Imputation in a National Health Survey: Balancing Data Quality with Respondent Burden in the Medical Expenditure Panel Survey (MEPS)

Emily M. Mitchell, Jerrod Anderson, and Samuel H. Zuvekas
Center for Financing, Access and Cost Trends, Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20850

Abstract
The Medical Expenditure Panel Survey (MEPS) collects payment data from households for reported medical events. This household-reported data, however, is often incomplete and inaccurate. More accurate data for medical services are collected from households' medical providers. Provider-reported payment information is applied to medical events when available. For events without provider data, but with complete household data, household payment data are applied. For the remaining cases, incomplete payment data are imputed. In this analysis, we assess the impact of reducing data collected from the household that is most likely to be inaccurate. Doing so would reduce the burden on the household respondent, but would rely more heavily on imputation to complete the missing data. We assess the accuracy of household-reported data by comparing it to provider-reported data. We then compare imputed household data to actual provider data to assess the performance of the imputation algorithm. Finally, we simulate imputing additional household expenditures to assess the impact of reducing payment data collected from the household on the accuracy of MEPS expenditure estimates.

Key Words: Medical Expenditure Panel Survey; survey data; imputation; measurement error; respondent burden

1. Background

The Medical Expenditure Panel Survey (MEPS) is a nationally representative survey of the U.S. civilian non-institutionalized population that has been conducted annually since 1996. Each year, MEPS collects utilization and expenditure data for healthcare events, which is then edited, processed, and disseminated as public use files (Cohen, 1997). These files are commonly used by economic and health services researchers to conduct statistical or econometric analyses related to healthcare services and costs. In fact, MEPS data are "the number one single source of data for papers published in *Health Affairs*,"[1] one of the leading health services research journals.

The Household Component (HC) survey of MEPS is conducted by interviewing households five times over approximately two years. In a typical interview, one respondent will answer questions related to healthcare use and expenditures (defined as payments in MEPS) for all persons living in the household during the reference period.

[1] Berk, M. and Wilensky, G. How to make the Medical Expenditures Panel Survey even more useful. *Health Affairs blog*.

In addition to general questions about socio-demographics, health insurance status, and health status, respondents are asked to enumerate their healthcare events during the reference period. For each reported event, they are asked additional details about the event, including date of service, location of service, type of provider, and detailed information on charges and payments. For each reported medical event, respondents are asked how much was paid by each available payment source, including out-of-pocket payments, payments by private insurance, Medicaid, Medicare, TRICARE, VA insurance, Worker's comp, and any other potential sources. However, households rarely have complete information on all payments for each medical visit. For instance, Medicaid recipients typically do not receive an explanation of benefits (EOB) form, thus are often unaware of how much Medicaid is actually paying to the medical providers. In 2016, respondents reported complete expenditure data for only 16% of medical events[2], and persons on Medicaid had complete expenditure data for less than 1% of these events.

In order to gather more accurate expenditure data, a follow-back survey of medical providers, called the Medical Provider Component (MPC), is conducted. After receiving permission from households, their medical providers are contacted to gather detailed expenditure data for survey participants. MPC data are collected for hospital-based events (inpatient stays, ER visits, and outpatient visits), as well as for office-based doctor visits. Data collected from the MPC is considered the 'gold standard' for MEPS expenditure data. When available, MPC expenditure data is applied to medical events. When MPC data is not available, expenditure data from the household is used. When household expenditure data is incomplete, any remaining expenditures are imputed at the event level using predictive mean matching algorithms.

Due to the detailed nature of MEPS interview questions, the interview process can be quite burdensome for household respondents, particularly those with many medical visits to report. For every medical event reported, respondents are asked to report total charges for the event, out-of-pocket payments, and third-party payments (e.g. private insurance, Medicare, Medicaid). Total expenditures are then calculated by summing the out-of-pocket and third-party payments. As mentioned previously, households rarely have complete and accurate expenditure data for all payment sources. In addition, if MPC data is collected for an event, then any household expenditure data is essentially discarded for that event. On the other hand, when the household reports partial data and MPC data is not available, the partial reports are included in the imputation process. Table 1 gives the distribution of the source of expenditure data by event type, for 2013-2016. The 'Partially imputed' group includes those events for which the household provides data for some payment sources (usually out-of-pocket payments).

Table 1. Distribution of source of expenditure data by event type, 2013-2016

	Hospital-based events			
Source of expenditure data	**ER visits**	**Hospital Inpatient Stays**	**Outpatient visits**	**Office-based doctor visits**
Not eligible for imputation*	6%	0%	1%	2%
Complete HC data	5%	4%	5%	10%
Complete MPC data	51%	56%	57%	35%
Fully Imputed	32%	34%	30%	37%
Partially imputed	6%	6%	6%	16%

* Events not eligible for imputation include events with $0 expenditures, such as flat fee 'leaf' events and ER visits that are included in a subsequent inpatient stay.

[2] 'Medical events' encompass inpatient stays, ER visits, outpatient visits and office-based doctor visits.

As shown in the table, MPC data are used for over half of hospital-based events and a third of office-based doctor visits. While respondents are asked to report payment data for all events, only 5% of hospital-based events and 10% of office-based visits utilize complete HC data. In this study we consider options to reduce burden on respondents by reducing the amount of expenditure data requested during the household interview. Instead of using household-reported data, these expenditures would be imputed. All respondents would benefit from a reduction in respondent burden; at the same time, only a small percentage of events (namely, those with complete HC data) would be noticeably affected in terms of data quality.

2. Analysis

As mentioned previously, households are asked to report total charge, out-of-pocket payments, and third-party payments for each event. Although households rarely have complete expenditure data for all payment sources, they often report out-of-pocket expenditures (for around 93% of events). Although total charge is only reported for about 17% of events, it has a strong predictive capacity for imputing total expenditures due to its high correlation with total expenditures. Thus, we consider the following scenarios in which we eliminate interview questions regarding third-party payments, total charge, and out-of-pocket payments:

A. Only collect out-of-pocket and total charge from HC
B. Only collect out-of-pocket from HC
C. Don’t collect any expenditure amounts from HC

For this analysis, we first assess the accuracy of HC data, using MPC data as the gold standard. We then assess the accuracy of the imputation algorithm, and compare it to the accuracy of the HC-reported expenditure data, considering each of the three above scenarios. Finally, we simulate potential changes to national estimates of total expenditures and average expenditures per event if we were to impute the events that are currently using HC expenditure data. We run these analyses using MEPS data from 2013-2016. All analyses are conducted in R version 3.6.0.

2.1 How accurate is HC data?

Even when households do report payment data, it may not be accurate. For instance, households may incorrectly assume that the total charge is equivalent to total expenditures (defined as payments in MEPS), when in fact a discount adjustment is often applied due to contractual agreements between insurers and “in-network” providers. To assess the accuracy of HC-reported payments, we consider the set of events where both the HC and MPC report complete expenditure data. Table 2 shows the percentage of events (out of events with complete HC and MPC data) for which the household-reported expenditure data is within $5 or 10% of the MPC-reported expenditure data (our criteria for “accurate” reporting), for data years 2013-2016. Table 3 gives the mean expenditures for each event, as reported by the HC and MPC.

Table 2. Percentage of events with accurate HC data (within $5 or 10% of MPC), out of events with complete HC and MPC data, 2013-2016. ER = Emergency Room Visits; HS = Hospital inpatient Stays, MV = Office-Based Medical Visits; OP = Outpatient Events

	N (%) with complete data*		Percent of events with accurate HC data			
Event type	Payments	Charges	Out-of-pocket payments	Third-party payments	Total expenditures	Total charge
ER	2,280 (8%)	1,778 (6%)	49%	47%	28%	44%
HS	679 (6%)	579 (5%)	56%	40%	36%	57%
MV	26,654 (5%)	23,156 (5%)	65%	46%	39%	57%
OP	5,267 (9%)	4,627 (8%)	62%	36%	30%	48%

* Complete payment data refers to events with complete HC and MPC expenditures for all sources. Complete charge data includes events with complete payment data, that also have complete charges for the HC and MPC.

Table 3. Mean payments and charges per event, for events with complete HC and MPC data, 2013-2016. ER = Emergency Room Visits; HS = Hospital inpatient Stays, MV = Office-Based Medical Visits; OP = Outpatient Events

	Out-of-pocket payments		Third-party payments		Total expenditures		Total charge	
Event Type	HC	MPC	HC	MPC	HC	MPC	HC	MPC
ER	$438	$123	$841	$653	$1,280	$776	$2,496	$2,912
HS	$1,151	$352	$14,924	12,243	$16,076	$12,595	$32,420	$39,569
MV	$65	$45	$202	$173	$267	$218	$457	$498
OP	$157	$91	$979	$821	$1,136	$912	$2,320	$2,596

From Table 2 we see that households accurately report out-of-pocket payments for between about half and two-thirds of events, and are most likely to report accurately for office-based and outpatient visits. Household-reported third-party payments are accurate for fewer than half of reported events, and accuracy of total charges ranges from 44-57%, depending on the event type. Table 3 suggests that, on average, households tend to over-report expenditures and under-report charges. This could be a consequence of the aforementioned confusion between charges and payments, in which households may assume that total payments are equal to total charges. This confusion could result in either an under-report of charges (if respondents know total expenditures and assume that charges are equivalent) or, more likely, an over-report of expenditures (if respondents know charges and assume that total expenditures are equivalent).

2.2 How accurate is imputation?

For this analysis, we again consider the set of events where both the HC and MPC have complete data. Before describing the details of this portion of our study, we first give a brief summary of the predictive mean matching algorithm, the current strategy for expenditure imputation in MEPS.

2.2.1 Predictive Mean Matching

Predictive mean matching (PMM) is an imputation algorithm that imputes observations with missing values (‘recipients’) by matching them to observations with observed values (‘donors’). In the PMM framework for expenditure imputation in MEPS, the donors are events with complete MPC data. Recipients are household events with incomplete or missing MPC data and incomplete HC data. Events with complete HC data do not enter into the imputation process. The process consists of four phases, which are detailed below:

1. **Model**

The first phase of PMM is to model the relationship between the outcome and predictor variables among the donor events, using multiple linear regression. For MEPS expenditure imputation, the outcome is the square root of total expenditures (the square root transformation was applied due to the skewed nature of expenditures). This process is conducted separately by event type, so that each model can have a different set of predictor values (for instance, length of stay is a predictor in the model for hospital inpatient stays). The following model is fit for each event type:

$$y_d = \boldsymbol{X}_d \beta + \epsilon$$

Where y_d is the square root of expenditures, $\boldsymbol{X}_d$ is a matrix of predictors, β is the vector of regression coefficients, and ϵ is the normally-distributed random error component. The d subscript indicates that this model is fit only on the donor events, since all recipient events have missing values of y.

2. **Fit**

The next step in PMM is to calculate the predicted outcome value ($\hat{y}$) for both donors and recipients, using the vector of estimated regression coefficients ($\hat{\beta}$) produced in step 1:

Donors: $\hat{y}_d = \boldsymbol{X}_d \hat{\beta}$
Recipients: $\hat{y}_r = \boldsymbol{X}_r \hat{\beta}$

3. **Match**

After a predicted outcome value for each observation has been calculated, each recipient is matched to a donor event with the closest predicted value:

$$\delta = min\,|\hat{y}_d - \hat{y}_r|$$

At this stage, additional steps are taken to help preserve the natural variation in the distribution of the outcome variable, such as limiting the number of times a donor can donate or randomly selecting a donor out of a pool of donors with similar predicted values.

4. **Allocate**

Once each recipient has been assigned a donor, the expenditure data from the donor is then applied to the recipient. In some cases, the recipient event contains partial expenditure data from the household. For instance, the household may report a total charge, but not third party payments, or the household may report an out-of-pocket amount, but not total charge or payments from other sources. In those cases, the household-reported data is retained, and the donor's expenditure data is used to fill in the incomplete portions of the payment data for that event.

2.2.2 Simulation: Event-level accuracy of HC vs. imputation

In order to compare the accuracy of imputation with that of the HC-reported expenditure data, we start with the set of events with complete HC and complete MPC data. Then, we remove the HC and MPC expenditure data for about half of the events to act as recipients in the simulation. We then impute total expenditures for these recipients and compare the accuracy of the new imputed values with the accuracy of the household-reported expenditures for that event. We repeat this process 100 times, splitting the data differently each time. As mentioned previously, we simulate the following scenarios:

A. Only collect out-of-pocket and total charge from HC
B. Only collect out-of-pocket from HC
C. Don't collect any expenditure amounts from HC

For each of the 100 simulations, the same set of 'recipients' is assessed across the four scenarios (Scenarios A, B, C, and original HC-reported data). Figure 1 displays the Mean Absolute Error (MAE) for each simulated scenario as well as the MAE when comparing the household-reported expenditure data to the MPC data ('Complete HC data'), where a lower MAE corresponds to more accurate data. Each point in the bee-swarm plots corresponds to a different random split of the data. Results are shown separately by event type and data year.

In general, we see that for most event types, complete HC data tends to be more accurate than the simulated scenarios, which require additional imputation. Scenario A, in which third-party payments are not collected from the household, is only slightly less accurate than the complete HC data in most cases. In addition, scenarios B and C tend to perform similarly with respect to accuracy of imputed total expenditures for HS, MV, and OP events. This is to be expected, since the household-reported out-of-pocket payments have only a small effect on the imputation of total expenditures. For ER visits, however, we see that Scenario C, in which no expenditure data is collected from the household, actually performs slightly better than the HC-reported data. This could be due to the nature of the relationship between certain ER events and hospital inpatient stays. For instance, when a patient is admitted to the ER and then transferred to an inpatient unit at the same hospital, often the payments and charges for the ER visit will be rolled into payments and charges for the inpatient stay. The household may report separate payments for the linked ER event and inpatient stay, whereas the MPC may combine all expenses into the bill for the inpatient stay, reporting $0 in expenses for the ER visit.

Overall, the loss of accuracy in the simulated scenarios is not drastically different from the complete HC data, and in the case of ER events, may prove to perform even slightly better.

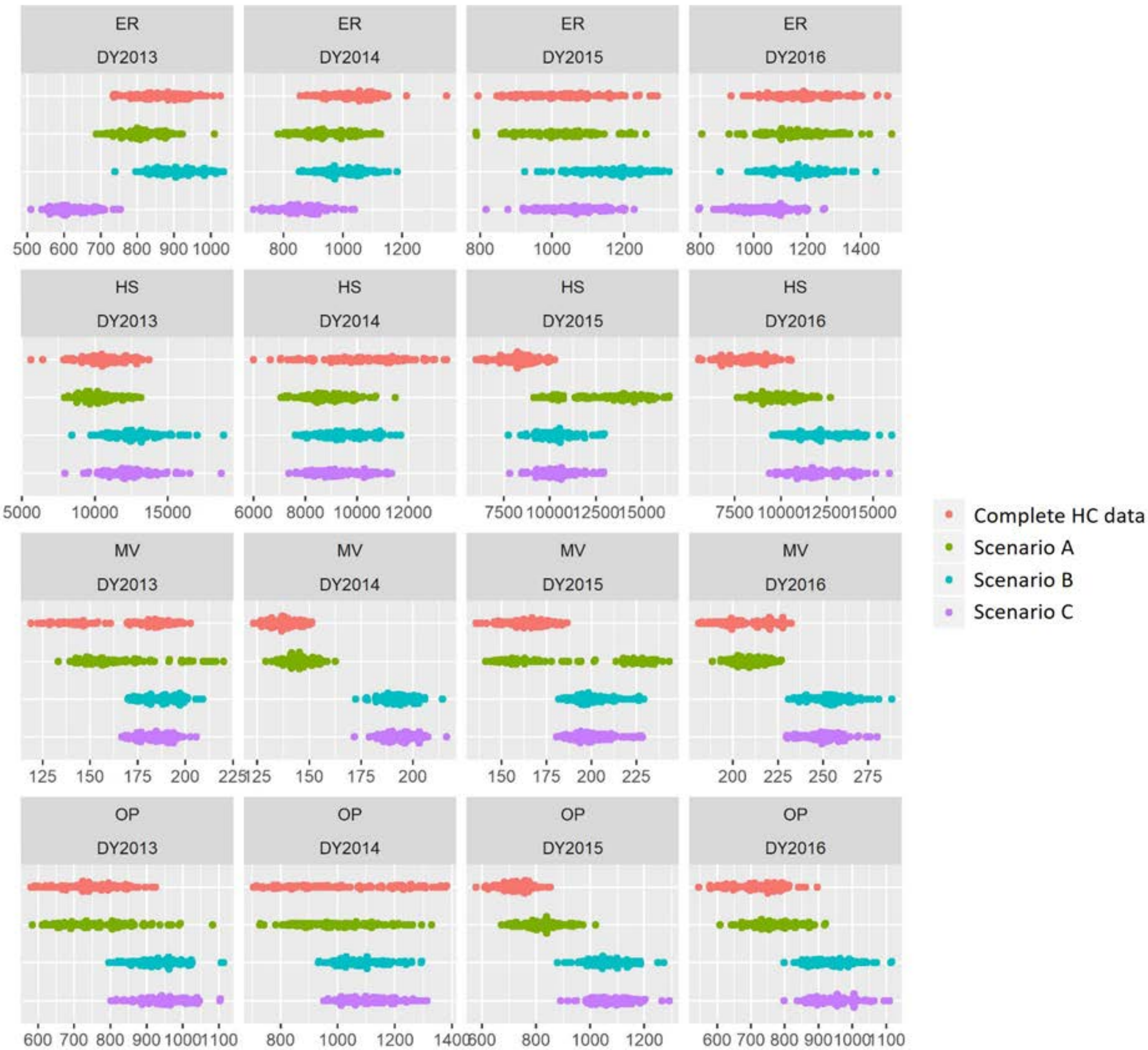

Figure 1. Mean absolute for each simulation run, by event type and year. ER = Emergency Room Visits; HS = Hospital inpatient Stays, MV = Office-Based Medical Visits; OP = Outpatient Events

2.3 How would increasing imputation affect national estimates?

In this section, we consider the potential changes to mean and total expenditures, for all events eligible for MPC data collection (specifically, ER visits, hospital inpatient stays, outpatient visits, and office-based doctor visits). Again, we compare the current strategy of using complete HC data (when available), to the three scenarios reducing the amount of household expenditure data collected. Figures 2 and 3 display the mean expenditure per event and total expenditures, respectively, by event type and data collection scenario considered, for data years 2013-2016. These Figures present results from a single simulation (differences between various runs of the simulation were negligible). The 95% confidence bands provided in the figures reflect sampling error from the complex survey design of MEPS.

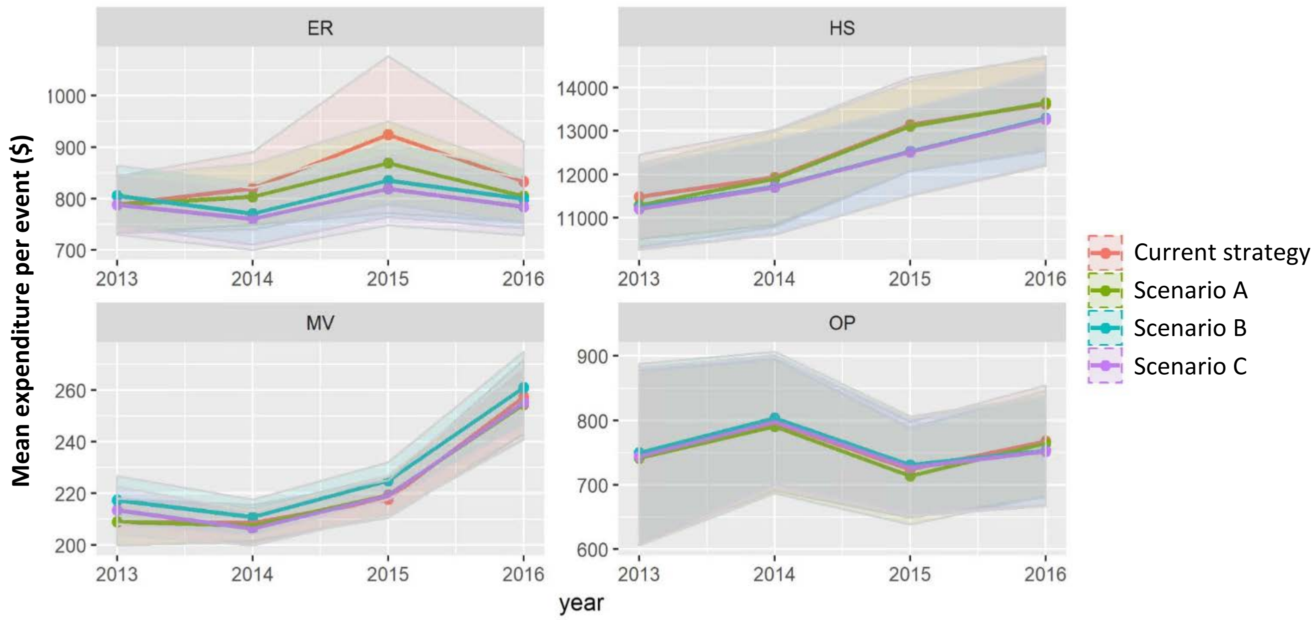

Figure 2. Mean expenditure per event and 95% confidence interval bands, by event type and collection scenario, 2013-2016.
ER = Emergency Room Visits; HS = Hospital inpatient Stays, MV = Office-Based Medical Visits; OP = Outpatient Events

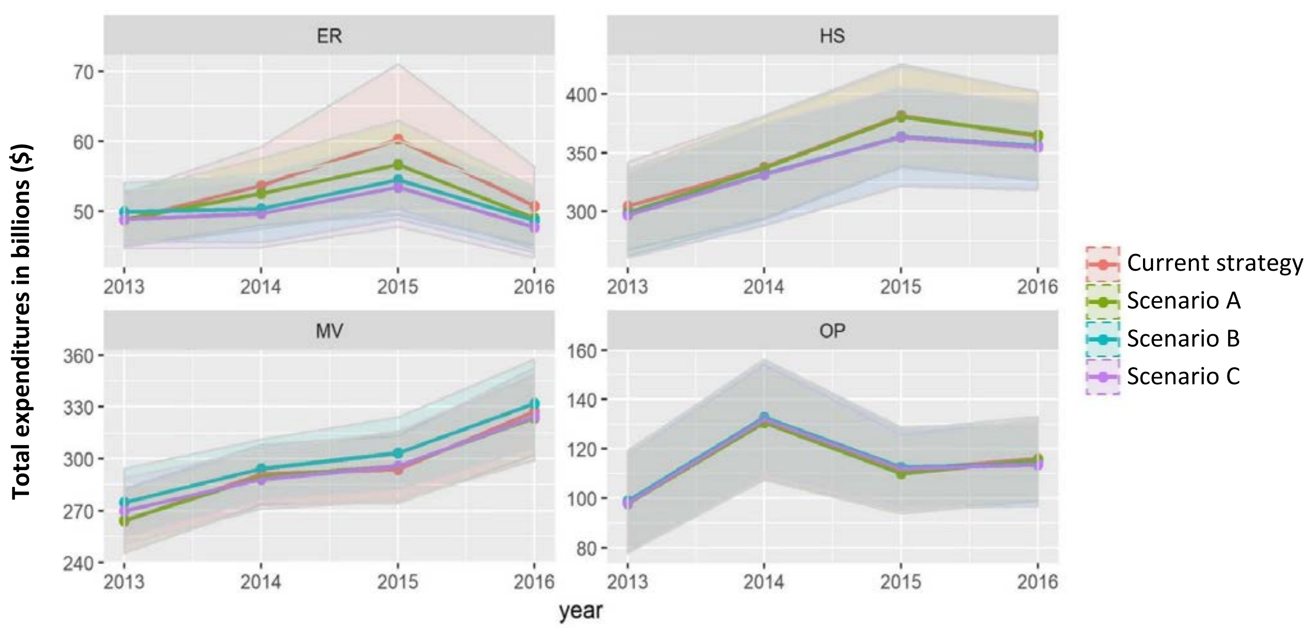

Figure 3. Total expenditures in billions and 95% confidence interval bands, by event type and collection scenario, 2013-2016.
ER = Emergency Room Visits; HS = Hospital inpatient Stays, MV = Office-Based Medical Visits; OP = Outpatient Events

These results suggest that introducing additional imputation by reducing the amount of household data collected has little impact on national expenditure estimates. In particular, there is virtually no impact on expenditure estimates for outpatient (OP) events. Hospital inpatient stays (HS) have a slight decrease in expenditure estimates for 2015 data when total charge is not collected from the household (Scenarios B and C), although the same trend is not as apparent for the other years. The three scenarios track closely to the current imputation strategy for office-based medical visits (MV), which have the highest percentage of events with complete HC data among these event types (and thus, the highest percentage of events impacted by the reduced data collection strategies). For ER visits, all of the scenarios appear to slightly lower the mean and total expenditure estimates for the years 2014-2016, although the confidence intervals are overlapping.

3. Conclusion

In summary, households often report out-of-pocket payments, but rarely report complete payment data from all sources. Compared to the predictive mean matching imputation algorithm, reported household data are slightly closer to MPC-reported expenditure data, based on the mean absolute error criterion. Simulations suggest that reducing the amount of household expenditure data collected would have only a slight impact on estimates of mean and total expenditures for the years 2013-2016. This minimal impact is most likely due to the small percentage of events that would be affected by the reduction in household data. Specifically, only around 10% of office-based doctor visits and 5% of the other event types would potentially experience a notable change in total expenditure estimates.

However, this study has some limitations. One such limitation is that we did not account for matching probability when assessing the accuracy of household-reported data. After collecting MPC data, each MPC-reported event needs to be matched to the appropriate HC-reported event. However, there is no deterministic identifier to match events from these two sources, and often, the number of reported events differs between the HC and MPC (Zuvekas, 2011; Zuvekas and Olin, 2009a, b). Thus, a probabilistic matching algorithm is implemented (Fellegi and Sunter, 1969), which attempts to match HC to MPC events based on characteristics such as event date, types of services received, and conditions treated. In this paper, we do not differentiate between ‘good’ and ‘bad’ matches. In other words, we may categorize HC-reported expenditure data as being inaccurate when compared to the MPC-reported expenditures for that event, when in fact, the HC event has just been linked with the incorrect MPC event.

Another limitation is that the selection of our study samples may not be representative of the entire set of events. For instance, we only use the set of events with complete HC and complete MPC data for the first two analyses (assessing accuracy of HC data and accuracy of imputation). However, results based on these events may not generalize to the events which have only partial HC data. In addition, this paper only assesses the impact on total expenditure estimates summed across all sources; accuracy by payment source is not assessed. In particular, while not collecting household-reported out-of-pocket expenditures may have minimal impact on overall expenditure estimates, estimates for out-of-pocket payments may be more susceptible to reduced accuracy.

Overall, this analysis demonstrates that reducing the amount of expenditure data collected from the household would reduce the accuracy of expenditure data on an event-level, but would have only a small impact on overall expenditure estimates. This is due in large part to the fact that only a small percentage of events have complete household data. One concern about respondent burden is that over the course of the five interviews, respondents may tend to omit events, since they have ‘learned’ that reporting events just extends the interview time and requires more effort on their part (Mitchell, Muhuri, and Machlin, 2016). Additional time during the interview as well as more cognitive capacity is required to answer specific questions about expenditures for each reported event. Our ability to reduce respondent burden in this area could increase the reporting of events, thus improving this vital estimate of utilization. While per-event expenditure data may be less accurate, overall utilization estimates may improve, which could result in better estimates of overall expenditures.

Some additional steps for future work include quantifying the level of respondent burden that is currently required for households to report all expenditure data. Respondent burden has multiple facets, including not only time during the interview, but also the additional time and cognitive burden required to prepare for the interview by collecting EOBs, logging into health portal systems, or interpreting medical bills, which in many cases is not a straightforward task. This analysis has helped quantify the potential impact of reducing household data collection on the accuracy of national expenditures. The next step is to quantify the impact of reduced HC data collection on reducing respondent burden, and subsequently, to assess the cost/benefit of these considered changes to the MEPS survey instrument.

Additionally, further steps could be taken to improve the imputation algorithm. Some potential avenues of consideration include replacing the linear regression model in the PMM framework with machine-learning algorithms with better predicative capacity, thus resulting in better matching between donors and recipients. Gaussian mixture models (Di Zio and Guarnera, 2009), sequential regression trees (Burgette and Reiter, 2010), and other machine learning algorithms (Bertsimas, Pawlowski, and Zhuo, 2017) are potential candidates. If improvements to the imputation algorithm can outperform household-reported data, then the choice to reduce data collected from households would be unambiguous.

References

Bertsimas, D., Pawlowski, C., & Zhuo, Y. D. (2017). From predictive methods to missing data imputation: an optimization approach. *The Journal of Machine Learning Research, 18*(1), 7133-7171.

Burgette, L. F., & Reiter, J. P. (2010). Multiple imputation for missing data via sequential regression trees. *American journal of epidemiology*, *172*(9), 1070-1076.

Cohen, J. *Design and Methods of the Medical Expenditure Panel Survey Household Component*. MEPS Methodology Report No. 1. AHCPR Pub. No. 97-0026. Rockville, MD. Agency for Healthcare Policy and Research, 1997. https://meps.ahrq.gov/data_files/publications/mr1/mr1.pdf

Di Zio, M., & Guarnera, U. (2009). Semiparametric predictive mean matching. *AStA Advances in Statistical Analysis*, *93*(2), 175-186.

Ezzati-Rice, T.M., Rohde, F., Greenblatt, J., *Sample Design of the Medical Expenditure Panel Survey Household Component, 1998-2007.* Methodology Report No. 22. March 2008. Agency for Healthcare Research and Quality, Rockville, MD. https://meps.ahrq.gov/data_files/publications/mr22/mr22.pdf

Fellegi, I. P., & Sunter, A. B. (1969). A theory for record linkage. *Journal of the American Statistical Association*, *64*(328), 1183-1210.

Machlin S.R., Chowdhury S.R., Ezzati-Rice T., DiGaetano R., Goksel H., Wun L.-M., Yu W., Kashihara D. *Estimation Procedures for the Medical Expenditure Panel Survey Household Component*. Methodology Report No. 24. September 2010. Agency for Healthcare Research and Quality, Rockville, MD. https://meps.ahrq.gov/data_files/publications/mr24/mr24.pdf

Mitchell, E.M., Muhuri, P., and Machlin, S. (2016). Changes in Health Care Use Reporting in the 2nd vs. 1st year: MEPS Household Component Overlapping Panel Design. In JSM Proceedings, Survey Research Methods Section. Chicago IL: American Statistical Association. 794-806.

Stagnitti, M.N., Beauregard, K., and Solis, A. *Design, Methods, and Field Results of the Medical Expenditure Panel Survey Medical Provider Component (MEPS MPC) - 2006 Calendar Year Data*. Methodology Report No. 23. November 2008. Agency for Healthcare Research and Quality, Rockville, MD. https://meps.ahrq.gov/data_files/publications/mr23/mr23.pdf

Zuvekas, S. H. (2011). The effects of recall length and reporting aids on household reporting of health care events in the medical expenditure panel survey. *Journal of Economic and Social Measurement*, *36*(4), 321-343.

Zuvekas, S. H., & Olin, G. L. (2009a). Validating household reports of health care use in the medical expenditure panel survey. *Health services research*, *44*(5p1), 1679-1700.

—(2009b) Accuracy of Medicare expenditures in the medical expenditure panel survey. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, *46*(1), 92-108.