

EC-OECD Workshops on Business and Consumer Surveys

Recent Improvements of BOK Consumer Survey : Adjustment for Non-response

Jang, Youn Kyung



Bank of Korea
B O K O F K O R E A

Table of Contents

- I Outline of BOK Consumer Survey**
- II Sampling Design**
- III Adjustment for Non-Response**

I

Outline of BOK Consumer Survey

1. Brief Overview

2. Consumer Survey Indexes

3. Composite Consumer Sentiment Index

1. Brief overview

◆ Objective :

To obtain qualitative data for use in monitoring consumer sentiment on the domestic economic situation, living standards and spending plans, etc.

◆ Survey Periodicity : Monthly (2008.7~)

Quarterly (3Q. 1995~2Q.2008)

1. Brief overview

◆ Survey Target :

Head of Household of Every City Nationwide

◆ Sample Size : 2,200

◆ Survey Method : Mail and E-mail survey (Self-administrative Questionnaires)

1. Brief overview

◆ **Timeframe of Survey :**

a week centered around the 15th day of every month

◆ **Publication :**

around the end of a month ,

Press release and Statistics database,

Bank of Korea website

I

Outline of BOK Consumer Survey

1. Brief Overview

2. Consumer Survey Indexes

3. Composite Consumer Sentiment Index

2. Consumer Survey Indexes(CSIs)

Survey Items	
Economic Recognition	◆ Current Living Standards of Household
	◆ Current Domestic Economic Conditions
Economic Outlook	◆ Expectations of Living Standards of Household
	◆ Expectations of Domestic Economic Conditions
	◆ Expectations of Job Opportunity
	◆ Expectations of Level of Interest Rate

2. Consumer Survey Indexes(CSIs)

Survey Items	
Income and Expenditure	◆ Expectations of Household Income
	◆ Expectations of Household Spending
Savings and Debt	◆ Current Household Savings
	◆ Expectations of Household Savings
	◆ Current Household Debt
	◆ Expectations of Household Debt

2. Consumer Survey Indexes(CSIs)

Survey Items	
Expectation Of Price	◆ Expectations of Price Level
	◆ Expectations of Housing Price
	◆ Expectations of Wage Level

2. Consumer Survey Indexes(CSIs)

◆ Calculation Method

Each Consumer Survey index is calculated based on these five answer options

- ① Considerably Better (Considerably increase)
- ② Slightly Better (Slightly increase)
- ③ Similar
- ④ Slightly Worse (Slightly decrease)
- ⑤ Considerably Worse (Considerably decrease)

2. Consumer Survey Indexes(CSIs)

◆ Calculation Method

1. Assign a weight to each answer option

Answer	weight
① Considerably Better	+1.0
② Slightly Better	+0.5
③ Similar	0.0
④ Slightly Worse	- 0.5
⑤ Considerably Worse	- 1.0

2. On the basis of the distribution of each answer option for each question, aggregate balances are calculated for each question

2. Consumer Survey Indexes(CSIs)

◆ Index Formulation

$$B = (PP + 1/2P) - (1/2M + MM)$$

$$CSI = B + 100$$

where, PP, P, M,MM are percentage of respondents choosing each option

PP : Considerably Better or Considerably Increase

P : Slightly Better or Slightly Increase

M : Slightly Worse or Slightly Decrease

MM : Considerably Worse or Considerably Decrease

2. Consumer Survey Indexes(CSIs)

◆ Interpretation of CSI

Current Living Standards of Household CSI > 100

: The Number of households who think their current living standards got better is larger than the number of households who think their current living standards got worse

Current Living Standards of Household CSI < 100

: The Number of households who think their current living standards got worse is larger than the number of households who think their current living standards got better

I

Outline of BOK Consumer Survey

1. Brief Overview

2. Consumer Survey Indexes

3. Composite Consumer Sentiment Index

3. Composite Consumer Sentiment Index(CCSI)

◆ Comprehensive indicator of consumer sentiment

- Overall perceptions and expectations in a one dimensional index
- To be able to evaluate overall economic activity
- CCSI is comprised of six individual Consumer Survey Indexes that are standardized using their respective means and standard deviations over the standardization interval(2003 to 2012)

2. Composite Consumer Sentiment Indexes(CCSI)

◆ Component Indexes

Component Indexes
◆ Current Living Standards of Household
◆ Current Domestic Economic Conditions
◆ Expectations of Living Standards of Household
◆ Expectations of Domestic Economic Conditions
◆ Expectations of Household Income
◆ Expectations of Household Spending

2. Composite Consumer Sentiment Indexes(CCSI)

◆ Standardize Individual Component series

- Make the individual series comparable in terms of their mean level and volatility

< Formula for standardizing individual indexes >

□ Individual index standardized: $Y_{i,t} = \frac{X_{i,t} - \bar{X}_i}{S_i}$

· $X_{i,t}$: individual CSI of particular month

* i : individual index subscript (1 to 6); t : relevant month

· \bar{X}_i : mean of individual index over standardization interval = $\frac{1}{T} \sum_{t=1}^T X_{i,t}$

* $t(=1, \dots, T)$: standardization interval (2003 to 2012), extended by one year every January

· S_i : standard deviation of individual index over standardization period = $\sqrt{\frac{1}{T-1} \sum_{t=1}^T (X_{i,t} - \bar{X})^2}$

2. Composite Consumer Sentiment Indexes(CCSI)

◆ Compute Composite Consumer Sentiment Index

Formula for Composite
Consumer Sentiment Index

$$\square CCSI_t^* = \left(\frac{Z_t - \bar{Z}}{S_z} \right) \times 10 + 100$$

* CCSI: Composite Consumer Sentiment Index

• Z_t (summation of six individual standardized indexes) = $\sum_{i=1}^6 Y_{i,t}$

• \bar{Z} (mean of Z_t) = $\frac{1}{T} \sum_{t=1}^T Z_t$

• S_z (standard deviation of Z_t) = $\sqrt{\frac{1}{T-1} \sum_{t=1}^T (Z_t - \bar{Z})^2}$

• $sd (=10)$: standard deviations of CCSI

- Each component series has the same weight
- Scaled to have a long-term mean of 100 and s.d of 10

* s.d of 10 implies that about 68% are within the range of 90 to 110

2. Composite Consumer Sentiment Indexes(CCSI)

◆ Interpretation of CCSI

➤ Composite Consumer Sentiment Index > 100

Present Consumer Sentiment is more Optimistic than Long-term Average(Average of CCSI during 2003~2012)

➤ Composite Consumer Sentiment Index < 100

Present Consumer Sentiment is more Pessimistic than Long-term Average(Average of CCSI during 2003~2012)

II Sampling Design

1. Sample Design Procedure

2. Enumeration Districts Extraction

3. Sample Design Results

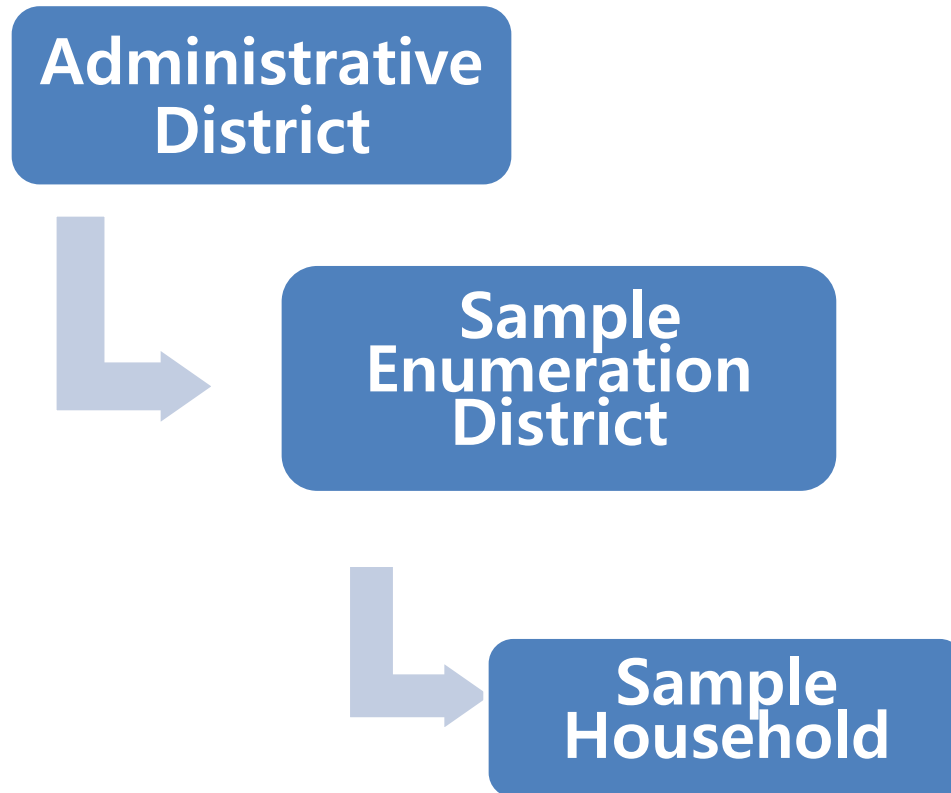
1. Sample Design Procedure

- ◆ Bank of Korea redesigned household sample in 2012 to reflect the changes in Population Structure.
- ◆ We extracted Sample households using the 2010 Population and Housing Census Roll conducted by The Korea National Statistical Office

1. Sample Design Procedure

- ◆ **Population** : National Urban Households
(91.2% of all households)
- ◆ **Sample size** : 2,200 Korea urban households
- ◆ **Sampling error** : less than $\pm 2\%$
- ◆ **Sampling methods** : Stratified two-stage Sampling

1. Sample Design Procedure



* Enumeration Districts : A group of about 60 households that are geographically close

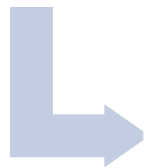
1. Sample Design Procedure

Administrative District

- **Stratified** to 19 administrative Districts



Sample Enumeration District



Sample Household

1. Sample Design Procedure

Administrative
District



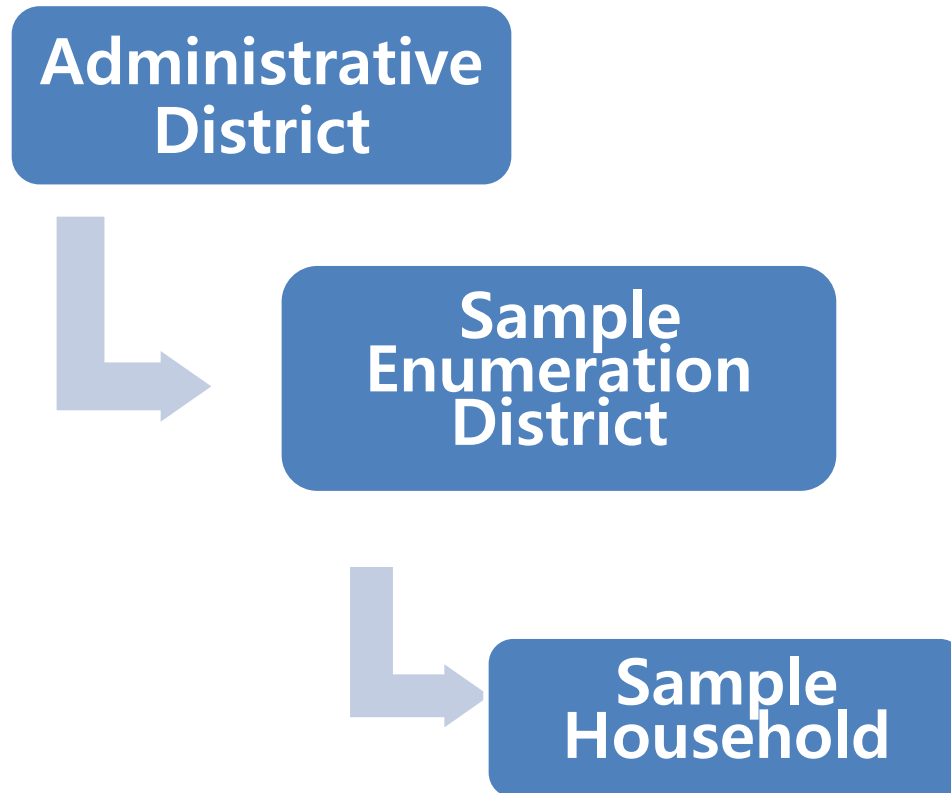
Sample
Enumeration
District

- **Proportional Allotment** by the number of households
- **Systematic Sampling** to reflect the characteristics of households



Sample
Household

1. Sample Design Procedure



- **Stratified by** the age of head of household
- **Systematic Sampling** to reflect the characteristics of household

II Sample Design Procedure

1. Sample Design Procedure

2. Enumeration Districts Extraction

3. Sample Design Results

2. Enumeration Districts Extraction

◆ The Extraction of Sample Enumeration Districts

The Type of Residence

```
graph TD; A[The Type of Residence] --> B[The Administrative District]; B --> C[The Age of the Head of the Household]; C --> D[The Average Total Floor Space];
```

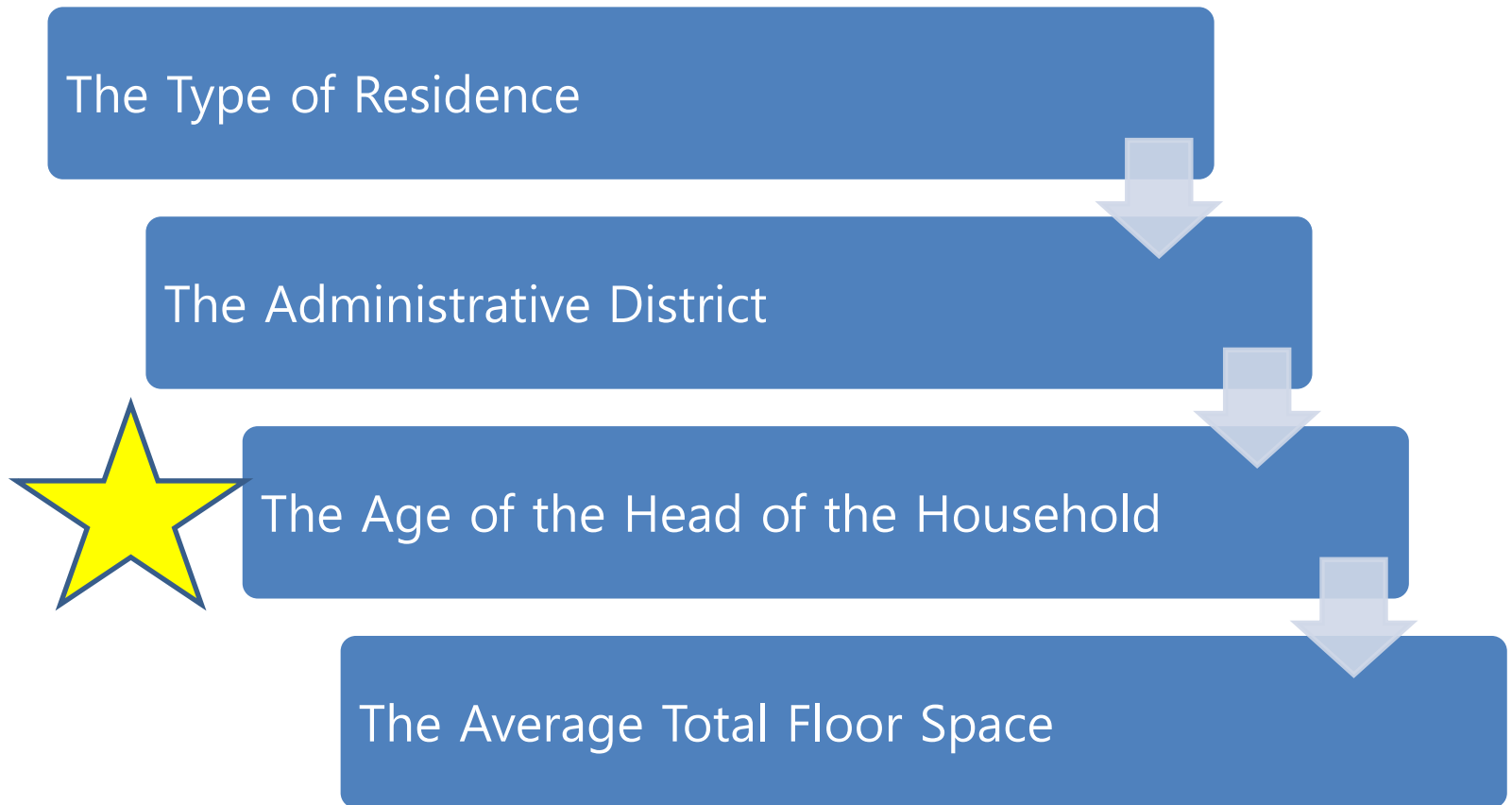
The Administrative District

The Age of the Head of the Household

The Average Total Floor Space

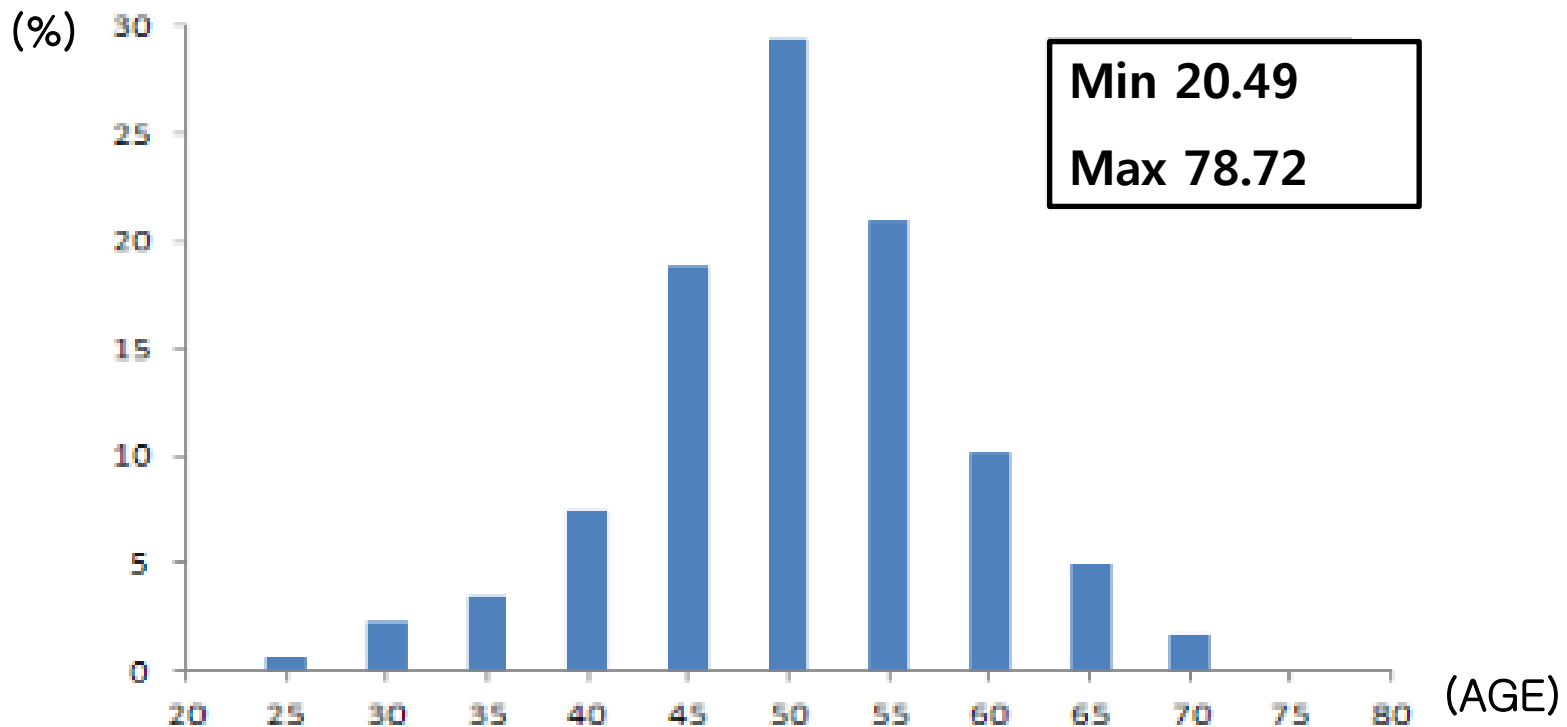
2. Enumeration Districts Extraction

◆ The Extraction of Sample Enumeration Districts



2. Enumeration Districts Extraction

- ◆ The average age distribution of the head of household of Enumeration Districts



2. Enumeration Districts Extraction

◆ The Extraction of Sample Enumeration Districts (Example : Seoul)

Enumeration Districts (population)	The average age of head of household in Enumeration Districts		Enumeration Districts (sample)	The average age of head of household in Enumeration Districts	The number of households in Enumeration Districts
1	20.49	⇒	1	20.49	63
2	21.22		2	22.15	55
3	22.15		3	23.06	51
⋮	⋮		⋮	⋮	⋮
⋮	⋮		⋮	⋮	⋮
55316	78.72		500	78.72	58
			SUM		28,508

2. Enumeration Districts Extraction

◆ The Extraction of Sample Enumeration Districts (Example : Seoul)

Age	The number of households in sample Enumeration Districts		Component ratio of Population
Under 30	1,837	(6.44%)	8.18%
30~39	5,498	(19.29%)	20.42%
40~49	6,793	(23.83%)	26.58%
50~59	6,826	(23.94%)	21.97%
Above 60	7,554	(26.50%)	22.86%
Sum	28,508		

II Sample Design Procedure

- 1. Sample Design Procedure**
- 2. Enumeration Districts Extraction**
- 3. Sample Design Results**

3. Sample Design Results

◆ Sample Distribution Table

		Previous		Present	
		Population	sample*	Population	sample
A G E	Under 40	33.9	10.0 (-23.9)	28.6	20.0 (-8.6)
	40~50	29.1	20.2 (-8.9)	26.6	24.1 (-2.5)
	50~60	18.9	29.5 (+10.6)	22.0	25.8 (+3.8)
	Above 60	18.1	40.3 (+22.2)	22.8	30.1 (+7.3)
R E G I O N	Seoul	26.0	25.1 (-1.1)	22.2	20.0 (-0.2)
	Metropolitan City	31.0	31.0 (same)	28.2	27.6 (-0.6)
	The Others	43.0	43.9 (+0.9)	49.6	50.4 (+0.8)

3. Sample Design Results

◆ Sample Distribution Table

		Previous		Present	
		Population	sample	Population	sample
INCOME (million won)	Under 1	-	8.6	16.9*	12.1
	1~2	-	20.3	18.5*	17.9
	2~3	-	22.5	17.6*	22.4
	3~4	-	18.0	14.2*	18.4
	4~5	-	13.5	10.5*	11.6
	Above 5	-	17.2	22.3*	17.5
Residence	Own	-	79.7	56.8*	70.0
	Rent	-	20.3	43.2*	30.0

* The results of Survey of Household Finances and Living Conditions, 2012
(survey 20,000 households)



Adjustment for Non-Response

1. Unit Non-Response

2. Item Non-Response

1. Unit Non-Response

- ◆ **Unit Non-Response :**

 - All survey item are unanswared**

- ◆ **Average response rate : About 90%**

1. Unit Non-Response

◆ Previous Consumer Survey

- **Compile Consumer Survey Indexes simply using the balance of respondents without adjusting sample weights**
- **Assume the response distribution of respondents the same as that of non-respondents**
- **Differences in Response Patterns between respondents and non-respondents affect the parameter estimate unfavorably**

1. Unit Non-Response

- ◆ **Difference of Response Pattern between respondents and non-respondents affect the parameter estimate unfavorably**
 - **Size of bias is a function of two factors**
 1. **proportion of population does not respond**
 2. **size difference in population between respondent and non-respondent**

1. Unit Non-Response

◆ Compensating for bias by non-response

- Make non-response as low as possible by using call-back method
- Weight-adjustment : to avoid bias by failing to include a particular portion of the population
(e.g. high income, younger age)

1. Unit Non-Response

◆ Post-stratification

- To compensate for unit non-response
- Adjust the weighted sample distribution for key variables to make it conform to a known population distribution

1. Unit Non-Response

◆ Bank of Korea Consumer Survey

- Survey data are weighted to two census variables
: region, age
 - Component Ratio of the Respondents
= Component Ratio of the Population
- Use an interactive raking ratio procedure
- Trim extreme weights

1. Unit Non-Response

◆ Raking Ratio Estimation method

- CALMAR(Calibration on Margins)

- ✓ SAS macro and calculate the calibrated weights when the auxiliary information in the survey consists of known marginal counts in a frequency table of arbitrary dimension
- ✓ The distance function used to measure the distance from the original weight to the final weight

- * Logit(L.U) method

- # L.U (lower, upper bound for weight ratio)

1. Unit Non-Response

◆ Before and After Weighting Adjustment

	Current Living Standards of Household	Expectations of Living Standards of Household	Current Domestic Economic Conditions	Expectations of Domestic Economic Conditions
Correlation coefficient	0.99	0.99	1.00	1.00
P value of PT test	9.9×10^{-9}	1.3×10^{-8}	6.6×10^{-8}	2.4×10^{-7}
	Expectations of Household Income	Expectations of Household Spending	Composite Consumer Sentiment Index	
Correlation coefficient	0.98	0.98	0.99	
P value of PT test	4.4×10^{-3}	7.6×10^{-2}	5.3×10^{-3}	

* Pesarnar & Timmermann test (1994) :

$$S_n^2 = \frac{(\hat{p} - \hat{p}_*)^2}{\text{var}(\hat{p}) - \text{var}(\hat{p}_*)} \sim \chi_1^2$$

1. Unit Non-Response

		Current Living Standards of Household	Expectations of Living Standards of Household	Current Domestic Economic Conditions	Expectations of Domestic Economic Conditions
Absolute difference	Mean	0.78	0.91	1.31	1.65
	Max	3	3	4	4
T-hac test statistic		0.34	1.58	0.27	-3.07
		Expectations of Household Income	Expectations of Household Spending	Composite Consumer Sentiment Index	
Absolute difference	Mean	2.24	2.4	1.81	
	Max	5	4	6	
T-hac test statistic		9.43	9.44	0.98	

* Newey and West(1987) T-hac test : $t-hac = \bar{R} / \sqrt{Var(\bar{R})}$

$$\widehat{Var}(\bar{R}) = \frac{1}{n(n-1)} \left\{ \sum_{t=1}^n \hat{\epsilon}_t^2 + 2 \sum_{\nu=1}^q \left(\left(1 - \frac{\nu}{q+1} \right) \sum_{t=\nu+1}^n \hat{\epsilon}_t \hat{\epsilon}_{t-\nu} \right) \right\} \quad \hat{\epsilon}_t = R_t - \bar{R}$$

1. Unit Non-Response

◆ Results

The results that some of the before and after weighting adjustment CSIs have significant level differences while having the same direction and high correlation means that the bias caused by non-response has been reduced.

2. Consumer Survey Indexes(CSIs)

◆ Variance Estimation

CV(Coefficient of variation) Using Jackknife method

$$CV(CSI) = \frac{\sqrt{\text{Var}(CSI)}}{CSI} = \frac{s.d.(CSI)}{CSI}$$

	May	June	July
Current Living Standards of Household	0.996	1.023	0.972
Expectations of Living Standards of Household	1.010	0.970	1.007
Expectations of Household Income	0.916	0.932	0.907
Expectations of Household Spending	1.024	1.027	1.024
Current Domestic Economic Conditions	1.267	1.104	1.075
Expectations of Domestic Economic Conditions	1.102	1.023	1.046

Weighting Adjustment

1. Unit Non-Response

2. Item Non-Response

2. Item Non-Response

◆ Item Non-response

- The respondent skips some survey items
- There was seldom Item non-response
 - ✓ All survey items of BOK Consumer Survey are simple Questions
 - ✓ Can obtain complete data by supplementary phone surveys

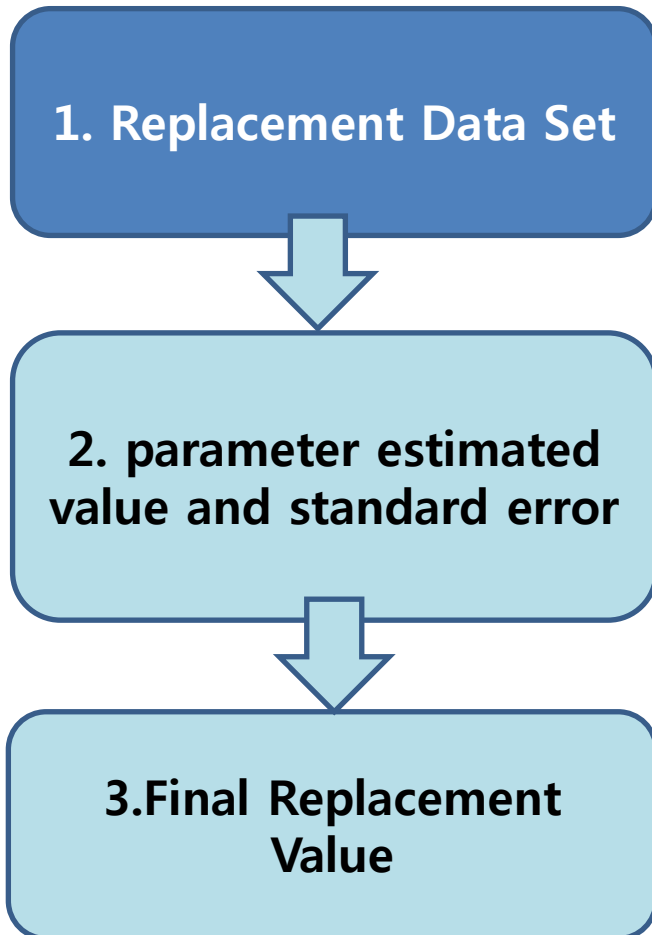
2. Item Non-Response

◆ Imputation method for Item non-response

- The missing value of personal information of respondent
 - Replaced with the information value of the previous month
- The missing value of CSI items
 - Replaced with an imputed value using multiple imputation method

2. Item Non-Response

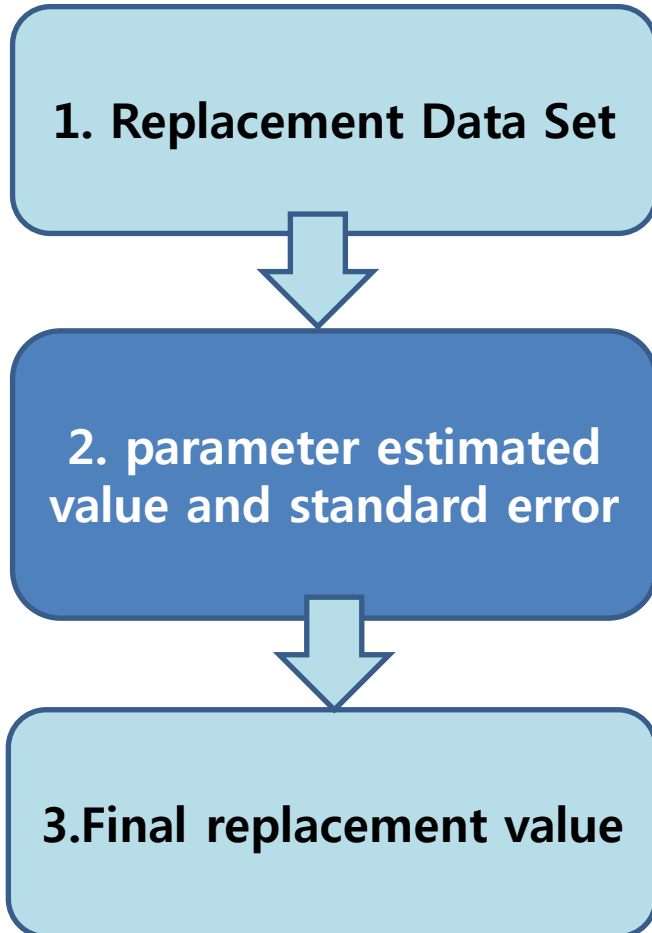
◆ Multiple Imputation Method(by SAS MI Procedure)



- Stepwise regression model
- From the survey Item which has the least missing values, second least missing values ...until all missing values are replaced
- Consider error variance for predictive value
- Increase explanatory variable until model estimation results don't show significant improvements
- Repeating all of step 1 to make m units of replacement Data set

2. Item Non-Response

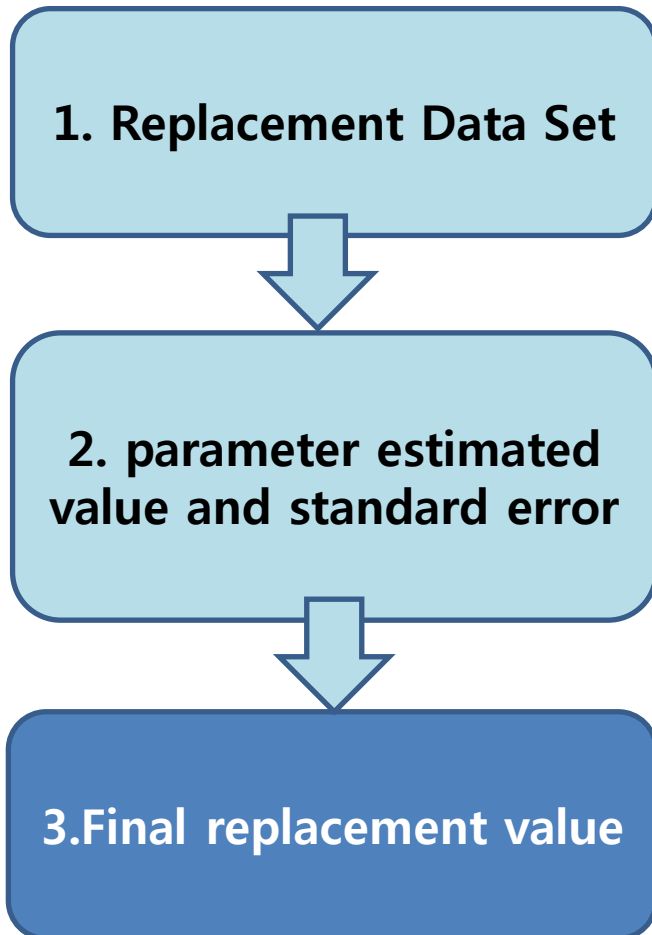
◆ Multiple Imputation Method(by SAS MI Procedure)



- compute parameter estimated value and standard error for each set

2. Item Non-Response

◆ Multiple Imputation Method(by SAS MI Procedure)



- calculate the final replacement value by composing parameter estimated value and standard error

2. Item Non-Response

◆ Simulation Results

		The number of household	Minimum value	Maximum value	mean	Standard deviation
set 1	All response	1889	1	5	3.034	0.712
	Item Non-response	99	0.918	4.913	3.009	0.817
	Whole(After Imputing)	1988	0.918	5	3.033	0.717
set 2	All response	1889	1	5	3.034	0.712
	Item Non-response	99	1.061	5.099	3.023	0.707
	Whole(After Imputing)	1988	1	5.099	3.034	0.711
set 3	All response	1889	1	5	3.034	0.712
	Item Non-response	99	1.172	4.576	3.056	0.643
	Whole(After Imputing)	1988	1	5	3.036	0.709
set 4	All response	1889	1	5	3.034	0.712
	Item Non-response	99	1.402	4.980	3.053	0.694
	Whole(After Imputing)	1988	1	5	3.035	0.711
set 5	All response	1889	1	5	3.034	0.712
	Item Non-response	99	1.524	4.716	3.038	0.651
	Whole(After Imputing)	1988	1	5	3.035	0.709

2. Item Non-Response

◆ Simulation Results

index	Imputing	Original	Difference
Composite Consumer Sentiment Index	104	104	0
Current Living Standards of Household	91	91	0
Expectations of Living Standards of Household	98	98	0
Expectations of Household Income	99	99	0
Expectations of Household Spending	106	106	0
Current Domestic Economic Conditions	79	79	0
Expectations of Domestic Economic Conditions	97	97	0
Expectations of Job Opportunity	94	94	0
Expectations of Level of Interest Rate	84	84	0
Expectations of Price Level	136	135	-1
Expectations of Housing Price	112	111	-1
Expectations of Wage Level	116	116	0

THANKS

Jang, Youn Kyung
(ykjang@bok.or.kr)



Bank of Korea
B S U K O F K O R E A