Journal of Data Science 24 (3), 616–629
July 2026

DOI: 10.6339/26-JDS1242
Data Science in Action

Predicting "Yes": Machine Learning and Diverse Data to Boost Respondent Cooperation

Rashi Saluja[1,*], Hanyu Sun[1], Gizem Korkmaz[1], Jill Carle[1], Ryan Hubbard[1], Brad Edwards[1], and Rick Dulaney[1]

[1]*Westat, Bethesda, MD, United States*

Abstract

In the evolving field of survey research, leveraging machine learning to predict response behavior has transformative potential for the efficiency of survey operations. Integrating multiple data sources may improve response prediction by providing more nuanced insights for household outreach. This study presents a model-driven approach to enhancing respondent cooperation in the Medical Expenditure Panel Survey (MEPS) by combining features from disparate data sources. MEPS is a longitudinal household survey with 5 rounds of interviewing over 2.5 years. Its sample is derived prior National Health Interview Survey (NHIS) participants. MEPS Round 1 response rates are critical for sustaining representativeness throughout each panel. In this study, we constructed a multimodal machine learning model to predict (1) the likelihood of a positive response for an upcoming contact attempt and (2) the likelihood that new panel households complete a Round 1 interview. The model integrates tract-level data from the American Community Survey (ACS), outcomes from the Advance Call Records (ACR) made prior to MEPS Round 1, and paradata from the early contact period. We also explored the relative contributions of these sources to model performance. Our model aims to help manage field labor by identifying complex cases needing specialized support. It can also assist in determining the optimal mode for the next contact to increase the chance of a completed interview. Beyond improving MEPS operations, this study offers a roadmap for incorporating additional data sources to support fieldwork.

Keywords *case prioritization; machine learning; paradata; predictive modeling*

1 Introduction

Probability-based surveys are a foundational tool for producing nationally representative statistics on health, employment, expenditures, and a wide range of social indicators. In recent decades, however, survey organizations have faced persistent challenges: declining response rates, rising costs of fieldwork, and increased difficulty in contacting and engaging households despite the use of multiple outreach modes such as telephone, mail, and in-person visits (Dillman et al., 2014). These pressures have prompted survey research organizations to explore adaptive and responsive design strategies that can make more efficient use of limited field resources while preserving representativeness (Schouten et al., 2017).

One promising avenue for improving survey operations is the use of paradata. Contact history is one type of paradata - broadly defined as data about the data collection process

*Corresponding author. Email: rashisaluja@westat.com.

Received April 16, 2026; Accepted July 14, 2026

itself (Chun et al., 2017), it includes contact-level details on when, how, and how often sampled households are contacted, along with the outcomes of those attempts. Contact history contains a wealth of detail on interviewer effort and respondent behavior that, if systematically analyzed, can inform adaptive case management strategies (Chun et al., 2017). For example, patterns in contact mode sequences may reveal which households are more likely to cooperate after certain types of contacts, or which require additional effort to encourage participation.

Advances in machine learning provide new opportunities to analyze the contact history data. By integrating the contact history data with auxiliary information such as Census tract-level characteristics or interviewer observations, machine learning models can uncover nonlinear relationships and interactions that are not easily captured through conventional regression methods. Recent studies have demonstrated the potential of predictive modeling for predicting response status and guiding adaptive designs in large-scale surveys (Eckman and Koch, 2019; Schouten et al., 2021). However, empirical evidence does not yet provide a clear consensus on how to best utilize multiple data sources to support case prioritization in longitudinal household surveys.

The Medical Expenditure Panel Survey (MEPS) provides a useful setting to investigate these issues. MEPS is a nationally representative longitudinal survey of U.S. households designed to measure healthcare utilization and expenditures. The survey follows sampled households over five interview rounds spanning 2.5 years, with the sample drawn from households that participated in the National Health Interview Survey (NHIS) in the previous year. Achieving high Round 1 cooperation rates is essential for ensuring representativeness and minimizing nonresponse bias throughout the panel. Declines in early cooperation increase the risk of attrition and compromise data quality over time.

This paper develops and evaluates a machine learning framework for predicting respondent cooperation in MEPS Round 1. Specifically, we address two research questions: (1) What is the probability that an upcoming contact attempt will result in a completed interview? and (2) What is the probability that a new panel household completes the Round 1 interview? To answer these questions, we construct predictive models that integrate three complementary sources of data: (i) tract-level sociodemographic indicators from the American Community Survey (ACS), (ii) outcomes from the Advance Call Records (ACR) (described in detail in 3.1.2) prior to the start of Round 1, and (iii) contact history data from early contact attempts. By combining these information streams, our approach seeks to provide a richer and more nuanced basis for predicting respondent cooperation.

In addition to predictive modeling, we examine the relative contributions of each data source to model performance, highlighting which variables are the most informative for case management. We further explore the use of clustering methods to group households with similar contact histories, with the goal of identifying distinct patterns of cooperation that may inform tailored contact strategies. Taken together, these analyses aim to support more efficient allocation of data collection resources, reduce wasted effort on low-probability cases, and ultimately improve completion rates through data-driven decision-making.

The remainder of this paper is organized as follows. Section 2 reviews related literature on predictive modeling and adaptive survey design. Section 3 describes the MEPS data sources, feature construction, and modeling framework. Section 4 presents the results of our predictive and clustering analyses. Section 5 discusses implications for survey operations, limitations, and directions for future research.

2 Related Work

A growing body of research in survey methodology explores how paradata and auxiliary information can be leveraged to predict survey outcomes and guide adaptive fieldwork strategies. Traditional approaches to handling nonresponse have relied primarily on weighting or imputation techniques that correct for observable differences between respondents and nonrespondents (Groves and Peytcheva, 2008; Kalton and Flores-Cervantes, 2003). However, these approaches do not fully exploit the operational data generated during data collection, such as contact records and interviewer notes, which can offer valuable insight into respondent behavior and interviewer effort.

Recent work has focused on using such paradata for predictive modeling and adaptive survey design. Schouten et al. (2017) and Kreuter and Olson (2013) highlight how adaptive survey designs can dynamically allocate resources or change data collection strategies in response to real-time indicators of survey progress. Paradata are particularly useful in this context because they capture the sequence, mode, and timing of contact attempts—variables that can be used to infer respondent availability or willingness to participate. Studies such as Bianchi et al. (2015) and Rybak (2023) further show that mode sequencing and mixed-mode strategies influence both participation rates and cost efficiency, reinforcing the value of data-driven design decisions. For example, a recent article by Wuyts et al. (2023) presents one of the most comprehensive applications of predictive modeling to paradata. Using call record data from the European Social Survey, the authors develop models to predict the probability that a case will require additional fieldwork effort, such as multiple visits or extended interviewer contact. They compare a range of machine learning algorithms—including gradient boosting, random forests, and logistic regression—and demonstrate that paradata can substantially improve prediction accuracy compared to demographic and design variables alone. Their findings suggest that integrating paradata with contextual information offers an efficient way to identify high-effort cases early in the field period.

One of the earliest applications of sequence analysis to call record data was (Kreuter and Kohler, 2009), who examined contact sequences across 14 countries in the European Social Survey and developed indicators to characterize non response adjustment potential. While studies like Wuyts et al. (2023) and earlier work by Durrant et al. (2019) show that predictive models can successfully estimate response propensities using call record data, most prior research focuses on binary outcomes (response versus nonresponse). Few studies model respondent cooperation as a sequence of contact patterns or use clustering to detect behavioral patterns.

Our work contributes to this literature in several ways. First, instead of predicting only whether a case responds, we model the cooperation trajectory by grouping contact sequences into behavioral clusters and then predicting cluster membership. This approach captures variation in both cooperation and interviewer effort, offering a more nuanced view of household participation dynamics. Second, we integrate multiple data sources— Census tract-level characteristics, Advance Call Record (ACR) outcomes prior to the Round 1 of MEPS, and early contact records of the fielding period of the Round 1—into a unified predictive framework. Combining these sources enhances the model's explanatory power and provides richer context for understanding differences in respondent behavior. Third, by using multinomial logistic regression with regularization, we retain interpretability of model coefficients while accommodating complex interactions across predictors. Our design allows us to benchmark all results against a high-cooperation, low-burden reference cluster (Cluster 5), facilitating direct comparison across groups. Finally, whereas prior studies have primarily demonstrated predictive feasibility, we link

our findings explicitly to operational applications, showing how predicted cluster membership can inform tailored contact approach in future survey rounds.

3 Data and Methods

This study integrates multiple data sources to develop predictive models of respondent cooperation in MEPS. In this section, we describe the survey design and auxiliary data, followed by the modeling framework used to predict outcomes of contact attempts and Round 1 interview response status.

3.1 Data Sources

3.1.1 Medical Expenditure Panel Survey (MEPS)

The MEPS Household Component (MEPS-HC) is a nationally representative longitudinal survey of U.S. households conducted by Westat for the Agency for Healthcare Research and Quality (AHRQ). The survey collects data on healthcare use, expenditures, insurance coverage, and access to care. A new MEPS panel is initiated each year, with each panel's sample drawn from NHIS respondents from the prior year, resulting in overlapping panels running concurrently. Households are drawn from NHIS in the prior year and followed over five interview rounds spanning 2.5 years. Achieving high Round 1 response rates is critical for maintaining representativeness and minimizing attrition over subsequent rounds.

Our analytic dataset consists of 8,593 households sampled for Spring 2022 Round 1. These households had at least two contact attempts in the fielding period from 10,071 eligible cases. Of these, 4,707 households completed the Round 1 interview, while 3,886 remained not responding, yielding a completion rate of 54.8%.

The MEPS sample includes household- and respondent-level demographic information collected from the NHIS sampling frame and MEPS case information. In our analytic dataset, available demographic variables included household language, race, household type, respondent relationship to the household, marital status, gender, age, and number of household members. However, some demographic information may be limited for non-responding households at the time of Round 1 fielding. For this reason, we supplemented the MEPS case-level variables with ACS tract-level measures to capture broader neighborhood socioeconomic and demographic context for all sampled households, including both responding and non-responding cases.

MEPS communication materials and the survey instrument are available in English and Spanish. Most MEPS interviews are completed in English, while approximately 5% are completed in Spanish. A very small number of interviews are conducted in other languages when trained interviewers are available. In some cases, initial communication may also be attempted in other languages, such as Vietnamese or Chinese, through letters or interviewer contact. If communication cannot be completed in English or Spanish, interviewers may present a language card to help identify the household's preferred language. Because MEPS is a household survey, interviewers also attempt to identify a knowledgeable household respondent who speaks English or Spanish or who can assist with translation for another household member.

3.1.2 Advance Call Records (ACR)

Prior to Round 1, interviewers conducted advance telephone calls to 51.1% of sampled households to confirm contact information and establish eligibility; however, 5.5% of sampled households

had no available phone number and therefore could not be contacted by telephone. We use information from these calls to capture early indicators of contactability, respondent availability, and expected cooperation before in-person fieldwork began.

Four analytic variables were derived from the ACR data. First, the recorded reason that advance contact could not be completed was grouped into a categorical measure of advance call outcome, including categories such as left message, final no answer, hung up, wrong number, no number, and other. Second, interviewer-recorded best times to reach the respondent were coded into a respondent availability measure based on reported day and time preferences, including weekday, weekend, and time-of-day categories. Third, whether the interviewer obtained updated contact information was coded as a categorical variable indicating yes, no, don't know, or refusal. Fourth, the interviewer's assessment of the household's likelihood of participation was retained as an ordinal 1–5 measure, where 1 indicated very likely to participate and 5 indicated not at all likely to participate.

These ACR-derived variables were included because they provide pre-fielding information about the likelihood of reaching a household and its expected cooperation, which may not be captured by demographic or neighborhood-level measures alone.

3.1.3 Census Tract-Level Characteristics

To capture broader contextual information, we linked MEPS households to tract-level characteristics from the American Community Survey (ACS). These variables supplemented the household- and respondent-level information available from MEPS, particularly because some demographic information may be limited or incomplete for nonresponding households during Round 1 fielding.

The ACS indicators included poverty rate, educational attainment, housing characteristics, racial/ethnic composition, and language use. Language was included because communication barriers may affect the ability to reach households, schedule interviews, and complete Round 1 data collection. These variables were selected because they represent neighborhood-level factors that may influence survey response propensity beyond what is captured by MEPS case information, Advance Call Records, and paradata.

3.1.4 Contact Records

Contact records document the process of data collection, including the mode, timing, and outcome of each contact attempt. For each household, we extract information on whether attempts were made by phone, text message, email, mail, or in-person visit. Contact attempts were primarily conducted in English and Spanish, with occasional use of other languages when needed and when trained interviewers or translated materials were available. These sequences provide insight into the intensity and mode of interviewer effort. Figure 1 illustrates the distribution of contact modes across attempts in Round 1. Completed cases (green) tend to reach resolution with fewer contact attempts, while incomplete cases (red) are more prevalent among households requiring a higher number of contacts.

3.2 Methods

We developed a multimodal predictive modeling framework with two primary outcomes: (1) whether an upcoming contact attempt would result in a completed interview, and (2) whether a sampled household would complete the Round 1 interview. The modeling proceeded in three

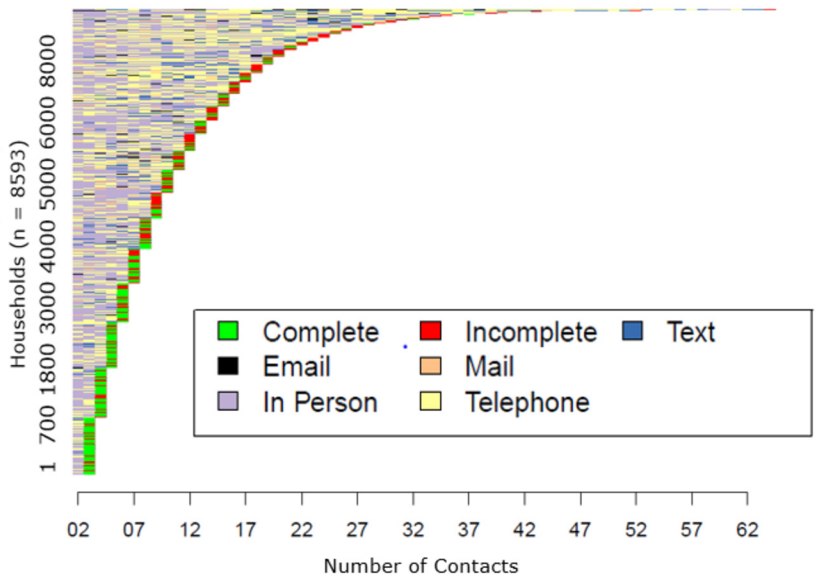

Figure 1: Distribution of contact modes in MEPS Round 1. Each bar represents a household's sequence of contact attempts, colored by mode, with outcomes indicating whether the case was ultimately completed or remained incomplete.

stages: feature construction, clustering analysis, and predictive modeling. First, we transformed information from ACS, ACR, MEPS, and contact records into analytic features used as independent variables in the models. Second, we used clustering analysis to construct the dependent variable for the multinomial logistic regression model: each household's ordered contact mode sequence cluster. Third, we used predictive modeling to examine which demographic, contextual, and contact-related features were associated with different contact sequence patterns and completion outcomes.

3.2.1 Feature Construction

We combined features from the ACS, ACR, and contact records. For ACS, we included sociodemographic variables at the Census tract-level. From ACR, we included outcomes and interviewer observations from advance calls. From contact records, we generated both static features (counts of attempts by mode) and dynamic features (sequence-based representations of contact histories). Table 1 summarizes the main features used in the models.

3.2.2 Clustering Analysis

To explore heterogeneity in household contact patterns, we applied sequence analysis and clustering to the contact records. Each household was represented by the ordered sequence of contact modes (e.g., phone → SMS → email). Sequence similarity was computed using Optimal Matching, and households were grouped using Partitioning Around Medoids (PAM). To determine the number of clusters, we examined an elbow plot of within-cluster dissimilarity across values of k. The plot showed dips at k=3 and k=6; however, solutions at these values did not yield operationally distinguishable groups upon visual inspection of the resulting contact sequence patterns. A solution of k=5 was selected as it produced clusters with meaningfully distinct cooperation

Table 1: Summary of feature sources and examples.

Source	Example Features
American Community Survey (ACS)	Population density, median household income, educational attainment
Advance Call Records (ACR)	Advance call outcome, best time to reach respondent, updated contact information obtained, interviewer-assessed likelihood of participation
Paradata	Number of contacts by mode, time of day of attempts, sequence of contact modes

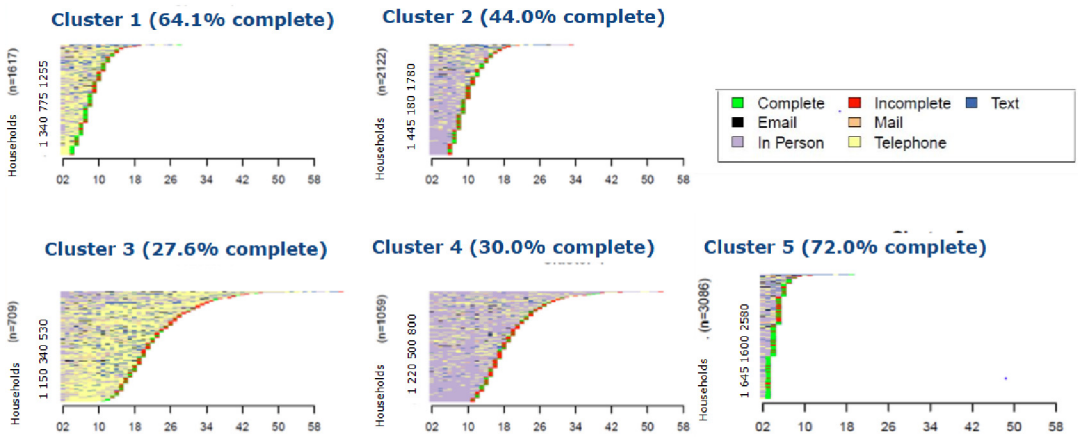

Figure 2: Completion rates by household contact sequence cluster.

profiles - varying in dominant contact mode, sequence length, and completion rate - that were both statistically reasonable and interpretable for field operations. Figure 2 summarizes the completion rates observed across the five clusters.

Cluster 1 (64.1% complete) was characterized by relatively successful, telephone-oriented sequences that often reached resolution after a moderate number of contacts. Cluster 2 (44.0% complete) showed more mixed contact histories and more moderate outcomes, with neither especially efficient nor especially unsuccessful sequences. Cluster 3 (27.6% complete) consisted of long contact sequences, suggesting households that required repeated follow-up and were more time-consuming to work. Cluster 4 (30.0% complete) was characterized by a greater reliance on in-person contacts, indicating a more resource-intensive and potentially more costly field effort. Cluster 5 (72.0% complete) represented the most efficient pattern, with short contact sequences, fewer attempts, and the highest completion rate. Together, these clusters provide an interpretable summary of distinct operational patterns of interviewer effort and household cooperation.

Cluster-level differences were further examined using multinomial logistic regression to assess how demographic and contextual features predicted cluster membership. This analysis provided insights into which types of households are most likely to follow particular response trajectories.

3.2.3 Predictive Modeling

We estimated multinomial logistic regression (MLR) models to examine how demographic and contact-related features differentiate households with distinct cooperation patterns. MLR is well suited for categorical outcomes with multiple classes, allowing us to compare households assigned to different contact sequence clusters. Cluster 5, which had the highest observed completion rate and the lowest number of contact attempts per household, was set as the reference category for interpretation of coefficients. We applied cross-validation to assess model stability and used Lasso regularization to aid in variable selection and prevent overfitting. The initial feature set included administrative and operational fields from the contact records system, such as interviewer identifiers, system-generated timestamps, task and activity log IDs, and case routing fields. Several of these were excluded prior to modeling due to high rates of missingness that could not be reliably imputed, or because they were administrative identifiers not conceptually linked to cooperation propensity. Of the remaining analytic variables, Lasso regularization further reduced the feature set by shrinking coefficients of low-signal predictors to zero. The variables shown in Figure 3 represent those with non-zero coefficients after regularization and showed distinctive patterns in different clusters. Model performance was evaluated by predictive accuracy, and coefficients were examined to interpret how features such as race/ethnicity, appointment timing, and interview language influence the likelihood of belonging to a given cluster.

In summary, the integration of multiple data sources and analytic strategies enables a comprehensive assessment of response propensity. The combination of predictive modeling and clustering provides both operational tools for case prioritization and conceptual insights into patterns of respondent cooperation.

4 Results

This section presents the results in three steps. First, we summarize the five contact sequence clusters and their completion rates. Second, we present the multinomial logistic regression (MLR) results, which estimate how household, demographic, and operational features differentiate each cluster from the reference group, Cluster 5. Third, we examine predicted probabilities to provide a more interpretable view of how key variables are associated with cluster membership. In Figure 3, positive coefficients indicate that a feature is more strongly associated with membership in a given cluster relative to Cluster 5, while negative coefficients indicate that the feature is less strongly associated with that cluster relative to Cluster 5.

4.1 Clustering Summary

As described in Section 3, sequence analysis identified five distinct clusters of household contact patterns. Completion rates varied substantially across clusters, from 27.6% in Cluster 3 to 72.0% in Cluster 5. Cluster 5 required the fewest contact attempts per household, making it both the highest cooperation and lowest burden group. For subsequent regression analysis, Cluster 5 was used as the reference category.

4.2 Multinomial Logistic Regression Results

The MLR model provided insights into which features differentiated households across clusters, using Cluster 5 as the reference category because it had the highest completion rate and required

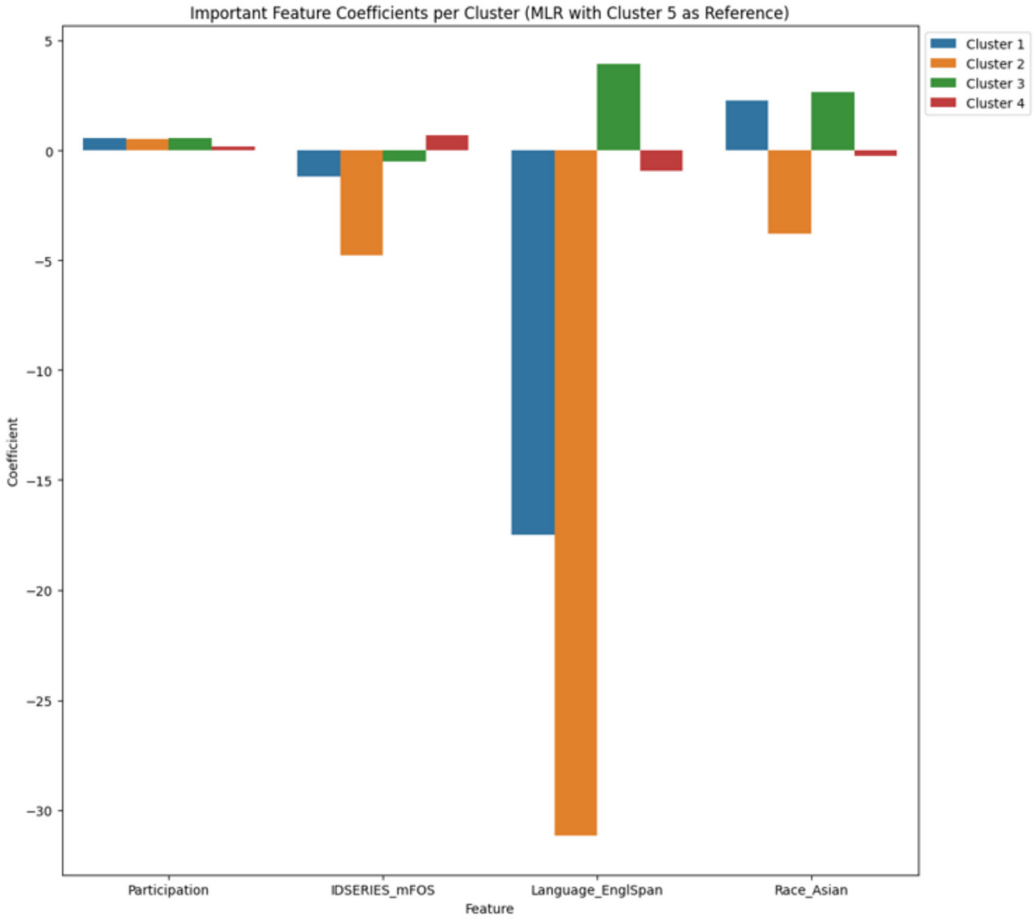

Figure 3: Important feature coefficients per cluster (MLR with Cluster 5 as reference).

the fewest contact attempts. Therefore, the coefficients in (Figure 3) should be interpreted as differences in the likelihood of belonging to Clusters 1–4 relative to Cluster 5, after accounting for other variables in the model.

Across all clusters, the interviewer-assessed likelihood of participation (from the ACR) showed modest positive associations relative to Cluster 5, though differences were small. This reflects that even lower-cooperation clusters contained households that interviewers rated as somewhat willing to participate, suggesting that interviewer assessment alone is not sufficient to distinguish high-cooperation households from those requiring more effort. Cluster 2 households were notably less likely to be recorded as responsive in the mobile field office system (MFOS) - the digital case management platform used by field interviewers to log contact attempts and outcomes - compared to Cluster 5; however, this distinction is likely operational, reflecting how interactions are recorded or conducted by field interviewers (FIs), rather than a true difference in respondent behavior. Language patterns also emerged: Clusters 1 and 2 contained significantly fewer English–Spanish bilingual households in their census tracts than Cluster 5, suggesting that language resources may play a role in cooperation. In terms of race and ethnicity, Clusters 1 and 3 included a greater proportion of Asian households in their census tracts relative to Cluster 5.

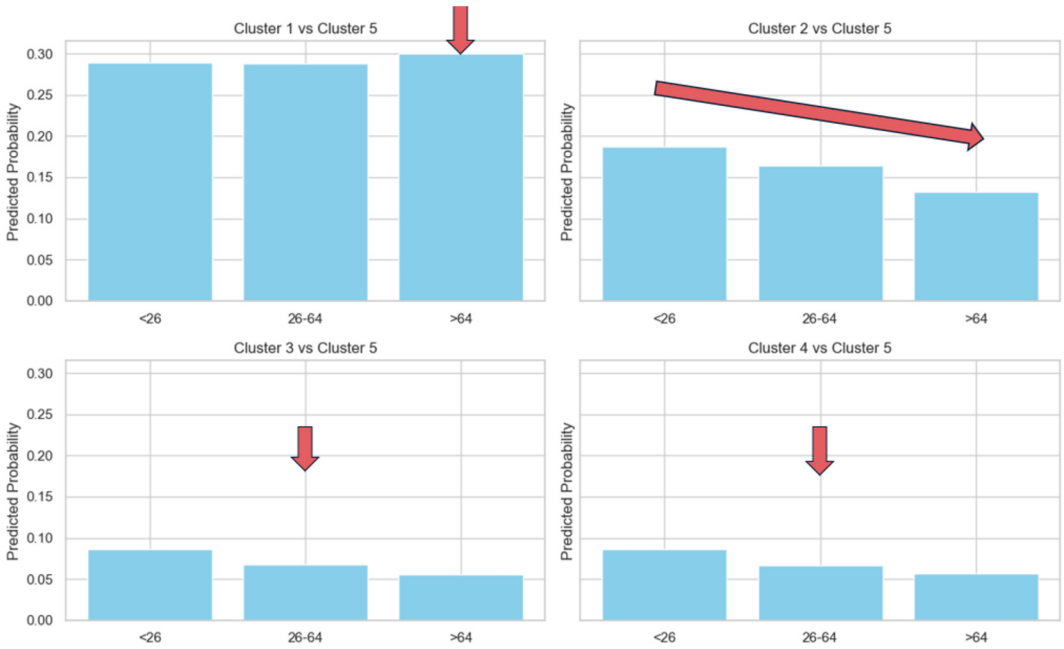

Figure 4: Predicted probabilities of cluster membership by age (reference: Cluster 5).

4.3 Predicted Probabilities

To aid interpretation, we examined predicted probabilities of cluster membership by key demographic and operational features, with Figures 4–7 illustrating marginal effects for age, race/ethnicity, initial contact mode, and reported participation likelihood. Younger households (<26) were more likely to fall into Clusters 1 and 2, while older households (>64) showed elevated probabilities of being in Cluster 1 but lower probabilities in Cluster 2. Cluster 1 also exhibited higher predicted probabilities for White households, whereas Cluster 3 showed a strong association with households reporting multiple racial identities. Mode-of-entry patterns further differentiated clusters: Cluster 1 households were more likely to be reached via SMS, Cluster 2 was more closely associated with MFOS entries, and Clusters 3 and 4 demonstrated lower responsiveness to SMS compared with Cluster 5. Finally, reported willingness to participate played a substantial role in distinguishing clusters—households that were "very likely" to participate were more often assigned to Cluster 1, while Clusters 3 and 4 maintained consistently low predicted probabilities regardless of stated participation.

4.4 Implications

The MLR findings underscore the joint role of demographic and operational features in shaping cooperation trajectories. In particular, language, race/ethnicity, and mode of contact differentiate less cooperative clusters from the highly cooperative Cluster 5. These insights provide actionable evidence for adaptive survey design: tailoring contact approach by language, adjusting contact mode sequencing, and allocating resources strategically to households predicted to fall outside the high-cooperation group.

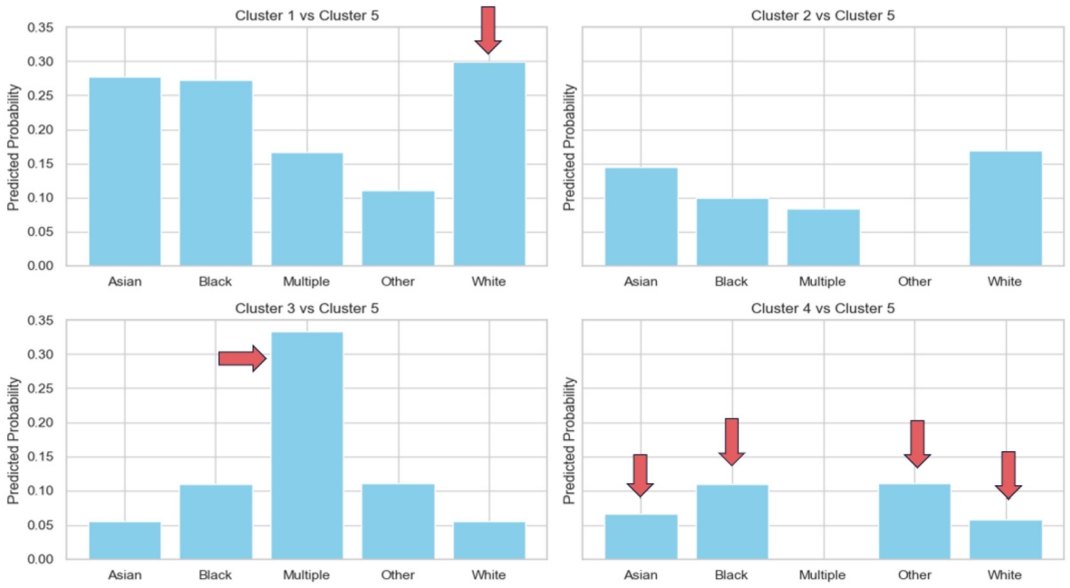

Figure 5: Predicted probabilities of cluster membership by census tract race/ethnicity (reference: Cluster 5).

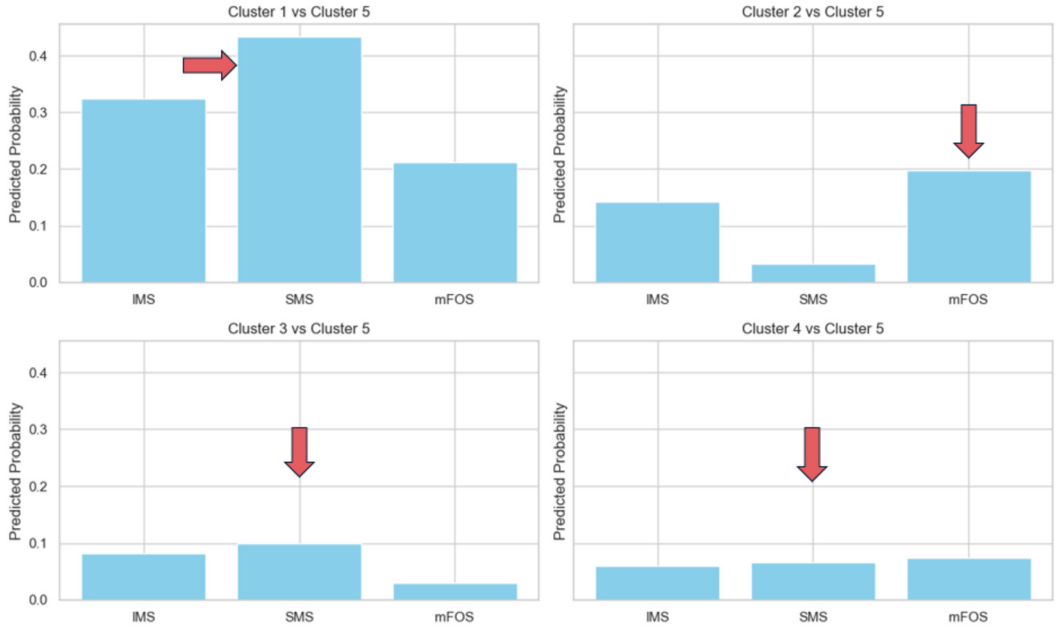

Figure 6: Predicted probabilities of cluster membership by initial contact mode (reference: Cluster 5).

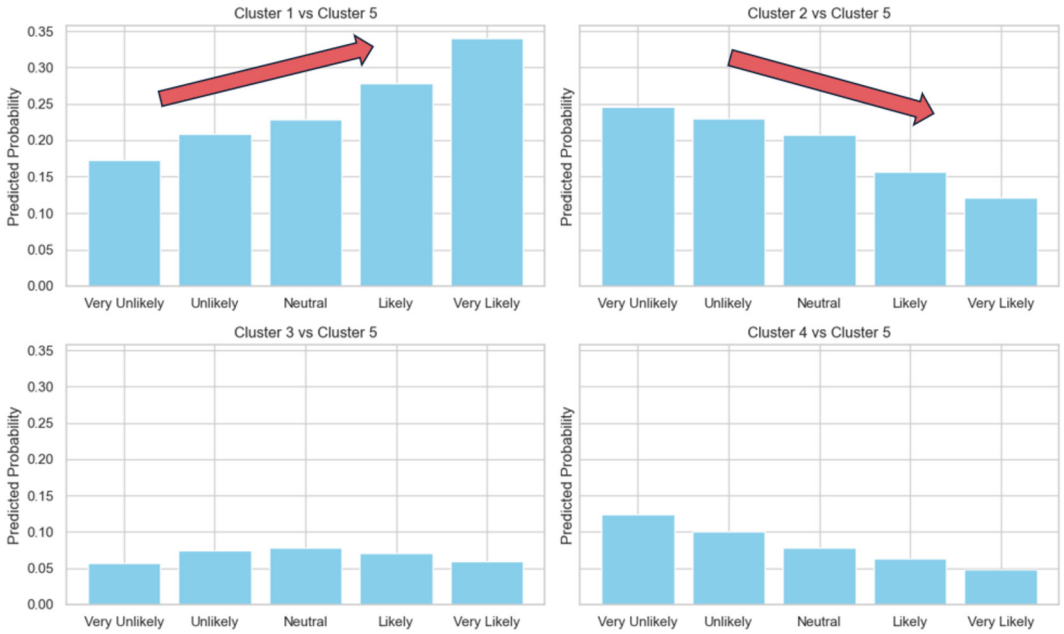

Figure 7: Predicted probabilities of cluster membership by stated participation likelihood (reference: Cluster 5).

5 Conclusion and Discussion

This study demonstrates the potential of integrating multiple data sources and machine learning methods to enhance operational decision-making in large-scale household surveys. Using the Medical Expenditure Panel Survey (MEPS) as a case study, we combined contact records, Advance Call Records (ACR) prior to the fielding of Round 1 of MEPS, and contextual Census tract-level characteristics from the American Community Survey (ACS) to model patterns of respondent cooperation. By analyzing contact sequences through clustering and predicting cluster membership using multinomial logistic regression, we provide a framework for understanding how census-tract demographic and operational factors interact to shape household response behavior.

Our results show that distinct patterns of cooperation emerge across households, reflecting differences in both respondent characteristics and interviewer effort. The highest-performing group (Cluster 5) achieved a 72% completion rate while requiring the fewest contact attempts, serving as an operational benchmark for effective outreach. In contrast, lower-performing clusters were associated with more contact attempts and specific census-tract demographic profiles, including higher proportions of Asian households or fewer bilingual respondents. Multinomial logistic regression results highlight that features such as interview language, race/ethnicity, and initial contact mode are significant predictors of cooperation trajectories. These findings confirm that contact records, when combined with auxiliary sources, can meaningfully predict differences in participation patterns.

Beyond prediction, the study underscores the practical value of such models for adaptive survey design. In particular, this work is intended to inform case prioritization strategies, where cluster-level insights can guide decisions around optimal contact modes and timing. Understanding which households are likely to follow certain contact trajectories can inform real-time

decisions on mode selection, outreach sequencing, and interviewer allocation. For example, households predicted to fall into low-cooperation clusters may benefit from targeted strategies—such as early bilingual contact or a switch to higher-engagement modes—while high-cooperation households could be managed more efficiently. In this sense, predictive modeling can support resource prioritization, helping balance fieldwork efficiency with representativeness.

We acknowledge that the operational utility demonstrated here remains prospective. A natural next step is to apply predicted cluster memberships in a future MEPS round — for example, by piloting modified contact protocols on a subset of households predicted to fall into low-cooperation clusters and comparing outcomes against a status quo control group. Such an implementation study would provide direct evidence of the model's field value. Like many paradata-based analyses, our work faces several limitations. The clustering and modeling are based on a single round (Round 1) of MEPS fieldwork, which may limit generalizability across time or survey contexts. The ACR and contact records sources, while rich, contain operational noise such as variations in interviewer practices or incomplete records for certain contact modes. In addition, while multinomial logistic regression offers interpretability, more flexible methods (e.g., gradient boosting or recurrent models) could capture higher-order interactions and temporal dependencies in future research. Finally, the analysis focuses on predicting cooperation patterns rather than downstream impacts on survey estimates or nonresponse bias; linking modeled cooperation patterns to data quality outcomes remains an important next step.

Future extensions of this work include applying the framework to later MEPS rounds and other longitudinal surveys to assess temporal stability in cooperation trajectories. Incorporating real-time predictions into field management systems could enable dynamic mode sequencing, where contact strategies adapt automatically based on predicted cluster membership. Developing interpretable dashboards that visualize predicted response likelihoods and contact patterns for field staff could further translate these models into operational tools. Ultimately, integrating predictive analytics with field operations has the potential to transform survey management—reducing costs, improving efficiency, and sustaining high-quality data collection in an era of declining response rates.

Supplementary Material

R and Python code for sequence feature construction, clustering analysis, and multinomial logistic regression modeling.

References

Bianchi A, Biffignandi S, Lynn P (2015). Web–face-to-face mixed-mode design in a longitudinal survey: Effects on participation, composition, and costs. *Survey Methodology*, 41(2): 301–318.

Chun AY, Schouten B, Wagner J (2017). JOS special issue on responsive and adaptive survey design: Looking back to see forward. *Journal of Official Statistics*, 33(3): 571–577. https://doi.org/10.1515/jos-2017-0027

Dillman DA, Smyth JD, Christian LM (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Wiley, Hoboken, NJ, 4th edition.

Durrant GB, Maslovskaya O, Smith PWF (2019). Investigating call record data using sequence analysis to inform adaptive survey designs. *International Journal of Social Research Methodology*, 22(1): 37–54. https://doi.org/10.1080/13645579.2018.1490981

Eckman S, Koch A (2019). Interviewer involvement in sample selection shapes the relationship between response rates and data quality. *Public Opinion Quarterly*, 83(2): 313–337. https://doi.org/10.1093/poq/nfz012

Groves RM, Peytcheva E (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2): 167–189. https://doi.org/10.1093/poq/nfn011

Kalton G, Flores-Cervantes I (2003). Weighting methods. *Journal of Official Statistics*, 19(2): 81–97.

Kreuter F, Kohler U (2009). Analyzing contact sequences in call record data: Potential and limitations of sequence indicators for nonresponse adjustments in the European Social Survey. *Journal of Official Statistics*, 25(2): 203–226.

Kreuter F, Olson K (2013). Paradata for nonresponse adjustment. *The Annals of the American Academy of Political and Social Science*, 645: 175–190.

Rybak A (2023). Survey mode and nonresponse bias: A meta-analysis based on the data from the International Social Survey Programme waves 1996–2018 and the European Social Survey rounds 1 to 9. *PLOS ONE*, 18(3): e0283092. https://doi.org/10.1371/journal.pone.0283092

Schouten B, Peytchev A, Wagner J (2017). *Adaptive Survey Design*. Chapman & Hall / CRC, 1st edition.

Schouten B, van den Brakel JA , Buelens B, Giesen D (2021). Adaptive mixed-mode survey designs. In: *Mixed-Mode Official Surveys*, 251–272. Taylor & Francis / CRC Press.

Wuyts C, Durrant GB, Kreuter F (2023). Predicting fieldwork effort in face-to-face surveys using call record data. *Journal of Survey Statistics and Methodology*, 11(2): 367–391. https://doi.org/10.1093/jssam/smab036