Journal of Official Statistics, Vol. 36, No. 3, 2020, pp. 489–505, http://dx.doi.org/10.2478/JOS-2020-0026

Continuing to Explore the Relation between Economic and Political Factors and Government Survey Refusal Rates: 1960–2015

Luke J. Larsen[1], *Joanna Fane Lineback*[1], *and Benjamin M. Reist*[2]

In the United States, government surveys' refusal rates have been increasing at an alarming rate, despite traditional measures for mitigating nonresponse. Given this phenomenon, now is a good time to revisit the work of Harris-Kojetin and Tucker (1999). In that study, the authors explored the relation between economic and political conditions on Current Population Survey (CPS) refusal rates over the period 1960–1988.

They found evidence that economic and political factors are associated with survey refusals and acknowledged the need to extend this work as more data became available. In this study, our aim was to continue their analysis. First, we replicated their findings. Next, we ran the assumed underlying model on an extended time-period (1960–2015). Last, since we found that the model was not an ideal fit for the extended period, we revised it using available time series and incorporating information about the CPS sample design. In the extended, refined model, presidential approval, census year, number of jobs and not-in-labor-force rate were all significant predictors of survey refusal.

Key words: Refusal rates; response rates; nonresponse; time series.

1. Introduction

Major government survey programs have many tools at their disposal for mitigating survey nonresponse. For example, they have access to high-quality sampling frames for correctly locating potential respondents and access to staff with expertise in converting nonrespondents. Additionally, they may be able to offer multiple reporting modes, offer monetary incentives, or extend data collection. However, recently in the United States, despite access to such tools, nonresponse rates – in particular refusal rates – have been dramatically increasing for unidentified reasons (see Subsection 1.1).

Here, we explore the recent increase by continuing the work of Harris-Kojetin and Tucker (1999), which focused on potential macro-level factors of survey refusal. In that study, the

[1] U.S. Census Bureau, 4600 Silver Hill Rd., Washington, DC, 20233, U.S.A. Emails: Luke.J.Larsen@census.gov and Joanna.Fane.Lineback@census.gov
[2] NASA Headquarters, 300 E St. SW Washington, DC, 20546, U.S.A. Email: Benjamin.Reist@nasa.gov
Acknowledgments: Thanks to Isaac Dorfman for his help analyzing the CPS Contact History Instrument data. Thanks to Kevin Younes for his help collecting relevant literature. Thanks to Timothy Gilbert for providing the historical National Crime Victimization Surveys refusal rates. Thanks to Shane Ball for proofreading the article. Thanks to the U.S. Department of Agriculture's National Agricultural Statistics Service and the U.S. Census Bureau for supporting this research. Finally, thanks to our reviewers for their helpful feedback. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official U.S. Department of Agriculture, U.S. Census Bureau, NASA or U.S. Government determination or policy.

authors used Current Population Survey (CPS) data and a time-series regression approach to examine economic and political influences (unemployment rate, presidential approval rating, inflation rate, consumer sentiment score, a census year indicator, and a March supplement indicator (see online Supplemental material, Appendix A (Data Sources)), for more information on these series) on CPS refusal rates over the period 1960–1988. The authors hypothesized that they would find evidence that environmental factors have an influence on the decision to participate in the CPS and suggested that a negative attitude about the government and a weak economy might decrease the likelihood of survey participation. However, they found that negative feelings about the government were associated with decreased survey participation, but that weak economic times were associated with increased survey participation.

Now, 20 years later, we replicate their findings and extend their model to the period 1960–2015. We also refine their model using available predictors and information about the CPS design. We expect to find the same relation between economic and political factors and survey refusal rates.

1.1. Increasing Refusal Rates in Major Government Surveys

Refusal rates in major government surveys in the United States have been increasing at an alarming rate, and they have been the main driver of increasing nonresponse rates. Increases in refusal rates are not unique to the United States; De Heer (1999) and De Leeuw and Luiten (2018) reported that refusal rates have been increasing since 1980 in many countries. To exemplify historical refusal rate patterns, below are plots of refusal rates over time for three such large-scale surveys: the CPS, the National Crime Victimization Survey (NCVS), and the National Health Interview Survey (NHIS). These surveys are conducted by the Census Bureau on behalf of the U.S. Bureau of Labor Statistics (BLS), the U.S. Bureau of Justice Statistics, and the National Center for Health Statistics, respectively. These surveys cover very different, but potentially sensitive subject matter: income, crime, and health, respectively. Each is primarily an in-person survey that has maintained a relatively stable design over an extended period, making the time series easy to interpret.

For interpreting the plots, it is important to understand the anatomy of these surveys' response rates. Households that were eligible for the survey but were not interviewed for some reason are referred to as Type A noninterviews. Each month, the Type A noninterview rate is calculated by dividing the total number of Type A households (refusals, temporarily absent, noncontacts, and other noninterviews) by the total number of eligible households, which includes Type A households and interviewed households. It follows that the refusal rate is the ratio of the total number of refusals to the total number of eligible households.

The CPS (U.S. Census Bureau and U.S. Bureau of Labor Statistics 2006) is the primary source of labor force statistics in the United States. Data are collected monthly in an electronic format by interviewers through in-person visits and telephone calls. As shown in Figure 1, the percentage of CPS refusals relative to the number of eligible sampled cases has been increasing over most of the 56-year period, 1960–2015. Around 1994, there was a sudden increase in refusals, as well as an increase in variability, that coincided with changes to data collection methods, including the introduction of computer-assisted in-person interviews (CAPI). Around 2010, the percentage of refusals began to increase

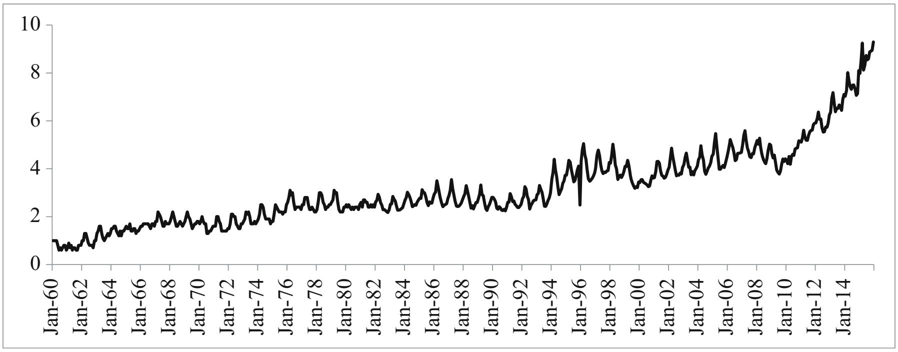

Fig. 1. CPS refusal rate by month: 1960–2015.
Source: U.S. Census Bureau, Current Population Survey, January 1960–December 2015 (unweighted).

sharply, doubling from around 4% in January 2010 to 8% in March 2014 with a high of 9.3% by the end of the study period in December 2015. As of this writing, CPS refusal rates have continued to increase, reaching a new high of 12.39% in February 2018.

Noncontacts are likely of interest to many readers, but the noncontact portion of the Type A rate was unavailable for much of the period being studied, so we can only comment on the non-refusal portion of the Type A rate. Harris-Kojetin and Tucker (1999) observed that, "Through the 1960s and 1970s, the non-refusal portion of the nonresponse rate (chiefly reflecting noncontacts) decreased at approximately the same rate as the refusal rate increased. However, when the refusal rate stabilized in the 1980s, the rate for other types of nonresponse did also." From 1989–2010, the ratio of CPS refusals to total Type A noninterviews hovered around 0.6, with the exception of the period 1999–2001 when refusals decreased as a percentage of Type A noninterviews (see Figure 2). Since 2010, refusals have increased as a percentage of total Type A noninterviews. Over the same period, the non-refusal portion of the Type A rate was increasing, but at a much slower rate.

The NCVS (NCVS 2017) is the primary source of crime victimization statistics in the United States. Data are collected in-person and by phone. The NCVS plot of the average yearly refusal rate since 1992 (shown in Figure 3) is strikingly similar to the CPS plot

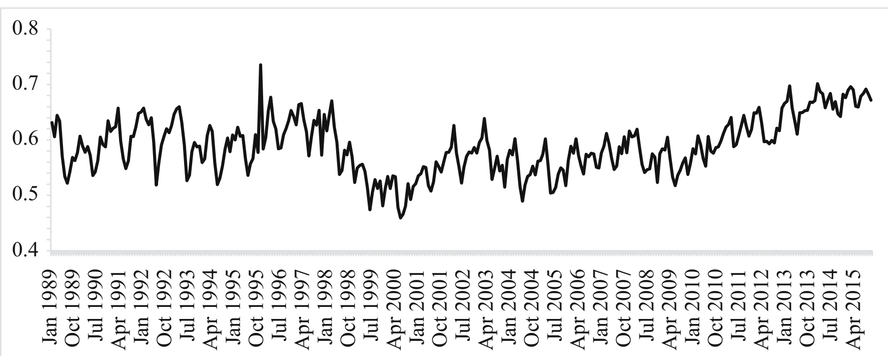

Fig. 2. Ratio of CPS refusals to total type A noninterviews (refusals, temporarily absent, noncontacts, and other noninterviews) by month: 1989–2015.
Source: U.S. Census Bureau, Current Population Survey, January 1989–December 2015 (unweighted).

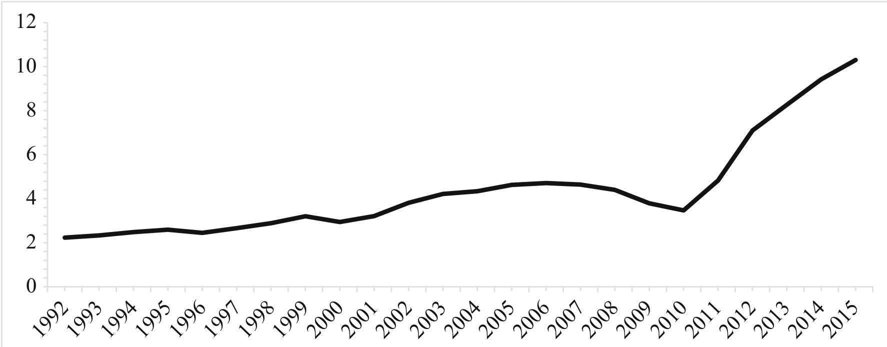

Fig. 3. NCVS average refusal rate by year: 1992–2015.
Source: U.S. Census Bureau, National Crime Victimization Survey, 1992–2015 (unweighted).

(Figure 1) over the same period. Specifically, refusal rates steadily increased until around 2010, when they began dramatically increasing.

The NHIS (National Center for Health Statistics 2016) is the main source of health statistics for the civilian, non-institutionalized population of the United States. Data are collected through in-person, household interviews. Like the CPS and the NCVS, NHIS refusal rates have seen an exponential increase in refusals over the period 2010–2015 with no signs of slowing down (see Figure 4), although NHIS refusal rates had already reached current CPS and NCVS levels around the time their refusal rates started their dramatic increase.

1.2. Changes in Data Collection Methods as a Possible Reason for Increased Refusal

From Figures 1, 3 and 4, one might wonder if the increase in refusals over time is due, at least in part, to changes in data collection methods. For instance, did working cases harder in the field lead to more contacts and ultimately more refusals?

While we do not have data before 2005 to help us answer this question, we do know that since the 1950s, the CPS has undergone regular questionnaire, sample design, estimation,

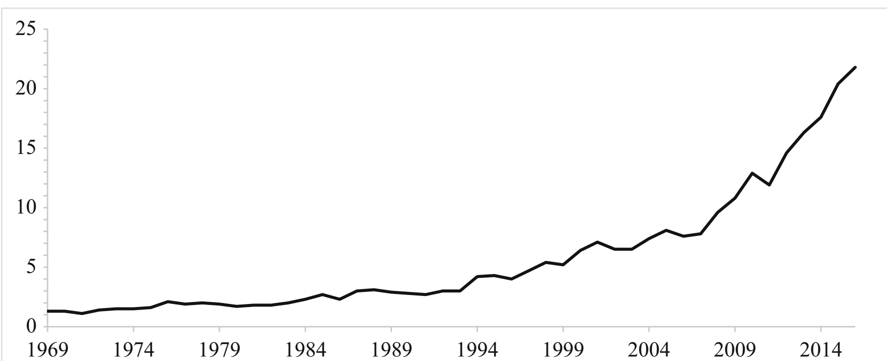

Fig. 4. NHIS refusal rate by year: 1969–2016.
Sources: 1969–2011, Gindi (2012); 2012–2016, National Center for Health Statistics, National Health Interview Survey.

and procedural data collection changes. As much as possible, these changes were planned to limit the amount of disruption to the survey. However, in many ways CPS data collection efforts have stayed the same over the years. It is still primarily an in-person survey, conducted monthly over a ten-day period, with cost-saving strategies built into field operations (such as limitations on number of contacts).

From the period 1960–1994, the Type A noninterview rate remained stable, although there were underlying changes in refusal and noncontact rates. Perhaps the most noteworthy procedural change happened in 1994, when the CPS began testing overlapping computer-assisted telephone interviewing (CATI) and CAPI. Around the same time, a new questionnaire and a sample redesign were introduced, and there was a noticeable increase in Type A noninterview rates. In late 1995 and early 1996, there was a disruption to data collection due to a government shutdown and another increase in Type A noninterview rates (U.S. Census Bureau and U.S. Bureau of Labor Statistics 2006). During 2011 and 2012, there was a restructuring of headquarters and field operations at the U.S. Census Bureau. Headquarters staff were realigned from survey-based to function-based units, with new survey directors managing each of the household surveys. Six of 12 regional offices were closed and the management structure of field operations changed. This was the first major restructuring of field operations in 50 years. Schafer (2014) found that there was no significant impact on response rates – at least for the NCVS – that could be attributed to the field restructuring.

Starting around 2005, the CPS, NCVS, NHIS, and other major government surveys conducted by the U.S. Census Bureau began collecting paradata that help analysts investigate whether changes in field efforts may have led to changes in refusal and noncontact rates. U.S. Census Bureau phone and field interviewers record information about contact attempts, including contact type, contact status, contact strategy, and reluctance-to-respond reason. From a cursory examination of these data for the CPS, which included an examination of the number of contacts and the distribution of reluctance reasons over time, there was also no evidence that data collection changes have contributed, at least since 2005, to changes in Type A noninterview rates.

1.3. Theoretical Background

Without evidence that data collection changes or any major event was the catalyst for the recent increase in refusal rates, we turn to the work of Harris-Kojetin and Tucker (1999) on large-scale factors of survey refusal. The authors ground their work in the research on an individual's decision to participate in a survey. They point out that an individual "can have well-founded rationales for not cooperating with a survey request that may be based on costs and benefits of responding." This is consistent with many of the theoretical frameworks for understanding the response process: for example, social exchange theory (Dillman et al. 2014), benefit cost theory (Singer 2011) and leverage-saliency (Groves et al. 2000). Harris-Kojetin and Tucker (1999) reason that the social, political, and economic environment may influence an individual's estimation of the costs and benefits of responding to a survey.

Based on this theoretical perspective, Harris-Kojetin and Tucker (1999) hypothesized that an individual's decision to respond to a government survey is, at least in part, related to his or her attitude about the government. They go on to propose that an individual's general feelings about the government can be captured by their approval of the chief executive. In addition, they surmised that, since the government seeks to manage the economy, measures of the nation's economic health may also reflect an individual's feelings about the government, further influencing one's decision to participate in a government survey.

Harris-Kojetin and Tucker's hypothesis was only partially supported by their model. The model showed that presidential approval has an inverse relation with refusal rates. On the other hand, the model showed that economic strength has a direct relation with refusal rates. In other words, the refusal rate decreased during weak economic periods and increased during strong economic periods.

We propose two alternative hypotheses for the relation between refusal rates and health of the economy. Note first that during weak economic times, more individuals in the United States rely on social programs (e.g., food stamps, unemployment insurance, Medicaid). A majority of these programs are facilitated – if not directly administered – by the federal government. We suggest that this increased interaction with the government increases an individual's estimation of the value of the government, and thus the benefit of responding to a government survey. However, neither we nor Harris-Kojetin and Tucker attempt to measure refusal at the individual level. Instead, as this is not practical, we used aggregate measures that reflect the broader survey climate. On the whole, the climate and contributing factors will influence some people's decisions more than others.

Our second hypothesis is that since the CPS is primarily a labor survey, the decision to participate is more salient during an economic downturn. The awareness and benefit of the unemployment rate maybe more apparent to the general individual because of the increased media coverage of the jobs report and the increased emphasis on the unemployment rate by politicians and policy makers. If this hypothesis is true, it would suggest that the relation between economic health and refusal rates might not generalize other types of surveys, such as health or crime surveys. However, pursuit of these hypotheses is outside the scope of this particular work, so we leave the topic to future research.

2. Methodology

This section details the three stages of this project, which included replicating Harris-Kojetin and Tucker's (1999) original findings, extending their model through 2015, and refining the model using additional data sources and information about the CPS sample design. In addition to outlining the methodology, this section discusses the rationale behind the use of alternative time series.

2.1. Replicating Original Findings

Harris-Kojetin and Tucker (1999) fit monthly CPS refusal rates (from January 1960 to December 1988) to a time series regression model using a select set of monthly time series data as regressors. (Hereon, we refer to this as the H-KT model.) In time series regression, the error term of the model is assumed to be decomposable into autocorrelated error that can be modeled with (1) autoregressive moving average (ARMA) model terms and (2) uncorrelated error that is assumed to be normally distributed with mean 0 and variance σ^2 (Ostrom 1978). The aim is that, once the autocorrelated error in the model has been controlled for, the response variable – in this case, the CPS refusal rate – can be fitted with predictor variables via typical multivariate linear regression techniques.

Most of the regressor series – CPS refusal rate, U.S. presidential approval rate, U.S. unemployment rate, and the Index of Consumer Sentiment – were differenced at both the first-order and seasonal-first-order, while the others – annual percent change of the 1982-basis consumer price index for urban consumers (CPI-U) inflation index and indicators for decennial census year and March CPS supplement month – were not differenced. (Throughout this article, we refer to the dual operations of first-order differencing followed by seasonal-first-order differencing as twice-differencing.) Harris-Kojetin and Tucker published the resulting model's coefficient estimates and the corresponding statistical significance for each model regressor, but the model's autocorrelated error structure, which takes the form of a seasonal autoregressive integrated moving average (SARIMA) model, $(p, d, q) \times (P, D, Q)_{12}$, was not identified.

In the first stage of data analysis, we attempted to replicate the original results using the same data sources over the same period. The exclusion of the autocorrelated error structure from the original article made it challenging to replicate the results of the original study exactly. Our work-around was to employ a brute-force technique to systematically fit the data under a wide variety of assumptions about the true error structure. For instance, the regression was first attempted with assumed error structure $(1,1,0) \times (0,1,0)_{12}$, then again under $(2,1,0) \times (0,1,0)_{12}$, and so forth. All data preparation and model fitting activities were conducted in R, with the SARIMA() wrapper handling the time series regression and residual diagnostics for over 200 variations of seasonal ARMA assumptions. The SARIMA() wrapper is part of the *astsa* package, which was produced for use with the textbook, *Time Series Analysis and Its Applications* (Shumway and Stoffer 2011). When adjusting the parameters of the autocorrelated error, the difference and seasonal difference parameters (d and D, respectively) were both fixed at 1, such that the first-order and seasonal-first-order differences were always in effect. However, the autoregressive (AR), moving average (MA), seasonal AR, and seasonal MA parameters (p, q, P, and Q, respectively) were allowed to vary between 0 and 3 to produce models under the various error structure assumptions.

Models in which the residuals were unstable or had significant autocorrelations were discarded from consideration; among those that remained, fit statistics – in particular, the corrected Akaike information criterion (AICc) – were used to assess good model fit, while the coefficient estimates were compared for accuracy against the "gold standard" coefficient estimates that were published in the original paper. Ultimately, one model was determined to have the best fit, while optimally minimizing the differences between the coefficient estimates and the coefficients of the original model.

2.2. *Extending the H-KT Model Through 2015*

In the second stage of data analysis, we applied the final model selected from the replication effort, including the finalized autocorrelation error structure, to an extended timeframe, January 1960 to December 2015. This particular model did not fit the new series of refusal rates very well, so we applied the brute-force procedure described previously to the extended timeframe to see if a different error structure might be more appropriate. This exercise resulted in a model with minimal differences in coefficient estimates and a different error structure that yielded acceptably uncorrelated residuals. However, the overall model fit was not as substantial over the full 55-year period, relative to the model fit of the original 28-year period. This outcome appeared to indicate that the original model design might not be appropriate for the more recent timeframe of 1988 to 2015.

2.3. Reevaluating the H-KT Model Construction

In the third stage of data analysis, we attempted to refine the H-KT model in order to obtain a better fit than that afforded by the second stage of analysis. To start, we vetted data sources that were available now that may not have been available during the original study and considered changes to the structure of the original model. We only considered including series that were comparable and available across time. Unfortunately, it was difficult to find a monthly or even quarterly time series in the social or political realms for the entire timeframe, so we ultimately focused only on new economic predictors.

In the end, the following additional regressor candidates were considered for inclusion in the model: U.S. quarterly gross domestic product (GDP), U.S. not-in-labor-force rate, number of U.S. jobs, raw CPI-Uinflation index, Standard and Poor's (S&P) 500 index end-of-month value, and party composition in the U.S. Congress. (For more information about each series, refer to online Supplemental material, Appendix A. The level of S&P 500 was chosen as a measure of wealth effect on refusal rates. The GDP was chosen as a measure of the general health of the economy. The number of jobs added was chosen as an alternative to the unemployment rate. The not-in-labor-force rate was chosen as a proxy for discouraged workers, as well as a measure of saliency of the labor force survey, since a labor force survey may not be salient to people not in the labor force. The series that was considered, but ultimately not included in the model, is congressional makeup (the ratio of Republicans to Democrats in each chamber of Congress), as congressional makeup stays relatively constant across consecutive months, which does not lend itself to the twice-differencing technique used in this analysis.

The final set of new and old regressor candidates were considered together for model inclusion on the basis of their pairwise correlations, excluding some to minimize multicollinearity concerns. All the original regressors were log-transformed prior to differencing, and the response series (CPS refusal rate) received a small constant addition prior to the log transformation in order to resolve some stationarity issues in the 1960s. The brute-force, trial-by-error approach to selecting an autocorrelated error structure was eschewed in favor of a more mindful strategy that incorporated information about the CPS sample design.

The monthly CPS sample design follows a rotating-panel structure, in which participating housing units are in sample for four consecutive months, then leave the sample for the next eight months, and then return to the sample for the following four months before leaving the CPS completely. For any given survey month, the CPS sample is comprised of members from each of eight different panels (U.S. Census Bureau and U.S. Bureau of Labor Statistics 2006). The rotating-panel structure results in significant correlation among estimates derived from monthly CPS files that are specific lags apart due to the sharing of some participating households between the two files. Any two consecutive monthly CPS files share about 75 percent of their samples by design, whereas any two CPS files that are a year apart (such as January 2000 and January 2001) share about 50% of their samples. From this structure, one can demonstrate that most CPS-based time series have significant correlation by as many as 15 lags apart. Keeping in mind that the intended response series to be modeled is actually a twice-differenced version of the monthly CPS refusal rates, the autocorrelation structure should be more complex in order to properly account for the presence of shared households between any two-point estimates of the twice-differenced series. Given the lag with significant correlation among the twice-differenced series can be as wide as 28 months, we chose a SARIMA model of $(28,1,0) \times (0,1,0)_{12}$ as the basis for the autocorrelated error of the time series regression.

3. Results

This section details the findings for each stage of this project: replicating Harris-Kojetin and Tucker's (1999) original findings, extending their model through 2015, and refining this model using alternative data sources and information about the CPS sample design.

3.1. *Replicating Original Findings*

After obtaining the CPS refusal rates series and the four regressor series used in the time series regression model showcased in Harris-Kojetin and Tucker (1999), some additional data preparation had to be applied prior to running the brute-force SARIMA modeling routine in R. Recall that the CPS refusal rates, presidential approval rates, unemployment rates, and consumer sentiment indices were all twice-differenced prior to modeling. In SARIMA(), all series in the model received the same differencing treatment, including those that we did not intend to difference (inflation rate, Census year indicator, and March supplement indicator). To counter this action, we applied "reverse twice-differencing" to the inflation rate series and two indicators using the diffinv() function in R. This allowed us to create a set of time series that could be twice-differenced within SARIMA() to get back to the original time series, while also applying the same differencing to the desired regressors and CPS refusal rates.

During this first stage of analysis (the primary work of which was completed between September 2015 and May 2016), we were aware that the "4-8-4" sample design employed by the CPS would be a key feature in determining which SARIMA parameters should be considered for the brute-force routine. It was known that the CPS in a given month shares a portion of its sampled housing units with other CPS sample months by up to 3 lags (months) forward and backward within a 12-month season and by up to 3 lags forward and backward after the first seasonal lag. Therefore, we decided that the brute-force routine would cycle through first-order lag parameters from 0 to 3 and seasonal lag parameters from 0 to 3. It was unknown whether the original model was strictly AR, MA, or some combination of the two types of ARMA terms, so we allowed the modeling routine to cycle both the AR- and MA-type error parameters (as well as their seasonal counterparts). From this design, the routine produced $4^4 = 256$ regression models to assess.

Each time series regression model produced by SARIMA() yielded model convergence status, coefficient estimates and corresponding standard errors, model fit statistics, and residual diagnostics charts. The residual diagnostics included the standardized residual

Table 1. Autocorrelation selection process.

Rank	Criterion	Condition	Action
1	Model convergence	Model does not converge	Eliminate model from consideration
2	Residual diagnostics	Residuals indicate autocorrelation or non-normality	Favor models with no or few residual issues
3	Coefficient estimates	Coefficient estimates not close to H-KT coefficient estimates	Favor models with estimates close to H-KT results
4	Model fit statistics	Lower AICc scores indicate better fit	Favor models with lowest AICc scores

plot over time, residual ACF plot, normal Q-Q plot of standardized residuals, and p-value plot of the Ljung-Box statistic over lags (in months). These charts were used to assess whether the standardized residuals are generally heteroscedastic, approximately normal, and contain little-to-no autocorrelation. Table 1 ranks the importance of each criterion in determining which model provided the best fit. Under this ranking, it is clear that convergence and residual behavior are critical elements in determining favorable model structures, while coefficient estimates and model fit statistics are important only when the critical elements are satisfactory. Of all 256 models under consideration, only 94 were viable choices in that they satisfied the critical convergence and residual behavior criteria. From these 94 options, the remaining criteria were assessed to make a selection of the autocorrelated error structure that yielded the best model fit.

Table 2 shows the original coefficient estimates (Harris-Kojetin and Tucker 1999) alongside those we obtained after determining the best-fit autocorrelated error structure: $(3,1,1) \times (2,1,1)_{12}$. One can observe that the point estimates were strikingly close to those of the original model, and yet two of the factors had differences in terms of statistical significance: Presidential approval rating was no longer significant under the replication effort, while the March supplement indicator was now significant. Nevertheless, the model based on $(3,1,1) \times (2,1,1)_{12}$ was convergent, featured acceptable residual behavior, and had a low AICc score of -524.90; therefore, we determined that this model structure

Table 2. H-KT versus replication model results.

	H-KT model: error structure unknown		Replication model: $(3,1,1) \times (2,1,1)_{12}$	
Predictor	Estimate	Std. Error	Estimate	Std. Error
Presidential approval (D)	-0.0026^{**}	0.0011	-0.0013	0.0012
Inflation rate	0.0000	0.0000	-0.0004	0.0002
Unemployment rate (D)	-0.0590^{**}	0.0180	-0.0540^{**}	0.0201
Consumer sentiment (D)	0.0042^{**}	0.0016	0.0043^{*}	0.0020
Decennial year	0.0084	0.0047	0.0095	0.0196
March supplement	0.0120	0.0073	0.0112^{*}	0.0046

Sources 1960–1988: Harris-Kojetin and Tucker (1999); U.S. Census Bureau; U.S. Bureau of Labor Statistics; University of Michigan; Gallup. All series are based on data from January 1960 to December 1988. (D) indicates a differenced time series. $N = 348$. $^{*}p < 0.05$, $^{**}p < 0.01$.

satisfactorily replicated the efforts of Harris-Kojetin and Tucker (1999) to fit the 1960–1988 monthly CPS refusal rate series to the featured set of predictors.

3.2. Extending the H-KT Model Through 2015

Next, we investigated whether the replication model identified in the previous section could adequately fit the refusal rates when the series was expanded to include monthly data up to December 2015. These expanded series contained several clear features, such as spikes in presidential approval following the 9/11 attacks and the 2008 presidential election, a sharp increase in unemployment rates during the Great Recession, and the relatively flat annual inflation rate since the 1990s. However, the CPS refusal rates after 1988 were particularly notable for a sustained growth trend until 2010, when the refusal rate series began a sharp increase, approaching 10% by the end of 2015. (Refer to online Supplemental material, Appendix B, for the plots of the expanded series for presidential approval, the inflation rate, the unemployment rate, and the index of consumer sentiment. Refer to Table 1 for the expanded CPS series.)

Because the trends in the original 1960–1988 window appear to be different from the trends in the 1989–2015 window, we compared pairwise correlations between the variables of interest for the entire 1960–2015 span with the correlations of the original span (Online Supplemental material, Appendix C). Most of the correlations were similar between the two efforts in terms of approximate magnitude, direction, and statistical significance. However, there were notable differences as well. For instance, the correlation between inflation and CPS refusal rate shifted direction significantly from 0.46 to −0.30, while the correlation between presidential approval and consumer sentiment shifted direction significantly from −0.60 to 0.48.

We ran the "best fit" model from Subsection 3.1 on the data from 1960–2015. Residual analysis did not indicate any problems, but the statistical significance of model coefficients for all factors except unemployment rate had shifted when comparing the 1960–1988 period to the 1960–2015 period (see Table 3).

Table 3. "Best fit" model parameters: 1960–1988 versus 1960–2015.

	1960 – 1988 $(3,1,1) \times (2,1,1)_{12}$		1960 – 2015 $(3,1,1) \times (2,1,1)_{12}$	
Predictor	Estimate	Std. error	Estimate	Std. error
Presidential approval (D)	−0.0013	0.0012	−0.0024*	0.0012
Inflation rate	−0.0004	0.0002	−0.0007**	0.0002
Unemployment rate (D)	−0.0540**	0.0201	−0.0714**	0.0222
Consumer sentiment (D)	0.0043*	0.0020	0.0033	0.0020
Decennial year	0.0095	0.0196	0.0350**	0.0082
March supplement	0.0112*	0.0046	0.0051	0.0038

Sources 1960–1988: U.S. Census Bureau; U.S. Bureau of Labor Statistics; University of Michigan; Gallup. All series are based on data from January 1960 to December 1988. (D) indicates a differenced time series. $N = 348$. $*p < 0.05$, $**p < 0.01$.
Sources 1960–2015: U.S. Census Bureau; U.S. Bureau of Labor Statistics; University of Michigan; Gallup. All series are based on data from January 1960 to December 2015. (D) indicates a differenced time series. $N = 672$. $*p < 0.05$, $**p < 0.01$.

Table 4. "Best fit" model parameters: 1960–1988 versus 1989–2015.

Predictor	1960 – 1988 $(3,1,1) \times (2,1,1)_{12}$		1989 – 2015 $(2,1,2) \times (0,1,1)_{12}$	
	Estimate	Std. error	Estimate	Std. error
Presidential approval (D)	−0.0013	0.0012	−0.0040	0.0024
Inflation rate	−0.0004	0.0002	−0.0015	0.0008
Unemployment rate (D)	−0.0540**	0.0201	−0.0768	0.0477
Consumer sentiment (D)	0.0043*	0.0020	0.0053	0.0039
Decennial year	0.0095	0.0196	0.0532**	0.0157
March supplement	0.0112*	0.0046	−0.0054	0.0084

Source 1960–1988: U.S. Census Bureau; U.S. Bureau of Labor Statistics; University of Michigan; Gallup. All series are based on data from January 1960 to December 1988. (D) indicates a differenced time series. $N = 348$. $*p < 0.05$, $**p < 0.01$.
Sources 1989–2015: U.S. Census Bureau; U.S. Bureau of Labor Statistics; University of Michigan; Gallup. All series are based on data from January 1989 to December 2015. (D) indicates a differenced time series. $N = 672$. $*p < 0.05$, $**p < 0.01$.

Analysis of the AICc between the two regression attempts yielded an interesting comparison: Under the 1960–1988 period, the AICc was −524.90, while under the 1960–2015 period, the AICc was −692.88. Because the 1960–1988 and 1989–2015 periods contain roughly the same amount of monthly data points (29 years for the former, 27 years for the latter), one might expect that the magnitude of the AIC for the combined periods would be approximately double that of the original period. Yet, that was not the case, which led us to suspect that the $(3,1,1) \times (2,1,1)_{12}$ error structure may not yield the best model fit for the more recent period. Another run of pairwise correlations, this time exclusively upon the 1989–2015 period, further corroborates this notion (see online Supplemental material Table 2 of Appendix C). Compared with the correlations from the 1960–1988 period, nearly half of the 15 pairwise correlations differed from the earlier period in magnitude, direction, or statistical significance.

Next, we returned to the brute-force modeling strategy from Subsection 3.1 to determine whether a different error structure might yield improved model fit for the 1989–2015 period. In this effort, the "best fit" model (see Table 4) featured an error structure of $(3,1,1) \times (2,1,1)_{12}$ with AICc $= -187.98$; though the replication model did have a better AICc of −190.87, it failed the residual assessment and was ineligible for further consideration. This is still less than half the magnitude of the AICc for the 1960–1988 replication model (AICc $= -524.90$), so the pursuit of obtaining a better fit to the more recent CPS refusal rate data may have to venture beyond the core construction of this particular model. To drive this point further, note that all but one of the coefficient estimates for the "best" model in the 1989–2015 window were not statistically significant. With only the decennial census year indicator having a significant effect upon predicted refusal rates, this clearly is not a very useful outcome.

3.3. Reevaluating the H-KT Model Construction

After determining that the original model could be reconstructed to fit the 1960–1988 data, but subsequently finding that process did not yield a comparably good fit to the expanded 1960–2015 data, we attempted a number of modifications in pursuit of a better model fit. First, we applied a log-transformation to the CPS refusal rate series – as well as to each of the regressors in the model – prior to the twice-differencing step in an attempt to improve the stationarity of the series before fitting the regression model. Also, because the raw refusal rates in the 1960s were so low, we found that we could further improve stationarity about that time period by adding a small constant to the entire raw refusal rate series prior to the log transform.

Next, we prepared a handful of other economic variables to be candidates for regressors in the new model (detailed in online Supplemental material, Appendix A). As mentioned previously, the current number of U.S. jobs regressor from the Current Employment Statistics by BLS was introduced to replace the CPS-based unemployment rate series in the model, while the raw inflation index (CPI-U, 1982 basis) was used to replace the 12-month percent change in the same index. The three additional regressors – not-in-labor-force rate from CPS, quarterly U.S. GDP, and end-of-month closing price of the S&P 500 – serve to provide additional dimensions of U.S. economic health that may be relevant to potential CPS respondents in determining their willingness to participate in government surveys. As with the other regressors, these five variables were log-transformed and twice-differenced prior to their inclusion in the time series regression model.

Pairwise correlations among the CPS refusal rate and the expanded set of regressors for the 1960–2015 window were analyzed (see online Supplemental material, Table 3 of Appendix C). Notably, all the regressors were significantly correlated with the refusal rate, while many of the regressor pairs had strong correlations between them (aside from those involving the census year and March supplement indicators). This is in line with findings from the correlational analysis done for Subsections 3.1 and 3.2. One should keep in mind that a few of these correlations are exceptionally strong, indicating that there may be a risk of overspecification in the model.

Finally, we reevaluated the manner of selection for the autocorrelated error structure. Since we were no longer trying to replicate the H-KT model results by guessing the error

Table 5. New model using expanded set of regressors: 1960–2015.

Regressor	Coefficient estimate	Standard error
Presidential approval (LD)	−0.0171*	(0.0104)
Consumer sentiment (LD)	0.0322	(0.0302)
Decennial year	0.0054**	(0.0018)
March supplement	−0.0001	(0.0019)
Number of jobs (LD)	0.9510**	(0.3345)
Inflation (LD)	−0.1851	(0.2687)
Not in labor force (LD)	0.5029*	(0.2160)
U.S. GDP (LD)	0.1133	(0.2017)
S&P 500 (LD)	−0.0040	(0.0231)

Sources: U.S. Census Bureau; U.S. Bureau of Labor Statistics; U.S. Bureau of Economic Analysis; University of Michigan; Gallup; Standard and Poor's. All series are based on data from January 1960 to December 2015. Results shown are for log-differenced CPS refusal rates.
(LD) indicates a log-differenced series (first order and seasonal first order). $N = 672$. $*p < 0.10$, $**p < 0.01$.

Table 6. New model fit: 1960–1888 versus 1989–2015.

	Early era (1960–1988)		Recent era (1989–2015)	
	Replicated model	Refined model	Replicated model	Refined model
Model fit (AICc)	−524.90	−1364.72	−187.98	−1324.75
Significant regressors	Unemployment, consumer sentiment, march supplement	Jobs, NILF rate, consumer sentiment	decennial year	Jobs, U.S. GDP, decennial year

Sources: U.S. Census Bureau; U.S. Bureau of Labor Statistics; U.S. Bureau of Economic Analysis; University of Michigan; Gallup; Standard and Poor's. All series are based on data from January 1960 to December 2015. $N = 348$ for 1960–1988. $N = 322$ for 1989–2015.

structure used in that study, the "brute force" iterative method used in Subsections 3.1 and 3.2 was not appropriate. Instead, we applied information about the CPS sample design to make reasonable assumptions about the autocorrelation present in the CPS refusal rate series and subsequently, the transformed refusal rate series to be fit by the time series regression model. Ultimately, we determined that the error structure for this model should be $(28,1,0) \times (0,1,0)_{12}$. (See Subsection 2.3 for more details.)

With these changes to specifications in place, the refined model was convergent and yielded satisfactory residual diagnostics. Table 5 shows the coefficient estimates of the expanded regressor set under this new model. We expected to see more evidence that positive feelings towards politicians are associated with a decrease in refusal rates and positive feelings about the economy are associated with an increase in refusal rates. In fact, four of the series are statistically significant predictors of CPS refusal rates – presidential approval and census year indicator from the original set of predictors, and number of jobs and not-in-labor-force status from the new set of predictors. From these results, increases in the number of U.S. jobs and the share of the population that is not in the labor force were predictive of increases in refusal rates. Being in a decennial census year was also linked to higher CPS refusal rates. However, increases in presidential approval were predictive of lower refusal rates. Compared with the Subsection 3.2 results, there was not a notable change in the point estimates of the coefficients – differences in statistical significance aside – but a comparison of the model fit statistics was particularly interesting.

Under the final model decided upon in Subsection 3.2, we found that the AICc was about −693, but under the new model described here, the AICc was about −2799 – about four times greater in magnitude. Note that the log transformation is the most likely driver of this difference in AICc values, so that difference in of itself is not an indicator of improved model fit between this effort and that of the replicated model in Subsection 3.2. However, the reader may also recall that one of the problems with the replicated model was that the model fit from the "recent era" (1989–2015) was not as good as the model fit from the "early era" (1960–1988) – the relevant AICc statistics were −188 and −524, respectively (Table 6). To see how the newer model shown here might compare, we re-ran the model for the two shorter timeframes and found that the AICc statistics between the two were roughly the same: −1365 for the early years and

−1325 for the recent years. This finding indicates that the refined model fits the refusal rates series about equally well in either of the shorter timeframes – a substantial improvement over the previous effort.

4. Discussion

Harris-Kojetin and Tucker (1999) initially considered the effect of large-scale political and economic factors on survey refusal rates using CPS refusal rates and relevant predictors over the period 1960–1988. They proposed that negative feelings about politics and and a weak economy would be associated with an increase in refusal rates. They found that disapproval of the president was associated with an increase in refusal rates, but a weak economy was associated with a decrease in refusal rates. With the rapid increase in government surveys' refusal rates over the past decade, it seemed like the ideal time to replicate and extend the work by Harris-Kojetin and Tucker (1999).

First, we replicated the results from the original H-KT model using similar time series methods and the same set of predictors (unemployment rate, presidential approval rating, inflation rate, consumer sentiment score, census year indicator, and March supplement indicator). We also found that presidential approval and unemployment rate were both negatively associated with refusal, while consumer sentiment was positively associated with refusal.

Next, we extended this model to the period 1960–2015, but found that the model did not extend well to the period 1989–2015. After refitting the model, the statistical significance of all model factors except unemployment rate changed from the original to the longer period. Presidential approval, inflation, and decennial year were all significant factors in the extended model, while consumer sentiment and March supplement month were no longer significant. These results may, in part, reflect that the original model was developed for a much more stable period of refusal rates. However, given the poor model fit, we have little confidence in these results.

Last, we refined the model using a modified set of predictors (presidential approval rating, consumer sentiment score, census year indicator, March supplement indicator, number of jobs, inflation rate, not in labor force rate, GDP, and the S&P 500 index). We achieved increased model fit over the original model. Increases in presidential approval were associated with lower CPS refusal rates, while U.S. jobs, the percentage of the population not-in-labor-force, and decennial year were all associated with higher refusal rates. It might not be obvious that strong economic times would lead to increased refusal, but if one considers that somehow people may feel less connected to the government during a strong economy, then this is reasonable. This result makes even more sense in the context of a labor force survey, such as the CPS.

It is important to underscore that these results may not be generalizable. The focus of this study was a United States labor force survey. The results may not extend to other countries. Within the United States, the results may not generalize to non-government surveys, which have very different response rates, and they may not even generalize to other government surveys.

Along these lines, a logical next step would be to replicate this analysis for other government surveys. The methodology of surveys like the NHIS and NCVS has stayed relatively stable for many years, giving us additional time series to study. At the same time, we should continue to take a closer look at the theory on survey nonresponse and collect or otherwise obtain measures that will help us understand more about the social aspect of the social-political-economic construct that is missing from these analyses.

In sum, we explored the recent increase in government surveys' refusal rates by continuing the work of Harris-Kojetin and Tucker (1999), which focused on potential macro-level factors of survey refusal. We refined and extended their model, and showed that presidential approval, census year, number of jobs and not-in-labor-force rate were all significant predictors of CPS refusal. While this model does not explain the changes in refusal rates, it can be used as a tool for monitoring possible causes of survey refusal over time. And while the recent spike in refusal rates is alarming, the good news for surveys like the CPS, NHIS, and NCVS is that overall response rates are still high. Government surveys, at least in the United States, still see response rates that far surpass response rates of most non-government surveys.

5. References

De Heer, W. 1999. "International Response Trends: Results of an International Survey." *Journal of Official Statistics* 15(2): 129–142. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/international-response-trends-results-of-an-international-survey.pdf (accessed may 2020).

De Leeuw, E., J. Hox, and A. Luiten. 2018. "International Nonresponse Trends across Countries and Years: An analysis of 36 years of Labour Force Survey data." *Survey Insights: Methods from the Field*. DOI: http://doi.org/10.13094/SMIF-2018-00008.

Dillman, D.A., J.D. Smyth, and L.M. Christian. 2014. *Internet, Phone, Mail, and Mixed Mode Surveys: The tailored design method (4th ed.).* Hoboken, NJ, US: John Wiley & Sons Inc.

Gindi, R.M. 2012. "Responsive Design on the National Health Interview Survey: Opportunities and challenges." Paper presented at FCSM Research Conference, Washington, DC, US. December 4th 2012. Available at: https://slideplayer.com/slide/7836641/ (accessed August 2018).

Groves, R., E. Singer, and A. Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an illustration." *Public Opinion Quarterly* 64(3): 299–308. DOI: https://doi.org/10.1086/317990.

Harris-Kojetin, B. and C. Tucker. 1999. "Exploring the Relation of Economic and Political Conditions with Refusal Rates to a Government Survey." *Journal of Official Statistics* 15(2): 167–184. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/exploring-the-relation-of-economic-and-political-conditions-with-refusal-rates-to-a-government-survey.pdf (accessed May 2020).

National Center for Health Statistics. 2016. "National Health Interview Survey Description." Centers for Disease Control and Prevention, U.S. Department of Health and Human Services. Available at: https://nhis.ipums.org/nhis/resources/srvydesc2016.pdf.

NCVS, National Crime Victimization Survey. 2017. “National Crime Victimization Survey, 2016: Technical Documentation.” Bureau of Justice Statistics, U.S. Department of Justice. Available at: https://www.bjs.gov/content/pub/pdf/ncvstd16.pdf (accessed August 2018).

Ostrom, C.W. 1978. *Time Series Analysis: Regression techniques*. Thousand Oaks, CA, US: SAGE Publications Inc.

Schafer, J.L. 2014. “Modeling the Effect of Recent Field Interventions in the National Crime Victimization Survey.” *Research Report Series:* Statistics 2014-02, U.S. Census Bureau. Available at: https://www.census.gov/srd/papers/pdf/rrs2014-02.pdf (accessed March 2019).

Shumway, R.H. and D.S. Stoffer. 2011. *Time Series Analysis and Its Applications (3rd ed.)*. New York, NY, US: Springer.

Singer, E. 2011. “Toward a Benefit-Cost Theory of Survey Participation: Evidence, Further Tests, and Implications.” *Journal of Official Statistics*, 27(2): 379–392. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/-toward-a-benefit%20cost-theory-of-survey-participation-evidence-further-tests-and-implications.pdf (accessed June 2020).

U.S. Census Bureau and U.S. Bureau of Labor Statistics. 2006. “Design and Methodology: Current Population Survey.” *Technical Paper:* 66. Available at: https://www.census.gov/prod/2006pubs/tp-66.pdf (accessed March 2019).

Received August 2018
Revised May 2019
Accepted November 2019