Assessing & Improving the Quality of Prescription Drug Data from Surveys: A Discussion

Geoffrey Paulin, Ph.D.
Senior Economist
Division of Consumer Expenditure Surveys

Federal Committee on Statistical Methodology
November 4, 2021
Virtual Conference

This session consisted of three papers:

- Khali Defever, Becky Reimer, Michael Trierweiler, and Elise Comperchio (NORC at the University of Chicago), “Improving self-reported prescription medicine data quality with a commercial database lookup tool and claims matching;”
- Becky Reimer, Elise Comperchio, Andrea Mayfield, and Jennifer Titus (NORC at the University of Chicago), “Exploring Potential Benefits of Enumerating All Prescribed Medicines as a Tool for Estimating Opioid Use in the Medicare Current Beneficiary Survey (MCBS);”
- Yao Ding and Steven C. Hill (Agency for Healthcare Research and Quality), “Evaluating Alternative Benchmarks to Improve Identification of Outlier Drug Prices for MEPS Prescribed Medicines Data Editing.”

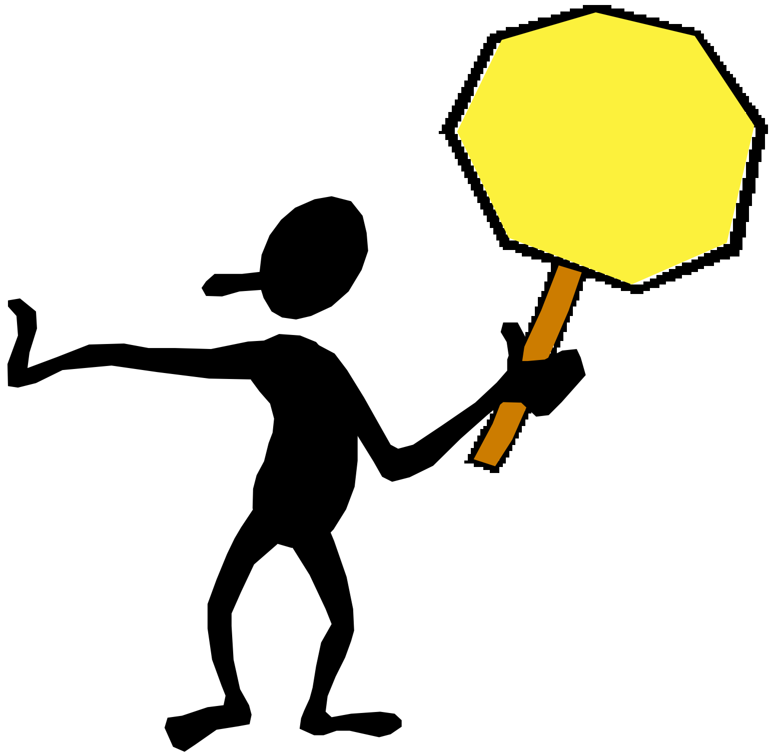
Commonalities

- Each collects data related to prescription drug purchases directly (types of medicine reported) or indirectly (prices of drugs).
- The first two use data from the Medicare Current Beneficiary Survey, while the third uses data from the Medical Expenditure Panel Survey.
- Each supplements survey data with administrative data, and finds that the final data set is improved over the survey data alone (e.g., higher levels of reported drug purchases in final set than from survey alone).

Differences:

- The first two papers use administrative data to adjust for nonresponse.
- The third paper uses administrative data to identify outliers, and improve editing procedures.

Caveat Emptor:



As discussant, it is important for me to mention the following "up front:"

- About the papers:
  - All focus on survey data that are subject to nonresponse.
  - All use outside sources—"administrative data"—to fill in the blanks.
- About the discussant:
  - I have no expertise in use of administrative data.
  - My expertise lies in conventional imputation methods.
  - Therefore, I shall compare and contrast these methods to the best of my knowledge, and provide some thoughts for the authors or others interested in mitigating nonresponse bias in their own survey-based analyses.

Why is “Missingness” Important?

- Missing data can lead to biases in:
  - Descriptive statistics
    - Mean
    - Median
    - Percent reporting
  - More sophisticated analyses
    - Regression coefficients
    - Standard errors of coefficients
- Measurement error within reported data can exacerbate these problems.

Defining terms:

For the purposes of this discussion,

- "Internal" imputation refers to methods for filling in the blanks using data reported within the survey of interest.
- "External" imputation refers to filling in blanks using administrative data, or other data obtained outside the survey of interest.

Common types of "internal" imputation:

- Hot-deck imputation
  - Select "n" characteristics that influence the item to be imputed. For example, expenditures (imputed when missing) are influenced by income and family size (so n=2).
  - Sort the data set by the "n" characteristics.
  - Find the record with values closest to the record with missing data, and fill in the value from the first "donor." (If the family on the missing record has $50,000 in income and two members, look for the first record with the same income and family size with the expenditure reported, and substitute the report for the missing value.)

Types of “internal” imputation (continued):

- Cold-deck imputation
  - Similar to hot-deck imputation, except that the “donor” comes from outside the current sample, such as from a previous edition of the survey.
  - Note the similarity between “cold-deck imputation” and “external” imputation. Both use data from outside the current sample. But there are differences. The administrative data are generally from a completely unrelated source. More on implications of this later.

Types of “internal” imputation (final example):

- Multiple imputation
  - Select a mechanism for conducting imputation. For example, regression-based prediction of missing values.
  - Modify the predicted value so that not one, but several estimates of the “true” value appear in the data set. (Example: Add random noise to the coefficients of the regression model, and predict the missing value based on the shocked model. Repeat several times. Each predicted outcome from the shocked models is one of the multiple imputations.)

Each of these "internal" imputation methods is used in major surveys today:

- Hot-deck imputation:
  - Current Population Survey (www.census.gov/programs-surveys/cps/technical-documentation/methodology/imputation-of-unreported-data-items.html)
- Cold-deck imputation:
  - Survey of Income Program Participation (www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2014-SIPP-Panel-Users-Guide.pdf)
- Multiple imputation:
  - Consumer Expenditure Surveys (www.bls.gov/cex/csxguide.pdf);
  - Survey of Consumer Finances (https://www.federalreserve.gov/econres/scfindex.htm);
  - National Health Interview Survey (www.cdc.gov/nchs/data/nhis/tecdoc16.pdf)

A Key Term for Any Imputation: Response Mechanism

- Missing Completely at Random (MCAR)
- Missing at Random (MAR)
- Nonignorable Nonresponse (NINR)

About Each Term: MCAR

- There is no pattern to the missingness. The probability of nonresponse is entirely random.
- This is almost never observed in “real” data.

About Each Term: MAR

- Missingness may be correlated with characteristics, but not with the value being imputed.
- Example: Reporting of income may vary by age or education of the respondent, but high income respondents and low income respondents with the same age and education have the same probability of reporting.

About Each Term: NINR

- Missingness may or may not be correlated with characteristics, but is correlated with the value being imputed.
- Example: High income respondents have a higher/lower probability of nonresponse than low income respondents, even with the same other characteristics (e.g., age and education).

“External” Imputation

- Most closely resembles cold-deck imputation, at least in the simplest form of external imputation.
- Key difference: External imputation may not match by individual records, but by descriptive statistics.
  - Suppose expenditures on rice are missing from the survey data, but an administrative source shows mean per capita expenditures on rice.
  - Multiply mean per capita expenditure by number of persons in the family to “externally impute” the expenditure on rice for that family.

Putting it all together…

The nature of the response mechanism is easiest to understand in the first two papers.

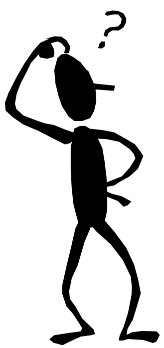
- It may be MAR in both cases.
- The evidence is particularly easy to understand from Defever, et al. (See slide 19.) The authors find "Factors associated with significantly more medicines discovered in claims matching" include:
  - Race and ethnicity
  - Number of chronic conditions reported (4 or more)
  - Income below 200% of the poverty line

However:

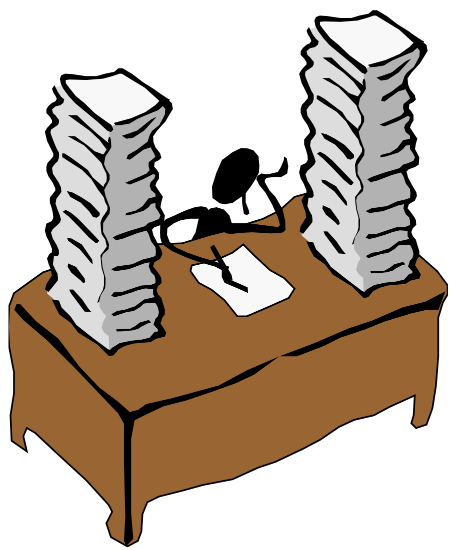
- Could it may be NINR?
- The first item on the same list (Defever et al., slide 19) of “Factors associated with significantly more medicines discovered in claims matching” is “Having a higher count of reported medicines.”

Whatever the mechanism:

- What role does use of administrative data play?
- Does it produce superior outcomes to using “internal” imputation?



I pose these two questions merely for the reader's consideration, not for the presenters' answers…



…Sorting through the implications could be considerable work.

Regarding the third paper:

- The nature of the problem is a bit different here. Rather than imputing for missing values, the authors use administrative data to identify outliers, and replace "incorrect" outliers (e.g., due to typographical error, rather than actual legitimate values) with more reasonable results.
- Nevertheless, are outliers more or less prevalent with some types of medications than others? If so, what are the characteristics?
  - Example: The authors correctly consider generic and "name brand" medicines.
  - Presumably, generics are lower in price.
  - Are they also more/less likely to have outliers? If so, what are the implications?

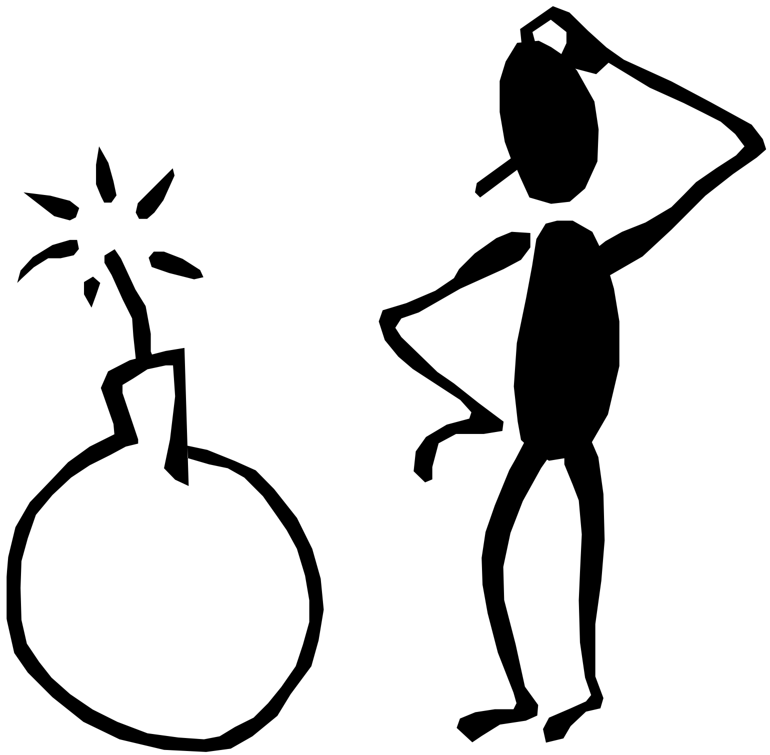
Reason This Matters:

Suppose the reader wants to implement imputation in a particular survey. What should the reader consider?

All imputation methods have strengths and limitations:

- Strengths of the methods discussed herein:
  - Generally, they preserve means of the imputed data
  - Allow use of a complete data set, reducing bias in certain estimates. (E.g., under MAR assumptions, if probability of reporting differs by age, as does level reporting, using reported data will result in a biased mean. Imputation, properly done, will correct this.)
- Limitations vary by method. Example: Hot-deck imputation can bias variance estimates downward.

What about “external” imputation specifically?



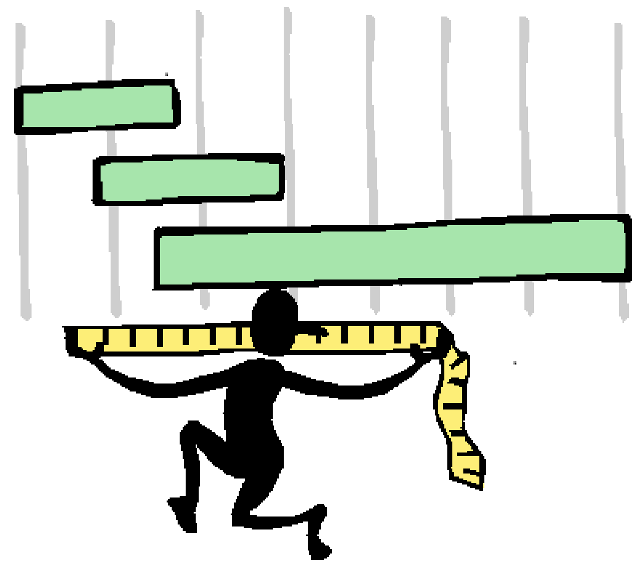
Factors to consider include:

Conceptual agreement.

- If the survey under study asks, "Did you purchase Prescription Drug X in the last month?" and the administrative data measure *all* prescription drugs purchased in a *calendar year*, the ability to provide good filling in of blanks is reduced.

Measurement Error

- This can affect both the survey of interest, and the external data.
- But external data may have an advantage: Reimer et al. (slide 24) mention “Recall bias due to self-report of opioid prescriptions.” Recall bias may be present for survey, but not administrative, data.

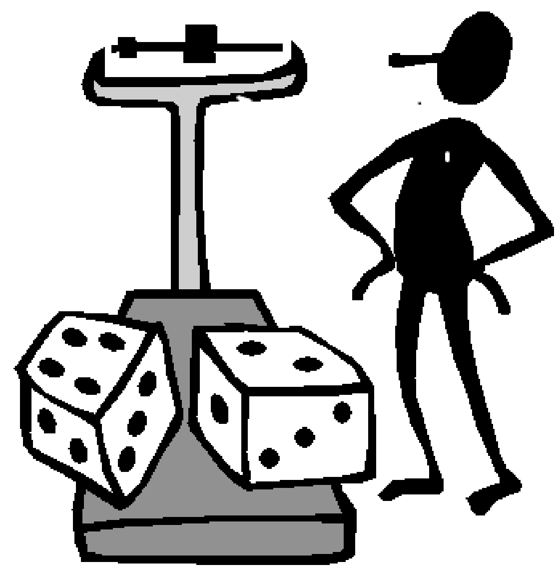


Use of Weights

- Whether to weight the data or not has implications both in imputation and use of complete data sets. (Note: Defever et al. use weights; Reimer et al. do not.)
- Particularly when conducting external imputation, how to weight is tricky: What if the survey is nationally representative, but the external data are not? (I.e., sample design differences.)
- Whether and how to use weights is important for the data provider to consider. A full discussion is beyond the scope of this presentation, but it merits mention.

Variances

- In addition to preservation of mean, preservation of variance is important.
- Hot-deck imputation biases variance estimates downward; multiple imputation requires special formulas to accurately estimate variance.
- Producers and users of imputed data need to be aware of these factors, and to weigh tradeoffs carefully.



Summary and Conclusions

- This discussion covers three papers that use administrative data to improve quality of survey data due to nonresponse or other factors (e.g., outliers).
- The author compares “internal” imputation (using reported data within the survey of interest) to “external” imputation (using administrative data), describing aspects of each, and considerations for selecting a method.
- Choosing a proper technique is not easy, but as the papers show, many techniques lead to improved data quality.

Contact Information

Geoffrey Paulin, Ph.D.
Senior Economist
Division of Consumer Expenditure Surveys
www.bls.gov/cex
paulin.geoffrey@bls.gov