The number of publicly reported surveys conducted using automated telephone interviewing techniques, or robo-polls, has increased dramatically over the last few years. Pressured to cut costs, many media organizations no longer commission traditional telephone surveys, sponsoring robo-polls instead. Often the results of different surveys are compared regardless of the mode by which they interviews were conducted. This paper conducts a time series analysis of the President's job approval question asked on 624 national surveys conducted with live operators, over the internet and using robo-polls. It concludes that while live operator surveys and internet surveys produce quite similar results, robo-polls produce a significantly higher estimate of the disapproval rate of the President and a significantly lower estimate for 'no opinion', attributing the difference in results to non-response bias resulting from low participation rates in robo-polls.

The same question asked of the same sample using different interviewing methods can generate different responses. A number of such 'mode effects' have been identified. Kiesler and Sproul (1986) compared self-administered computer surveys to paper mail surveys, finding more extreme, less socially desirable responses for computer interviews. Chang and Krosnick (2009) compared oral interviewing to self-administered computerized questionnaires, finding greater social desirability response bias and satisficing for telephone interviewing. Dillman et al. (2000), found that automated voice and live operator phone respondents gave more extreme responses on scaled questions than mail and Internet respondents did.

In each of these studies the authors used experimental designs comparing a number of different questions across modes. Our approach consists of a statistical analysis of a single question, President Obama's job approval rating, asked in 624 publicly released national surveys; these surveys used three different modes: live operator, robo-poll and the Internet. We use the

---

[*] **Institution:** Voter/Consumer Research

[†] **Institution:** University of Santa Catarina

President's approval rating as the dependent variable and dummy variables representing the three interview modes as the explanatory variable. Of the surveys in our dataset 374 were conducted with live operators, 72 were Internet surveys and 178 were robo-polls. Since our main interest is the impact of interview mode, we control for the effect of other variables. We included as controls dummy variables representing the type of sample [3 levels: general population, registered voter (RV) and likely voter (LV)], the funding source of the survey (5 levels: media funded, Republican funding, Democratic funding, university funding, other private funding) and variables representing sample size and the timing of the survey (the start date and field duration).

We use a time series regression model to control for changes in the President's approval ratings over time. The time variable is represented by the number of days elapsed since the date of the first survey in the dataset. The surveys included in our dataset span a period starting in November of 2008 and ending in April of 2010. Specifically we use time series regression for proportional-odds models in a logistic equation. In proportional odds regression analysis the dependent variable is represented as the logarithm of the ratio of two percentage responses. In practice, this transformation usually enhances the robustness and reliability of the statistical results (Agresti 200 2). Since the job approval question has 3 possible responses, Approve (A), Disapprove (D) and Undecided (U), this generates a total of 6 possible regression equations with log (A/U), log (U/A), log (D/U), etc. as dependent variables. Since A, D and U must sum to 1 only two regressions are needed to estimate the three percentages. We estimated all six equations to make sure that the results were consistent and found this to be the case. The typical regression equation (for the case of A/U for instance) is[1]

log(A/U)=f(Robo-Poll, Internet, Live Operator; Republican Funding, Democratic Funding, Media Funding, Private Funding, University Funded, Bipartisan Funding; Registered Voter, Likely Voter, General Public; Sample size, days in the field, time, $time^2$, $time^3$, $time^4$)+e

The time variable is a polynomial transformation of the time series trend up to the fourth degree.[2] Table 1 displays the parameter estimates for the regression equation estimating log (A/U).

---

[1] Theoretically, there is no essential difference between a regression with log(A/U) and the inverse log(U/A). This is true for all odds and their inverses. All results are available from samohyl@deps.ufsc.br. [1] The regression equation has no constant term nor is every variable represented; for instance, a variable flagging general public samples is not explicitly present in the regression equation to avoid the dummy variable trap. See Chatterjee et al. (2006), Chapter 5, on qualitative variables as predictors.

[2] $Y = a_o + b_1 t + b_2 t_2 + b_3 t_3 + b_4 t_4$

Table 1   Parameter estimates for Log (A/U).

	Estimate	Std. Error	t value	Pr(>|t|)
Robo-poll	1.703	0.483	3.523	0.000
Internet	0.851	0.473	1.799	0.072
Live Operator	0.319	0.456	0.698	0.485
Likely voter	0.562	0.101	5.578	0.000
Registered voter	−0.329	0.113	2.919	0.004
Time	0.009	0.002	4.548	0.000
Time^2	−0.00003	0.000008	3.146	0.002
Time^3	0.00000002	0.00000001	2.276	0.023
Democratic Funding	−0.370	0.435	0.850	0.396
Media Funding	0.713	0.427	1.670	0.095
Privately funded	0.197	0.433	0.456	0.648
Republican Funding	0.824	0.452	1.824	0.069
University funded	0.055	0.450	0.122	0.903
Sample size	0.0003	0.0001	4.347	0.000
Days in the field	0.043	0.020	2.161	0.031

Residual standard error: 0.582 on 609 DF

Multiple R-squared: 0.959, Adjusted R-squared: 0.958

F-statistic: 944.9 on 15 and 609 DF, p-value: =0.000

Table 2   Comparison of Respondent Ideology.

	Moderate/ Liberal	Somewhat conservative	Very conservative
Live operator	32%	25%	40%
Robo-poll	19%	27%	53%

Most t-statistics (the last column of Table 1) are significant and the adjusted R^2 is high. The variables representing mode effects had a significant impact, especially for robo-polls, and this finding will be analyzed in detail below. The variables representing the funding source are marginally significant but only in the two cases of media funding and Republican funding but not at all conclusive (p values of 0.095 and 0.069 fall outside the traditional value of 0.05). Martin, Traugott and Kennedy compare the results of election surveys to election outcomes, and report a systematic tilt by polling firms with a partisan affiliation (Martin, Traugott, and Kennedy 2005). Surprisingly, our results show that Republican pollsters tend to report somewhat higher approval ratings for President Obama. However, in light of the weak t statistics, these results demand further investigation. Sample type (likely vs. registered vs. general public) is also associated with different approval rates. This again is consistent with the finding that surveys that screen for registration and turnout generate different results, usually the more tightly screened the greater the Republican tilt. However, in this case the likely voter sample produced higher

approval ratings for the President. Finally, the larger the sample size and the longer the field period of the survey, the lower the President's approval rate.

We calculate changes in the estimated Presidential approval rate by first transforming $e^{\log(A/U)}=A/U$, followed by the following transformations

A=(A/U)/(1+A/U+D/U)U=1/(1+A/U+D/U)D=(D/U)/(1+A/U+D/U)

This allows us to plot the estimated trend in approval ratings after controlling for the effects of the independent variables:

President Obama's approval rating reached its highest value around May 5, 2009. The undecided rate dropped steadily over time, while the President's disapproval rate increased steadily, accelerating roughly in March of 2009.

The next three charts plot responses by interview mode. Figure 2 displays the trend in the approval rate differentiated by interview mode:
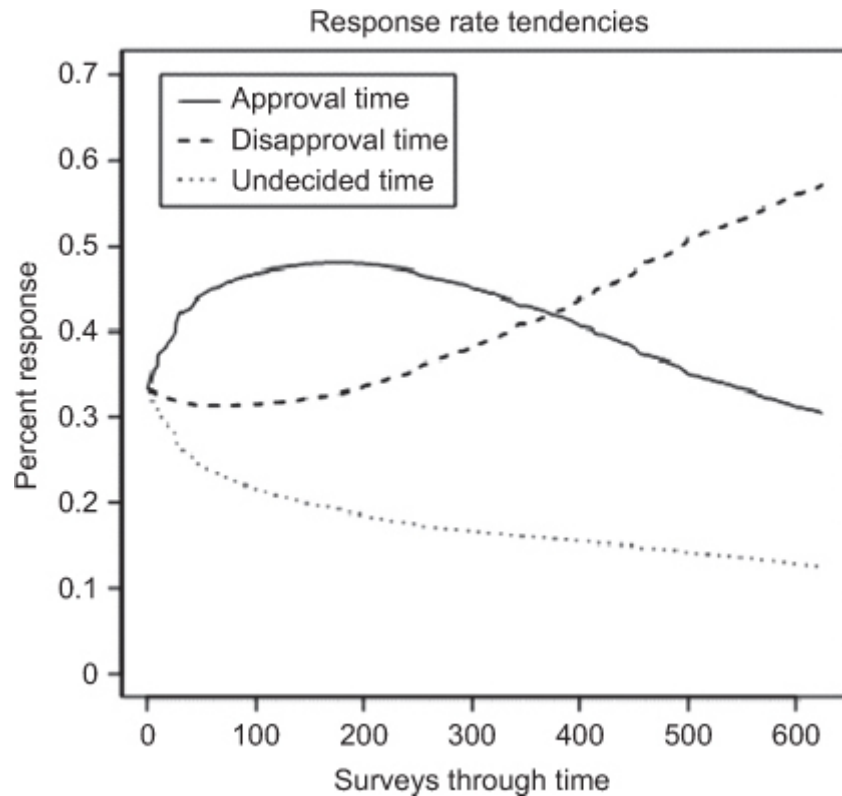

Figure 1 Approval, disapproval and indecision by day between November 2008 and April 2010, after controlling the effects of the independent variables.
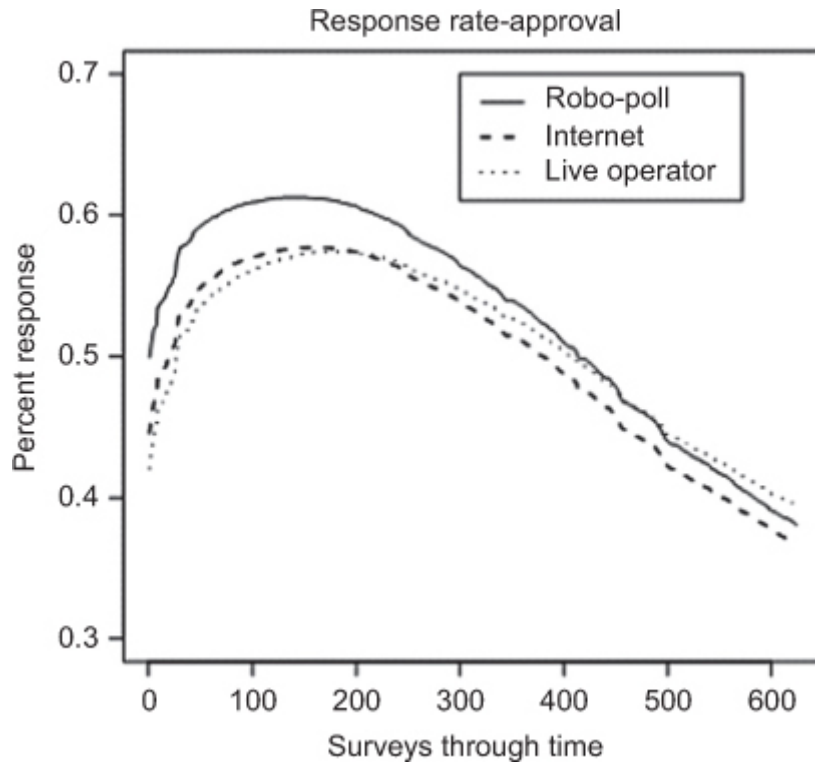
Figure 2 Approval rate by day between November 2008 and April 2010, by interview mode over time.

Figure 2 shows that the three trend-lines are practically indistinguishable, actually crossing at various points in time; in other words, interview mode has little or no effect on the President's approval rate. However, a very different picture emerges for the disapproval and undecided rates. Figure 3 shows that the undecided rate is much lower for robo-polls than for either live operator interviews or Internet interviews:
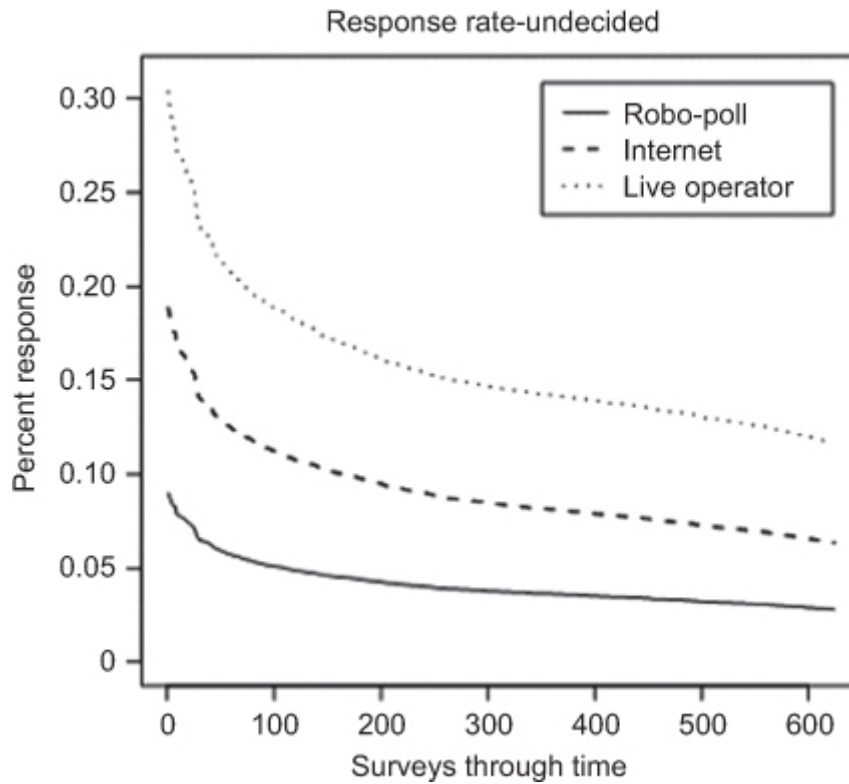
Figure 3   Undecided rate by interview mode day between November 2008 and April 2010.

Finally, Figure 4 shows that interviews done by live operators report consistently lower disapproval rates than robo-polls and Internet surveys. Also note that the differences are constant for the entire time period covered by the surveys:
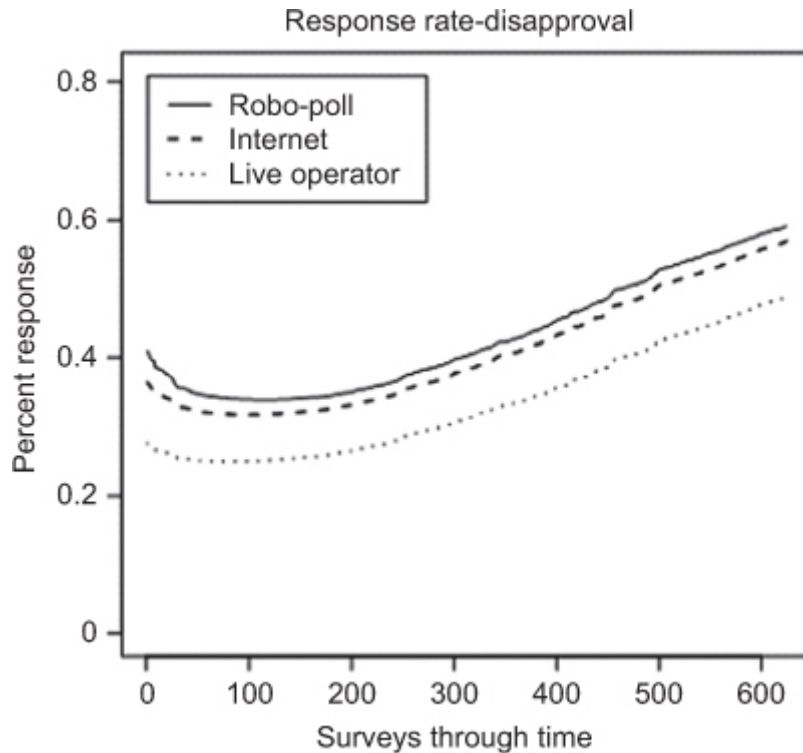
Figure 4   Disapproval rate by interview mode day between November 2008 and April 2010.

Since in any month dozens of surveys estimate the President's approval ratings, we suggest that news reports should routinely include the interview mode. Comparisons between surveys using different interview methods should be taken with a grain of salt; different disapproval rates may result from differences in interview mode rather than a change in public opinion. Comparing disapproval ratings among different Presidents is certainly unfair if data collection methods in constant evolution are not considered.

The literature has pointed to social desirability as a possible reason for mode effects. However, in this particular case that seems an insufficient explanation. A social desirability effect on the President's approval rating might be defined as follows: some respondents who are critical of the President's performance do not want to say so, instead replying that they are not sure. As more people disapprove of the President, the number of people unwilling to say so should increase and lead to a larger increase in the undecided rate for live operator surveys than for robo-polls. This clearly did not happen; in fact the difference shrunk over time (Figure 3).

An alternative explanation for mode effects might be found in what Groves (1989) defines as a non-response effect: different non-participation rates for robo-polls and live operator surveys might produce different audiences. People who stay on the phone for robo-polls might be different from people who stay on the phone for a live operator survey.

To assess this possibility we returned to the experimental design approach

used by others (Chang and Krosnick 2009; Dillman et al. 2001; Kiesler and Sproull 1986). We interviewed Republican primary voters in Delaware, and administered nearly the same questionnaire[3]: 400 respondents were interviewed using live operators, while 300 were interviewed with robo-polling. The interviews were conducted at the same time.

We obtained a higher response rate for the live operator sample then for the robo-poll. RR5 (as defined in AAPOR's standard definitions[4] for the live interviews was 23%, compared to 9% for the robo-poll. The two samples also produced very different responses to a key political variable: participants in the robo-poll were considerably more conservative[5]

Our experiment is incomplete in that we only interviewed Republican primary voters. Finding that Republicans who participate in robo-polls are more conservative is consistent with the finding that robo-polls produce higher disapproval rates for the President. Broader experimentation is needed, however. It may very well be that the type of person who is more likely to respond to a robo-poll has a stronger desire to be heard, is less likely to be undecided and is more likely to express an opinion. If that is the case, a parallel experiment conducted with Democratic primary voters might find that participants in robo-polls are more liberal than participants in live operator surveys.

---

3 Some of the questions asked in the live operator survey had more than 10 response codes, which robo-poll surveys, relying on a telephone keypad could not handle; the number of codes was reduced for the robo-poll survey. Also, open ended questions on the live operator survey were dropped for the robo-poll survey.

4 Standard Definitions, Final Dispositions of Case Codes and Outcome Rates for Surveys, revised 2009, the American Association for Public Opinion Research.

5 Question wording: On political issues do you consider yourself (ROTATE) a liberal, a moderate or a conservative? (IF CONSERVATIVE:) And would you say you are very conservative or just somewhat conservative?

REFERENCES

Agresti, A. 2002. *Categorical Data Analysis*. 2nd ed. New Jersey: Wiley.

Chang, L., and J.A. Krosnick. 2009. "National Surveys via Rdd Telephone Interviewing versus the Internet, Comparing Sample Representativeness and Response Quality." *Public Opinion Quarterly* 73 (IV).

Chatterjee, S., B. Price, and A.S. Hadi. 2006. *Regression Analysis by Example*. 4th ed. John Wiley & Sons.

Dillman, D.A., G. Phelps, R. Tortora, K. Swift, J. Kohrell, and J. Beck. 2001. "Response Rates and Measurement Differences in Mixed Mode Surveys Using Mail, Telephone, Interactive Voice Response and the Internet." Draft Paper.

Groves, R.M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley & Sons.

Kiesler, S., and L.S. Sproull. 1986. "Response Effects in the Electronic Survey." *Public Opinion Quarterly* 50 (III).

Martin, E.A., M.W. Traugott, and C. Kennedy. 2005. "A Review and Proposal for a New Measure of Poll Accuracy." *Public Opinion Quarterly* 69 (V).

R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.