRW 70,000 on earned income other than wage.</td></tr>
</table>

〈Appendix Table 4〉 EITC and CTC Eligibility Criteria

Type	Year	Household criteria	Income criteria	Formula
EITC	2008-2010	With at least one child	Less than KRW 8 million	Total income*15/100
			Less than KRW 12 million	KRW 1.2 million
			Less than KRW 17 million	(KRW 17 million-total income)*24/100
	2011-2013	With no child	Less than KRW 6 million	Total income*7/60
			Less than KRW 9 million	KRW 0.7 million
			Less than KRW 13 million	(KRW 13 million-total income)*7/40
		With a single child	Less than KRW 8 million	Total income*7/40
			Less than KRW 12 million	KRW 1.4 million
			Less than KRW 17 million	(1,700-total income)*28/100
		With two children	Less than KRW 9 million	Total income*17/90
			Less than KRW 12 million	KRW 1.7 million
			Less than KRW 21 million	(2,100-total income)*17/90
		With three or more children	Less than KRW 9 million	Total income*2/9
			Less than KRW 12 million	KRW 2 million
			Less than KRW 25 million	(2,500-total income)*2/13
	2014 and beyond	Single-person household	Less than KRW 6 million	Total income*70/600
			Less than KRW 9 million	KRW 0.7 million
			Less than KRW 13 million	KRW 0.7 million-(total income-900)*70/400
		Single-breadwinner household	Less than KRW 9 million	Total income*170/900
			Less than KRW 12 million	KRW 1.7 million
			Less than KRW 21 million	KRW 1.7 million-(total income-1,200)*170/900

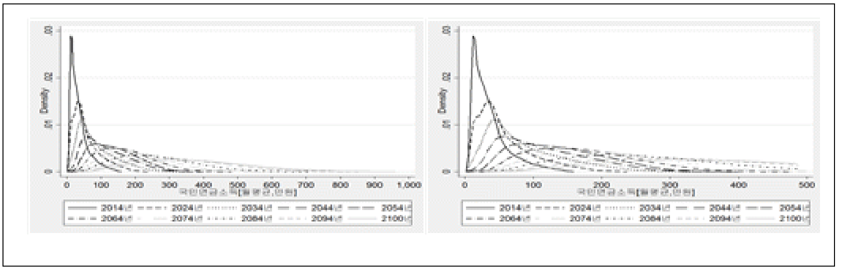
Type	Year	Household criteria	Income criteria	Formula
		Double-breadwinner household	Less than KRW 10 million	Total income*210/1000
			Less than KRW 13 million	KRW 2.1 million
			Less than KRW 25 million	KRW 2.1 million-(total income-1,300)*210/1200
CTC		Single-breadwinner household	Less than KRW 21 million	Number of children*KRW 0.5 million
			Less than KRW 40 million	Number of children*[KRW 0.5 million -(total income-KRW 2.1 million)*20/1900]
		Double-breadwinner household	Less than KRW 25 million	Number of children*KRW 0.5 million
			Less than KRW 40 million	Number of children*[KRW 0.5 million -(total income-KRW 2.5 million)*20/2500]

〈Appendix Figure 1〉 Wage and Business Income Distributions (Converted into Current Price Using CPI)



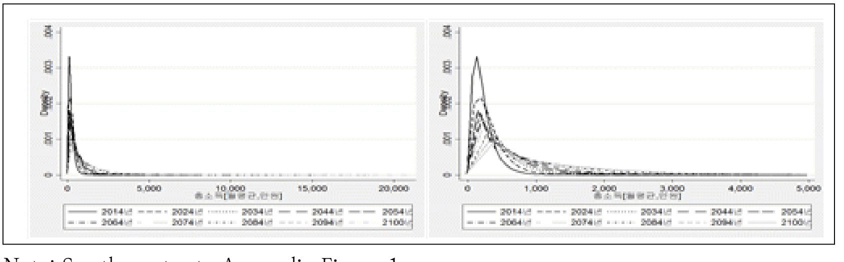
Notes: 1) Average monthly wage and business income, converted into the current price of 2012 using the inflation rate.
2) Persons whose wage and business income is zero are excluded.
3) Kernel density functions, smoothed using STATA's Epanechnikov.

〈Appendix Figure 2〉 Pension Income Distributions (Converted into Current Price Using CPI)



Note: See the notes to Appendix Figure 1.

〈Appendix Figure 3〉 Total Income Distributions (Converted into Current Price Using CPI)



Note: See the notes to Appendix Figure 1.