

LWS Missing Values Policy

General Policy

LWS treats missing information differently depending on whether the variable in question is continuous (e.g., income or wealth amounts) or categorical (e.g., race or gender).

Continuous variables:

- If a variable is not present in the dataset, it is coded as 0 for all observations.
- If a variable is present, households to whom the question does not apply (not in sample) are coded as 0.
- If the variable is available and is applicable to the household, but the information was not supplied, the observation is coded as missing.

Categorical variables:

- If a variable is not present in the dataset, it is coded as -8 for all observations.
- If a variable is present, individuals to whom the question does not apply (not in sample) are coded as -8.
- If the variable is available and is applicable to the household, but the information was not supplied, every attempt is made to provide the reason for the non-answer:
 - 8 = "not applicable"
 - 7 = "refusal"
 - 6 = "don't know"
 - 10 = "proxy"

If the reason is unknown, the observation is coded as missing.

This policy was created in order to preserve the original information provided in the surveys. Therefore, if the detail is not provided, the original non-response remains as missing (.) in the final LWS data set.

Presence of missing values in LWS

As the new data sets are added and existing data sets are updated, every attempt will be made to abide to a consistent policy for handling missing values (as stated above). In the meantime, some inconsistencies may arise due to the coding in the original data. The following list provides details of the types of missing values utilized in each country.

Austria (AT04)

. = missing/not in sample/not stated value

Canada (CA99)

. = missing/not in sample/not stated value

Cyprus (CY02)

. = missing/not in sample/not stated value

-8 = "not applicable"

-7 = "refusal"

-6 = "don't know"

Finland (FI98, FI94)

no missing values

. = not in sample

Germany (DE01, DE06)

. = missing/not in sample/not stated value

-8 = "not applicable"

Italy (IT02, IT04)

. = missing/not in sample/not stated value

Norway (NO02)

no missing values

. = not in sample

Sweden (SE02)

. = missing/not in sample/not stated value

United Kingdom (UK00)

. = missing/not in sample/not stated value

-8 = "not applicable"

-7 = "refusal"

-6 = "don't know"

-10 = "proxy"

United States PSID (USP00)

. = missing/not in sample/not stated value

-8 = "not applicable"

-7 = "refusal"

-6 = "don't know"

United States SCF (USS00, USS03, USS06)

. = missing/not in sample/not stated value

-8 = "not applicable"