

# The Use of Alternative Data Source As a Proxy to Approach More Frequent Updates of CPI Expenditure Weight

Fathia Utami Afdi<sup>1</sup>, Fenanda Dwitha<sup>2</sup>, Zaradia Permatasari<sup>3</sup>

<sup>1</sup>Statistics Indonesia ([fathia.utami@bps.go.id](mailto:fathia.utami@bps.go.id))

<sup>2</sup>Statistics Indonesia ([fenanda.dwitha@bps.go.id](mailto:fenanda.dwitha@bps.go.id))

<sup>3</sup>Statistics Indonesia ([zaradia.permatasari@bps.go.id](mailto:zaradia.permatasari@bps.go.id))

**Abstract.** The Consumer Price Index (CPI) is one of the key economic indicators widely used as a guide to make decisions and determine government policy effectiveness. This index is the most well-known inflation indicator that shows the household's price level of goods and services. BPS-Statistics Indonesia updates CPI monthly, and the compilation uses the expenditure weight from the Household Budget Survey held every five years. Facing a dynamic and rapid change in consumption behavior, using a fixed expenditure weight over a long period can be no longer relevant in describing the enormous shift in household expenditure patterns. The updated expenditure weight data would continuously provide the data user needs. However, conducting more frequent HBS requires additional cost. Regarding the importance of CPI, the index's quality and measurement approach are necessarily improved and revised, including the methodologies and data sources. In Statistics Indonesia, one of the alternative data sources that portrays household expenditure is National Socio-Economic Survey (SUSENAS). The focus on this study is to exercise the approach in compile new CPI using weight expenditure derived from the alternative data sources, which are more frequently updated. The updated weights show slightly different of consumption pattern compared to the existing CPI weights. The CPI obtained from the updated weight are evaluated with distribution pattern, Relative Mean Square Error (RMSE), and Mean Percentage Absolute Error (MAPE). The alternative CPI is closely similar to some points of the existing CPI series, in which MAPE are 12.11% for weight expenditure and 0.30% for CPI. also, RMSE are 1.74 and 0.39 for the same indicators. It may conclude that using Susenas data as alternative data sources can be considered as a feasible method for generating more frequent of CPI expenditure weight, since particularly it can address towards latest condition in society and more relevant to updated consumption patterns in force majeure conditions specifically, such as The Covid-19 pandemic. However, this study is subject to several limitations. It can only update consumption expenditure despite the basket of commodities and challenges faced in mapping commodities captured from the alternative data source. BPS-Statistics Indonesia will continue to review the strategy to implement the alternative weight in compiling CPI and we will conduct further studies as necessary.

**Keywords:** *CPI, inflation, expenditure weight, annually re-weight*

## 1. Introduction

### Backgrounds

The calculation of the Consumer Price Index (CPI) is one of the main indicators used to measure changes in the price level of goods and services consumed by the public. The CPI is the most commonly used tools to measure inflation and deflation, which is important indicator of an economy's health<sup>1</sup>. An accurate and up-to-date CPI plays an important role in economic policy analysis, business decision-making, measuring purchasing power, and monitoring inflation in a country.

One of the crucial aspects in the calculation of CPI is the weighting method used. These weights refer to the relative weight given to each commodity included in the commodity package as the basis for CPI calculation. The weight attached to each good or service determines the impact that its price change will have on the overall index<sup>2</sup>.

In Indonesia, CPI calculation weights are obtained from the Household Budget Survey (HBS) which is conducted every 5 years. The five-year period between surveys causes limitations in representing dynamic changes in the structure of household expenditure. In a dynamic economy, consumer preferences and spending patterns also change over time.

In recent years, we know that many countries from all over have faced big challenges since the onset of the COVID-19 pandemic. The pandemic has caused significant shifts in consumption patterns. Lockdowns and social restrictions have affected accessibility to certain goods and services and changed consumption preferences. Some sectors of the economy experienced a drastic decline, while others increased. Just for example, the tourism and transportation sectors have been severely affected by the travel restrictions. In this context, the changes of consumption patterns have shown the need for more frequent weight updates to obtain accurate CPI data.

In this study, a new approach is taken using the National Socio-Economic Survey (SUSENAS) data as an alternative data source to obtain more accurate and more frequently updated CPI calculation weights. SUSENAS is a household survey conducted periodically by Central Bureau of Statistics (BPS)-Indonesia that includes more information on the socioeconomic characteristics of households as well as comprehensive information on household expenditures, including prices and quantities of goods consumed.

By utilizing Susenas data as an alternative data source, this study aims to develop an alternative method in calculating the CPI by updating the weights more frequently. This approach is expected to provide a more accurate picture of changes in consumption patterns that occur in households. By considering SUSENAS data available every year, the resulting CPI can reflect more representative price changes and provide more accurate information for economic decision-making.

This study also aims to analyze the effectiveness of using Susenas data for update the CPI weight in Indonesia. In this study, the process of selecting commodities, calculating weights, evaluating the quality of Susenas data, and comparing the results of CPI calculations using the weight from HBS and the Susenas-based method are carried out.

The use of Susenas data as an alternative data source in CPI calculation is expected to help overcome the challenges faced by HBS, such as the cost and time required in conducting more frequent surveys. In the context of a fast-changing economy, the ability to update CPI calculation weights more regularly can provide significant benefits in monitoring inflation, planning price policies, and better understanding public consumption trends. This also relates to the empirical support found by Australia Bureau of Statistics that higher frequency re-weighting of the CPI better captures consumer' substitution effects<sup>3</sup>.

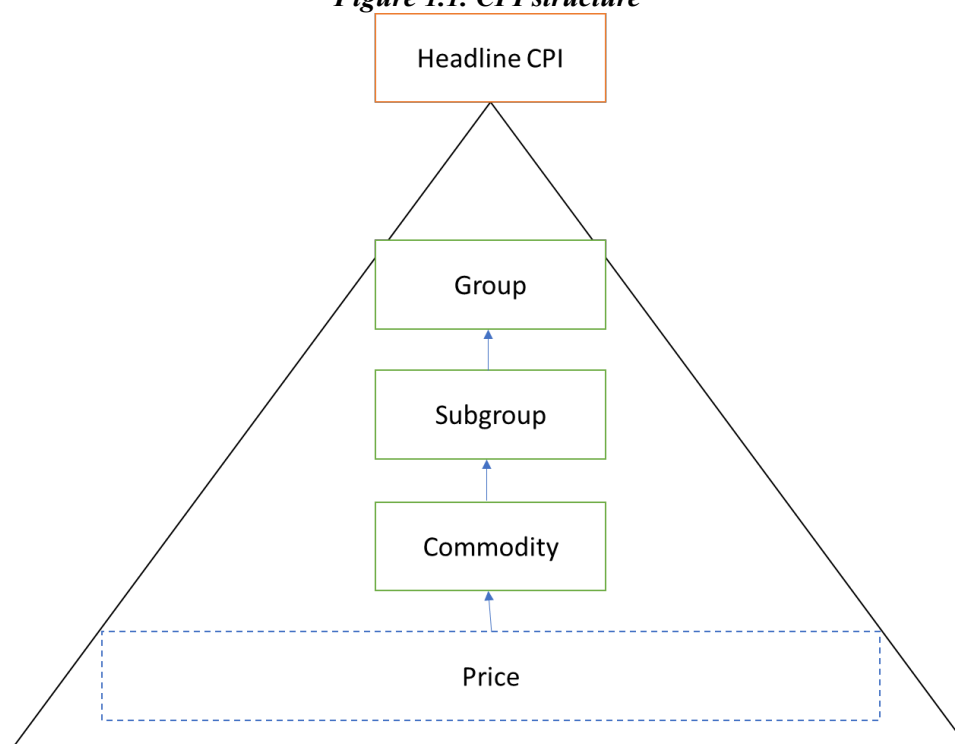
## Overview of Household Budget Survey and Consumer Price Survey in Indonesia

Consumer Price Survey (CPS), conducted by Statistics Indonesia, is a survey of transactional price between seller (retailer) and buyer (consumer). This survey aims to provide Consumer Price Index (CPI), one of strategic indicators collected by Statistics Indonesia for policy making. The difference in percentage of CPI, known as inflation or deflation rate, is essential economic indicators.

In history, CPS was held for the first time in 1953 in Jakarta. In 1968 CPI was calculated in eight selected capital cities. Starting from 2018 until now, CPS has been held in 90 cities in Indonesia. Statistics Indonesia releases CPI and inflation/deflation every month, disseminated in 90 cities and in national level.<sup>4</sup>

In calculating the headline CPI in a city, aggregation is done from a top-down approach, as seen in the figure below.

*Figure 1.1. CPI structure*



1. Price  
The data collected is transactional price data between sellers and buyers at the retail level. Prices are collected in traditional and modern markets. Microdata are not disseminated.
2. Commodity  
Commodity is the types of goods or services whose prices are monitored in consumer price survey. The number of commodities in the basket is different for each city, according to the consumption pattern of the society in each region. The lowest level index of the CPI is calculated at the commodity level. Currently, at the national level there are 835 commodities.
3. Subgroup  
All commodities are grouped to a higher level based on Classification of Individual Consumption by Purpose (COICOP), with minor modification for Indonesia. Based on the HBS 2018, there are a total of 43 subgroups.
4. Group  
Based on HBS 2018, there are 11 groups of household expenditure. Detailed descriptions of the 11 groups for national are as follows.

Table 1. The CPI Classification by Groups

Code	Group	The number of subgroups	The number of commodities
01	Food, Beverages, and Tobacco	4	348
02	Clothing and Footwear	2	101
03	Housing, Water, Electricity, and Household Fuel	4	39
04	Furnishings, Household Equipment, and Routine Household Maintenance	6	76
05	Health	4	27
06	Transport	4	39
07	Information, Communication, and Financial Services	4	27
08	Recreation, Sport, and Culture	6	37
09	Education	4	15
10	Food and Beverage Serving Services/Restaurant	1	68
11	Personal Care and Other Services	4	58
	Total	43	835

#### 5. Headline CPI

Headline CPI is the highest level of the index containing all groups, all subgroups, and all commodities, calculated at the city and national level. The change in headline CPI called inflation/deflation shows the rate of increase/decrease in the price of goods/services in general at the level of dissemination.

In Indonesia, the current CPI is disseminated for 90 cities and at the national level. The basket of goods/services and the weights are obtained from Household Budget Survey (HBS). HBS is a household consumption expenditure survey to obtain patterns of consumption as a material for preparing weights and basket commodities in the CPI calculation. HBS is held for a full year to capture an overview of consumption patterns throughout the year. The data collected includes expenditure on food and non-food consumption. HBS was first held in 1977/1978, and continued in 1988/1989, 1996, 2002, 2007, 2012, and the last is 2018. Currently, the CPI is calculated based on the HBS 2018, which is the 7th survey since it was first implemented. Household Budget Survey (HBS) held every five years. So, the fixed weights are used for five years until the next HBS.

#### Overview of National Socio-Economy Survey (SUSENAS)

Statistics Indonesia is responsible for the availability of data needed for sectoral and cross-sectoral development planning. One of the data sources needed especially for planning in the Socio-Economic Population sector is the National Socio-Economic Survey (Susenas) which is held by BPS every year. In general, the purpose of data collection through the Susenas is the availability of data on household welfare, including education, health, and purchasing power.

Susenas was first held in 1963. Susenas collects KOR data (main information) and module data (special information). Module data is collected once every 3 years covering the population consumption and income module, the social, cultural, and educational module, and the housing and health module. Since 2015, Susenas has captured the Household Consumption/Expenditure Module every semester, data will be collected every year in March and September. Data on semiannual enumeration results can only be presented in September for both the national and provincial levels, while for March the data can be presented up to the district/city level. Therefore, the data analyzed in this study is data from the Susenas in March, from household consumption/expenditure modul.<sup>5</sup>

The household consumption module captures food and non-food consumption expenditures. The grouping of goods and services captured in the consumption module Susenas is also guided by the COICOP, the same as was done for the HBS and CPS. Therefore, this data can be used as an alternative approach to annual CPI reweighting.

	Expenditure Percentage			
	2018	2019	2020	2021
Cereals and Tubers	6.45	6.08	5.97	6.14
Fish/shrimp/common squid/shells, Meat, Eggs, and Milk	8.76	8.80	8.80	9.20
Vegetables and Fruits	6.06	5.61	6.16	6.33
Legumes, Oil, and Coconut	2.21	2.10	2.11	2.31
Spices and Miscellaneous food items	1.87	1.79	1.83	2.05
Beverages Stuff and Alcoholic Beverage	1.53	1.44	1.50	1.54
Tobacco and betel (Cigarettes)	5.82	6.05	5.99	6.06
Prepared food and beverages	16.82	17.26	16.87	15.63
Housing and household facilities	25.29	25.49	25.19	26.33
Goods and services	12.39	12.40	12.42	12.17
Clothing, footwear, and headgear	2.92	3.03	2.95	2.51
Durable goods	5.14	5.04	4.96	4.60
Taxes and insurance	2.81	3.01	3.38	3.92
Parties and ceremonies	1.94	1.89	1.90	1.21
Total	100.00	100.00	100.00	100.00

According to Susenas data, consumption patterns have changed from year to year. For example, from 2018 to 2019, there has been a significant shift in consumption patterns from the grains and tubers group to the prepared food group. In 2018, grains and tubers were originally 6.45%, down to 6.08%, and processed food increased from 16.82% to 17.26%. Entering the COVID-19 pandemic, food consumption began to fall. In 2020 it fell to 16.87% and 2021 15.63%. People reduce buying prepared food to minimize contact with people and choose to cook at home. For four years, the consumption of protein in the community has continued to increase, as seen in the percentage of the fish, meat, eggs, and dairy group which has consistently increased, 8.76% in 2018 to 9.20% in 2021. From 2019 to 2020 and 2021, entering the COVID-19 pandemic, people tend to reduce spending on non-food items, meanwhile people still need to eat. This causes the percentage of food expenditure to increase, and non-food expenditure to decrease. For example, the percentage of consumption for the clothing, footwear and headgear group decreased from 3.03% in 2019 to 2.95% in 2020 and 2.51% in 2021. Another example is the consumption of durable goods which fell from 5.04 in 2019 to 4.96% in 2020 and 4.60% in 2021. In 2021, when the peak of the COVID-19 pandemic occurred, large-scale social restrictions were implemented causing spending on parties and ceremonies to drop significantly from 1.90% to 1.21%.

It should be a concern that consumption patterns have changed even in just one year, especially with the COVID-19 pandemic which has caused people's lifestyles to change. Susenas data available every year can be an alternative to capture information gaps of changes in consumption patterns between years, from one HBS to the next.

## 2. Methodology

### Scope

The data used for this study is micro data from Susenas–The Indonesian National Socio-economic survey, an annual household survey conducted by Statistics Indonesia every March and September. Statistics Indonesia collected social and demography data, specifically to measure poverty indicator, education, health, consumption, housing, and other socio-economic indicators through this survey. The susenas questionnaire consists of many modules, each of which is asked according to the needs of the indicators to be generated [3].

In this study, we use the Susenas dataset for 2018-2021. This is in line with the research objective which is to capture differences in consumption patterns in this period, including during the covid-19. Besides, we focused on consumption and expenditure's modul in Susenas' questionnaire to obtain consumption value in each commodity that available in Susenas. The number of commodities collected from this survey is 296 commodities, 539 fewer than collected commodities in HBS.

### Using National Socio-Economic Survey to Conduct Alternative Expenditure Weight

#### 1. Data Processing

First, we process raw data of Susenas, consists of information on household expenditure, including the price, amount, and quantity of goods and services consumed. The Susenas data used in this study is Susenas data for 2018-2021. In this stage, appropriate statistical data processing techniques are used to obtain relevant and representative information.

#### 2. Conducting the alternative weight based on Susenas data

In order to obtain new weights for CPI calculation from alternative data sources (Susenas), it is necessary to establish commodity baskets and updated weight based on Susenas data, as was done when processing data from the Household Budget Survey. In updating the new weights, there are several challenges, they are:

##### a. Mapping the Commodities

In categorizing goods and services, there are differences in classification between the commodity coverage of Susenas and SBH 2018. To overcome the mismatch between Susenas and SBH commodity codes, a commodity code mapping process was carried out, by connecting or linking commodities in the Susenas data with commodities in the SBH 2018. This was done to ensure conformity and consistency in the grouping of goods and services and the determination of their weights.

During the mapping process, several conditions occurred, (1) one commodity in Susenas matched to 1 commodity in SBH, (2) one commodity in Susenas was mapped to 2 or more commodities in SBH, (3) two or more commodities in Susenas were mapped to 1 commodity in SBH, (4) there were commodities covered by SBH that were not included in Susenas. This condition is affected by the differences in commodity details between Susenas and SBH, which SBH collects more detailed commodity data than Susenas.

##### b. Calculating the Consumption Value and Imputation Process

The formula used for calculating the expenditure value of the commodity baskets based on Susenas data is

$$NK_i' = \frac{NK_i^{(hbs)}}{\sum_{1,2,\dots,l} NK_i^{(hbs)}} \times \sum NK_j^{(ssn)}$$

*Notation :*

$NK^i$  : Updated expenditure value for  $i$ th commodity based on Susenas

$NK_i(hbs)$  : expenditure value for  $i$ th commodity based on HBS

$NK_j^{(ssn)}$  : Susenas expenditure value for commodity  $j$  which is mapped to  $i$ th commodity

$i$  : index for commodity in HBS

$j$  : index for commodity in Susenas

To complete the expenditure value of commodities that not covered in Susenas, imputations were made by utilizing available information, the household expenditure weights for those commodities based on SBH2018 results. Using imputation techniques, missing values were filled in with reasonable estimates, thus ensuring data completeness and consistency in the formation of weights.

c. Adjustments

In developing the new weights based on Susenas data, adjustments are also needed to obtain more accurate weights. Adjustments are made with the intention of considering changes in consumption patterns or consumer preferences over time.

**Differences in HBS and Susenas commodity details**

In the HBS, the list of commodities is larger and more detailed, while in the Susenas data, the coverage of commodities is more limited, and some commodities are combined as one unit without more specific details. The limited commodity detail in Susenas may cause expenditure on certain commodities to be aggregated into one broader category, resulting in lower expenditure values than actual. If expenditure on certain commodities is lower than it should be, then the weight on other higher commodities may be too large.

To overcome this problem, it is necessary to adjust the process of developing commodity weights. This adjustment aims to ensure that the resulting commodity weights reflect a more accurate proportion of expenditure and are not distorted by an underestimate of expenditure on certain commodities.

Adjustments were carried out such as:

- a. estimating the true value of expenditure based on other information in the Susenas data or using more detailed HBS data.
- b. Proportional approach: used to adjust commodity weights by considering differences in commodity classification between Susenas and HBS data. In this case, the proportion of expenditure on more detailed commodities in the HBS is allocated to broader categories in the Susenas data.

**Adjustments in the Insurance commodity, as well as the party and ceremonies category**

The Insurance Cost commodity covered in Susenas has a different concept from HBS2018. Insurance costs covered in HBS 2018 are only administrative costs, while insurance costs covered in Susenas include premi and other costs. The commodities in the party/event category covered in Susenas are also not covered in SBH as household expenditure. Therefore, for these commodities, adjustments were made by excluding the expenditure value from the new weight calculation.

3. Evaluation the weight based on Susenas Data

At this stage, the results of the weight conducted from Susenas Data are evaluated. The evaluation is carried out by comparing the expenditure weight from HBS-2018 with the expenditure weight from Susenas-2018. The purpose of this comparison is to ensure the suitability and accuracy of the weights generated from the Susenas data with the published weights from HBS 2018. In

addition, an analysis of the weight movement between Susenas periods is also conducted to identify changes in trends of household expenditure patterns year to year.

In statistics, there are several tests and procedures that can be done in a study, where one of them is determining the level of correlation between variables. Correlation itself is one of method that studies the degree of relationship between two or more variables. to determine the level of relationship between the two variables, we can evaluate from the size of the correlation value or what is commonly referred to as the correlation coefficient. The pearson correlation, is a correlation which coefficient measures the linear relationship between pairs of numerical codes for categories of each variable. In this study, we calculate the correlation of the elementary weights (in commodity level) between HBS 2018 and Susenas. The formula used as follow:

$$r_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

#### 4. Compiling the Alternative Price Indices

After evaluating the Susenas data to be an alternative data source in determining household expenditure weights, the weights are used to create the alternative price indices, which is then compared to the published CPI. The expenditure data from Susenas is available every year, enabling expenditure weights to be derived annually for the approach of CPI weights. Thus, the expenditure weights can be updated periodically and can captured the changes of household consumption patterns reflected in the Susenas data each year. This is yet another challenge to update the CPI weight regularly every year. In term of implementing the annually update of CPI weight, it needs to be considered whether the weights that have been formed based on Susenas data will be directly used as the weights for CPI calculation. Another choice is by using the Susenas weight movement to update the HBS expenditure in the commodity level. This practice is in line with the ABS (Australian Bureau of Statistics) which updates the CPI weights annually using the HFCE data movements. Meanwhile in this preliminary study, we started by comparing the published CPI with the Susenas price index, to see the feasibility of using the Susenas data as an approach to calculate the CPI. Further research is needed to determine the method that will be used to update the HBS weights continuously.

The alternative price index for all groups (headline CPI) is constructed using the weighted index of 835 commodities. The method used to construct the alternative price index at both the elementary level index and the upper level index is the same as that used in the published CPI.

##### a. Elementary Level Indices

Both series, the published CPI and the alternative price index, use the same published commodity price change in the city level. In practice, Indonesia use the Jevon formula to calculate the indices in the elementary level

$$I_{Jevons}^{0;t} = \prod_i \left( \frac{p_i^t}{p_i^0} \right)^{1/n}$$

$$I_{Jevons}^{0;t} = \frac{\prod_i (p_i^t)^{1/n}}{\prod_i (p_i^0)^{1/n}}$$

Where,

$p_i^t$  = price in current period

$p_i^0$  = price in base period

##### b. Upper Level Indices

As the current method used to compile the published CPI in Indonesia is Modified Laspeyres, the updated price index also uses the same method, which the weight reference period is the same with the price reference period. Since the Susenas expenditure data



collected on March, so the reference period for the updated price index is on March every year. The formula of the Modified Laspeyres is shown as below:

$$I_t = \frac{\sum_{i=1}^n p_i^t q_i^0}{\sum_{i=1}^n p_i^0 q_i^0}$$

The availability of expenditure weight for every year, means there are several new CPI series with the difference reference period. As the new weights are introduced, the new series is not comparable to the previous series. However, to fulfill the user's need for a CPI time series that covers a long period of time and provides historical content, we will need to link the series together. In line with CPI Manual (2020), if this linking process continues for multiple years, the linking factors for each year must be derived from the indices on the fixed index reference period or made cumulative by chaining the annual series through time. For example, in this study we make the long CPI series with the index reference period on 2018, so the chain linking method can be expressed as:

$$I^{19/21} = I^{18/19} \times I^{19/20} \times I^{20/21}$$

Where,

$I^{19/21}$  = continuous chain-linking factor for annual indexes from 2018-2019 to 2020-2021

## 5. Validation

For testing the method used, we compare result of IHK based on Susenas' expenditure weight with HBS'. MAPE (Mean Absolute Percentage Error), introduced by Lewis (1982) are performed as the base of prediction test [4]. In addition, MAPE can be used to evaluate the accuracy of Susenas' sampled data as more frequent alternative expenditure weight before we estimate IHK further in the following years. Similarly, RMSE (Root Mean Square Error) is alternative method measuring the error of prediction model. The formula of MAPE and RMSE are shown as follows,

$$MAPE = \frac{\sum \left| \frac{CPI_{sbh} - CPI_{ssn}}{CPI_{sbh}} \right|}{n} \times 100\% \quad \dots 1)$$

$$RMSE = \sqrt{\frac{(CPI_{sbh} - CPI_{ssn})^2}{n}} \quad \dots 2)$$

$IHK_{sbh}$  represents the monthly IHK based on HBS' expenditure weight, whereas  $IHK_{ssn}$  based on Susenas'. N is the total number of observations, which is 12 months in this study. As a result, we also can interpret the MAPE value according to Lewis' categorization (1982) as shown in Table X. Meanwhile, small RMSE value explains that predicted model is fitted to its true value, in this study represented by IHK based on HBS.

MAPE	Interpretation
<10%	Highly accurate forecasting
10-19%	Good Forecasting
20-49%	Reasonable Forecasting
>50%	Inaccurate forecasting

### 3. Result

#### HBS-2018 and Susenas

Obtaining the new expenditure value of the commodity packages based on Susenas-18 data, we conduct the comparative analysis between the resulted weight and the published weight. This comparison is carried out to see the suitability between the two data sources in forming the weight of household expenditure used in the CPI calculation. A comparison of group level expenditure weights for HBS-18 and Susenas-18 is shown in the table below.

Table 3.1. Expenditure Weight in Group Level for HBS-18 and Susenas-18

<b>Code</b>	<b>Group</b>	<b>HBS-18</b>	<b>Susenas-18</b>
01	Food, Beverages, and Tobacco	25,01	29,49
02	Clothing and Footwear	5,41	4,37
03	Housing, Water, Electricity, and Household Fuel	20,45	18,25
04	Furnishings, Household Equipment, and Routine Household Maintenance	5,97	5,49
05	Health	2,62	2,62
06	Transport	12,38	10,82
07	Information, Communication, and Financial Services	5,83	5,49
08	Recreation, Sport, and Culture	2,15	1,58
09	Education	5,62	5,52
10	Food and Beverage Serving Services/Restaurant	8,67	10,77
11	Personal Care and Other Services	5,89	5,60
	<b>Total</b>	<b>100</b>	<b>100</b>

From the table above, we can see that in the group level category, the HBS-18 and Susenas-18 show similar pattern of expenditure weight. Both based on HBS 2018 and Susenas 2018, household consumption is dominated by expenditure on the food, beverages, and tobacco group, which amounted to 25.01% based on HBS 2018 and 29.49% based on Susenas 2018. The next largest consumption is followed by the Housing, Water, Electricity, and Household Fuel group, which amounted to 20.45% and 18.25% based on HBS 2018 and Susenas 2018, respectively. The transportation group contributed the 3rd largest consumption value based on both data sources. And the group with the smallest percentage of consumption value based on both surveys is the recreation, sport and culture group. These results show that, although there are slight differences in the percentage of expenditure value between the two data sources, there are similarities in the composition and consumption patterns between groups. This comparison provides valuable information regarding the potential use of Susenas data as an alternative in forming expenditure weights.

In addition to descriptively comparing the weights between HBS 2018 and Susenas 2018, we also tried to see how closely the two weights match and relate. At the commodity level, the SBH-18 and Susenas-18 results have a correlation value of 0.9812, which means a very strong correlation. This indicates the methods used to adjust the Susenas data to obtain alternative weight for CPI give consistent result to those derived from household budget survey.

Furthermore, we will compare HBS-18 with the following year's Susenas. HBS-18 with Susenas-19, Susenas-20, and Susenas-21. The purpose of this analysis is to evaluate how far the weight generated from SBH 2018 can represent changes of household consumption patterns in the following years. The correlation value at the commodity level can be seen in the table below.

Figure 2. The scatter plot of HBS 2018 and Susenas 2018-2021

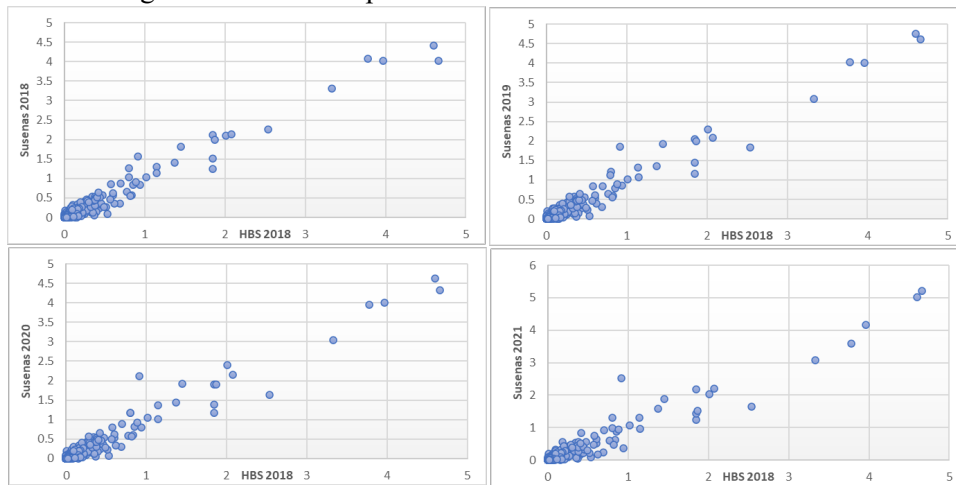


Table 3.2 Correlation Coefficient between HBS-18 and Susenas 2018-2021

Year of Susenas	Correlation with HBS-18 at Commodity Level
2018	0.9812
2019	0.9788
2020	0.9741
2021	0.9659

Based on the correlation test results, it can be seen that there is a strong correlation between the HBS-2018 weights and the updated weights of Susenas 2018, 2019, 2020, and 2021. However, the further the year difference between the HBS-2018 and Susenas, the lower the correlation. For example, between HBS 2018 and Susenas 2019, the correlation decreased from 0.9812 to 0.9788. Then when there is a 2-year gap, the correlation further decreases to 0.9741, and after 3 years the correlation again decreases to 0.9659. The larger the year gap between the HBS and Susenas, the less conformity in the weights produced by the two data sources. This also strengthens the hypothesis that using a fixed expenditure weight for years can be no longer relevant in describing the enormous shift in household expenditure patterns.

### Changes in Consumption Patterns Between Years Based on Susenas

Susenas which is available annually, can be an alternative data source to capture changes in household consumption over time. We can analyze and evaluate changes in expenditure patterns over a year by comparing the latest susenas with the previous period. The following table shows the percentage of expenditure by group, based on Susenas data from 2018 to 2021. It can be seen that there are changes in consumption patterns from year to year. Some expenditure groups have increased or decreased, which may be influenced by several factors, such as changes in consumer preferences, or events that have an economic impact, such as the COVID-19 pandemic.

Table 3.3. Result of Expenditure Weight with Alternative Data Sources

Susenas	2018	2019	2020	2021
Food, Beverages, and Tobacco	29.49	28.35	29.33	30.57
Clothing and Footwear	4.35	4.59	4.42	3.55
Housing, Water, Electricity, and Household Fuel	18.24	18.71	18.24	19.72
Furnishings, Household Equipment, and Routine Household Maintenance	5.47	5.45	5.15	5.61

Health	2.61	2.61	2.56	3.19
Transport	10.93	11.32	11.48	9.41
Information, Communication, and Financial Services	5.48	5.28	5.25	5.64
Recreation, Sport, and Culture	1.57	1.63	1.57	1.20
Education	5.50	5.40	5.56	5.57
Food and Beverage Serving Services/Restaurant	10.77	11.12	10.80	9.67
Personal Care and Other Services	5.58	5.55	5.63	5.88

The weight of food, beverage, and tobacco group decreased from 2018 to 2019, while in 2020 and 2021 it increased. One of the reasons for this significant increase in 2021 is the COVID-19 pandemic, where people tend to spend more time at home, and reduce nonfood consumption. Large-scale social restrictions, restaurant closures, and health concerns may have encouraged people to make and consume more food at home. This phenomenon is also in line with the percentage of consumption in the food and beverage service / restaurant group, which has decreased in 2021. For four years, 2018 to 2021, the commodities with the biggest decrease weight are rice with side dishes  $-0.468$  and meatballs ready to eat  $-0.121$ .

On the other hand, the clothing and footwear group shows a downward trend from 2019 to 2021. Factors such as changes in fashion trends, reduced social activities, and decreased purchasing power due to the pandemic may affect the interest and need for new clothing and footwear.

Furthermore, the health expenditure group showed a significant increase in 2021. This could be related to the COVID-19 pandemic and people's increasing need for health service, including medical expenses and the purchase of medicines and vitamins. The increasing awareness of the importance of health could also be a factor influencing changes in health spending. Weight increasing for some commodities in health group can be seen in table 3.4.

Table 3.4. Result of weight increasing in Health Group in 2018-2021

Commodities in Health Group	Increase of Weight 2021 from 2018
Check Up	0.292
Vitamin	0.118
Laboratory	0.118
Hospital	0.050
Medicines By Prescription	0.036
Jamu (Herb)	0.005
Cough Medicine	0.005
Cold Medicine	0.003

The transportation group showed a drastic decline in 2021 due to reduced mobility and the prohibition of domestic and international flights. The impact of the COVID-19 pandemic, where travel restrictions, working from home, and decreased social activities led to a reduced need for daily transportation. The gasoline consumption decreased by  $-0.493$ . Air transport fell by around  $-0.473$ . Consumption of online taxibike decreased by  $-0.123$ , city transportation decreased by  $-0.122$ , and inter-city transportation decreased by  $-0.069$  (more details in the table below).

Table 3.5. Result of weight decreasing in Transportation Group in 2018-2021

Commodities in Transportation Group	Decrease of Weight 2021 from 2018
Gasoline	$-0.493$

Air Transportation Rates	-0.473
Online Taxibike	-0.123
City Transportation	-0.122
Inter-City Transportation	-0.069
Car	-0.068
Train	-0.058
Motorcycle	-0.052
Online Taxi	-0.047
Rental Vehicles	-0.034
Taxi	-0.014
Travel	-0.012
Sea Transportation	-0.009

The development of technology and expansion of the digital economy can also be shown by the increasing percentage of spending on the information, communication, and financial services group. This is especially shown in the increase in spending in 2021. In addition, the Covid-19 pandemic has also had an impact on increasing community activities that require internet services, such as online schooling and working from home, shown in table below.

Table 3.6. Result of weight increasing in Information, Communication, and Financial Services Group in 2018-2021

<b>Commodities in Information, Communication, and Financial Services Group</b>	<b>Increase of Weight 2021 from 2018</b>
Internet Subscription Fees	0.954
Laptop/Notebook	0.009

The impact of the Covid-19 pandemic is also reflected in the expenditure of recreation, sports, and culture group, which showed a decline in 2021. Large-scale social restrictions led to the closure of recreational venues, and a shift in household preferences in spending leisure time. The consumption in recreation decreased  $-0.198$  and bioskop was decrease around  $-0.036$  (2021 from 2018).

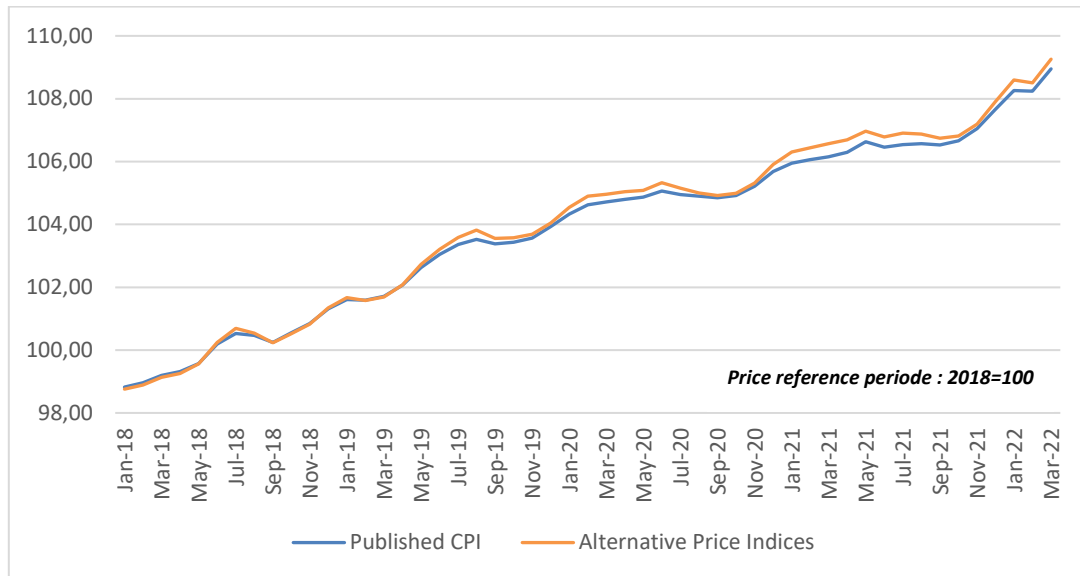
It can be concluded that changes in consumption patterns reflected in Susenas data show trends related to phenomena that occur in the year concerned, such as the Covid-19 pandemic which began to have an economic effect in 2020 and 2021. Such shifts may indicate significant economic changes in the pattern of consumer purchases. In addition, changes in consumer needs and preferences, technological developments, price fluctuations and socioeconomic developments can also affect changes in the composition and pattern of expenditure.

#### **Alternative CPI using Annual Updated CPI Weight**

Based on the evaluation of annual updated weight conducted by comparing the weights of Susenas 2018 and SBH 2018, it can be said that there are similarities in consumption patterns from both sources. In addition, changes in consumption patterns captured based on Susenas 2018-2021 data can also describe the phenomena that occur. This means that Susenas data can be used as an alternative to describe household consumption patterns and as an approach to obtain annual updates of CPI weights.

A comparison of the headline CPI from both sources can be seen in the graph below, which the alternative price indexes conducted appear similar and close to the published CPI.

Figure 3. Index Comparison between Susenas and HBS.



From the graph above, we can see that there is no significant difference between the published CPI (based on the SBH 2018 weight) and the Susenas prices indexes throughout 2018. It means that the expenditure weight from Susenas-18 can predict the published CPI throughout 2018 appropriately. The similar conditions of household expenditure in 2018 depicted in the two data sources provide similar results in the CPI. The similarity of the 2018 CPI data generated from both sources indicates that using the Susenas weights directly can be considered as a method to update the HBS weights periodically. However, further research should be conducted to corroborate these hypotheses.

From the graph above, in some points, the two series indexes began to show slight differences, especially in 2020 and 2021, where the covid pandemic greatly affected changes in people's consumption patterns. Overall, the publish CPI and the alternative index grew by 10,25% and 10,63% respectively from Januari 2018 to March 2022.

We already know that during a pandemic, the weight of several commodities changed significantly, such as increased consumption of the health and communication information group and decreased consumption of the transportation, recreation, and restaurant group. Just for example, the increased demand of health goods and services, causing the prices of several goods in health group to increase. The increase in the price of these commodity should have a significant impact during the pandemic and result in a higher headline CPI, due to the higher portion of expenditure for the commodities.

### Validation

We check validity Susenas dataset as alternative data source whether appropriate to use them to produce new expenditure weight. According to our result, as predictor consumption pattern, represented by weight expenditure in the same year in 2018, RMSE has small or nearly zero value. In accordance with this result, MAPE also show relatively small value at 12.11% which means that consumption pattern of Susenas' sampled data in 2018 can be used as predictor for consumption pattern of HBS in the same year.

The table below also depicts the feasibility of alternative price indices through alternative data sources (Susenas). The RMSE of alternative price indices is approximately zero (0.43), that means no difference between alternative price indices as resulted of Susenas and published CPI. It is followed

also with the value of MAPE of 0.34% so that we can conclude this predicted PI is highly accurate forecasting for the CPI in 2018.

Validity	Weight Expenditure	CPI
RMSE	1.74	0.39
MAPE	12.11%	0.30%

In summary, using an alternative dataset is able to describe consumption patterns and predict the price index as accurately as using SBH data. Based on these results, we have a stronger belief that to describe consumption patterns and the consumer price index, we can also use Susenas data in subsequent years.

#### 4. Conclusion and Implementation Plan

##### Conclusion

Indonesia currently updated expenditure weights through Household Budget Survey every five years at the commodity level. Using alternative data sources can be considered as a feasible method for generating more frequent of CPI expenditure weight, since particularly it can address towards latest condition in society and more relevant to updated consumption patterns in force majeure conditions specifically, such as The Covid-19 pandemic.

In this study, the alternative data source used is Susenas, which is available every year. The alternative weight and CPI conducted from the Susenas data show a price index that is more dynamic and responsive to changes in household consumption patterns. We also conducted empirical testing through validity tests, with MAPE and RMSE indicators, where Susenas data can be considered as an alternative data source for generating CPI expenditure weight annually to describe consumption patterns and conduct the consumer price index.

This study is expected to be a preliminary to the development of a more up-to-date and accurate CPI calculation methodology, as well as strengthening the understanding of the use of alternative data sources in compiling the CPI. However, since Household Budget Survey (HBS) can capture the whole picture of household expenditure and provide more detailed data, it is considered as the most established survey to obtain CPI weight. More frequent of HBS is still expected as the main source to compile the CPI.

##### Implementation Plan

Through this study, we can see the feasibility of Susenas data that can be used as an alternative source to describe household expenditure patterns. Further studies can be conducted in building a superlative index to estimate the substitution bias of the CPI. We also plan for future research to approach National Account data, namely HFCE as another data source that also captures household consumption expenditure, as one of the surveys recommended in the CPI manual.

Since BPS-Indonesia currently used the modified Laspeyres method for compiling the current Indices, we also plan to exercise the annual updated CPI using the alternative method, Young or Lowe. This study is the preliminary of BPS-Indonesia's plan to conduct more frequent updates of the CPI weight, so we will continue to review the strategy to implement the Susenas weight or other alternative data sources in compiling the CPI.

## 5. Limitation

Susenas is one of the most massive surveys in Indonesia, yet conducted only twice in a year, in March and September. It captures the consumption and expenditure of households in only those months, whereas HBS is carried out five-yearly and able to capture monthly consumption patterns over a year in the survey period with more detailed commodities.

On the other hand, there are differences in the classification and coverage of commodities in the Susenas data, where the coverage of commodities is more limited and not as detailed as the commodities in the HBS. This causes less coverage of expenditure on some commodities, which also has implications for the expenditure weights. BPS-Indonesia will continue to review the adjustment methods used to solve these issues.

## 6. Reference

- [1] Masterson, V. 2022. What is a consumer price index and why is it important?, World Economic Forum. Available at: <https://www.weforum.org/agenda/2022/05/what-is-the-consumer-price-index/>.
- [2] Graf, B. 2020. Consumer Price Index Manual 2020 Expenditure Weights and Their Sources. Washington DC : International Monetary Fund
- [3] ABS (Australia Bureau of Statistics). 2017. An Implementation Plan to Annually re-weight the Australian CPI. Australia.
- [4] Lewis, C.D. 1982. Industrial and business forecasting methods. London: Butterworths.
- [5] Statistics Indonesia.
- [6] ABS (Australia Bureau of Statistics). 2016. Information Paper: Increasing the Frequency of CPI Expenditure Class Weight Updates,. Australia
- [7] Household Budget Survey Metadata 2018 and Consumer Price Survey Metadata 2020 and 2021.
- [8] National Socio-Economy Survey Metadata. 2020.