IHP
INSTITUTE FOR HEALTH POLICY

Sri Lanka Opinion Tracker Survey

2024-16

FOR RELEASE NOVEMBER 05, 2024

Performance of the IHP SLOTS MRP model in polling the Presidential Election 2024, methodological issues, and improvements

BY *Ravindra P. Rannan-Eliya*

FOR MEDIA INQUIRIES
Secretary/Administrator
Email: info ‘at’ ihp.lk

RECOMMENDED CITATION

Ravindra P. Rannan-Eliya, October 2024, “Performance of the IHP SLOTS MRP model in polling the Presidential Election 2024, methodological issues, and improvements”, Institute for Health Policy: SLOTS Report #2024-16.

About the IHP Sri Lanka Opinion Tracker

The Institute for Health Policy (IHP) conducts the SLOTS survey to track changes in health and social conditions, and public opinion in the country, on behalf of the Sri Lanka Health and Ageing Study (SLHAS) consortium of Sri Lankan academic and research institutions. IHP is solely responsible for commissioning and designing the survey, and it takes full responsibility for it. IHP is an independent, non-partisan research institution based in Colombo, Sri Lanka. The SLOTS lead investigator is Dr Ravi Rannan-Eliya of IHP, who was trained in public opinion polling at Harvard University, and who has conducted numerous opinion surveys over three decades.

SLOTS interviews representative samples of Sri Lankan adults every day by telephone to gather their current views and situation. All interviews include a core set of common questions, with additional rotating sets of other questions that examine issues of topical importance. Interviews are done daily by phone by IHP employees, with respondents recruited by a national field survey or by randomly dialling mobile phone numbers. SLOTS fieldwork since 2021 has been supported by a range of funders, who play no role in question design, data analysis, or reporting. Funders have included the Neelan Tiruchelvam Trust, Asia Foundation, European Commission, UK National Institute for Health and Care Research, Foundation Open Society Institute, USAID, Velux Stiftung Foundation, New York University Abu Dhabi, and the IHP Public Interest Research Fund. The survey findings do not necessarily reflect the views or positions of past and present funders. Interested parties can contact IHP for more detailed data and results.

SLOTS respondents consist of a mix of respondents reached by random digit dialling of mobile numbers, and others coming from a national panel of respondents who have agreed to be re-interviewed, and who were previously recruited using random selection. As with any survey, bias can arise from the sampling design and non-response, which means that respondents are not representative of the underlying population. To adjust for this, unless otherwise noted, all reported estimates and analyses use data that have been weighted to ensure that they are representative of the national adult population. This weighting process uses propensity weighting and iterative proportional fitting (raking) to match the national population according to age, gender, ethnicity, religion, socioeconomic ranking, education, sector, and geographical location, and where appropriate by voting history. All survey results reflect the views of respondents only at the time the survey was fielded, as indicated in this report.

The survey has an omnibus design, and the Institute welcomes sponsorship to continue the survey, to add new questions, or to undertake tailored analyses of the data. Potential sponsors should contact the Institute for further details.

Performance of the IHP SLOTS MRP model in polling the Presidential Election 2024, methodological issues, and improvements

Summary

IHP released its final pre-election Presidential Election voting intent update on 18 September, three days prior to the 2024 Presidential Election. This final update was not based on the SLOTS MRP model that had been used in previous estimates, but on an improvised model prepared in haste following detection of a large social desirability effect leading to substantial over-reporting of past voting for AKD. This over-reporting led the original IHP model to increasingly underestimate AKD support.

The final released estimates for the major candidates correctly identified the winner, but they had an average error of 4.1%. This might be considered respectable for a first effort, but performance fell short of IHP expectations.

The primary reason for the underperformance was the practical difficulty in finding a rapid solution to the problem of a rapid increase in social desirability bias in survey responses during the previous three months. This report reports the accuracy of the final estimates, discusses issues that arose in the last few weeks of survey tracking, undertakes further analysis of alternative solutions, and develops methodological improvements. The analysis suggests that using modelled estimates of individual level past voting choice instead of self-reported past voting choice would have circumvented the issue of social desirability bias and produced estimates with a more satisfactory level of prediction error. This result may have wider relevance to the general research on modelling vote intent in conditions of significant reporting bias.

How well did we do?

Short answer: Not so badly in the circumstances, but not as well as we would have liked.

Our tracking of voting intent correctly identified the winner–AK Dissanayake–and that he would fail to win on the first count. Our internal estimates indicated also that he would win on a count of second preferences, with most people not casting a second preference and those that did breaking 2:1 in favour or Sajith Premadasa. We also correctly placed the three other major party candidates (Exhibit 1).

Exhibit 1: Comparison of MRP final estimates and actual election results

Candidate	Weighted responses 01/01–17/09/24	Final MRP estimate (18/09/24)	Election result (21/09/24)	Error
AK Dissanayake	43.6	47.9	42.3	5.6
Sajith Premadasa	33.5	25.0	32.8	-7.8
Ranil Wickremesinghe	16.1	20.0	17.3	2.7
Namal Rajapaksa	6.3	4.7	2.6	2.1
Others	0.5	2.4	5.1	-2.7
			RMSE =	4.2%

We overestimated AK Dissanayake's {AKD} vote by 6%, and underestimated Sajith Premadasa's by 8%. The Root Mean Squared Estimate (RMSE), a standard measure of polling accuracy that averages the errors, was 4.1%. This is almost double that in US state level Presidential Election polls, which are the most comparable, although polling a three or four-way race is harder than the two-way races that are the norm in the US.

Our previous tracking from 2023 through mid-2024 was notably closer to the actual result. The average difference between the average vote shares of candidates throughout 2024 and the final election result would have had a RMSE of just 2.3%. SLOTS data also showed for more than two years that the NPP was quite likely to win a Presidential Election.

I next discuss the discrepancy between our long-term tracking and the final estimates, and why we were forced to use an improvised model for the final estimates.

What went wrong in the IHP MRP model

IHP uses Multilevel Regression and Poststratification (MRP) to model voting intent based on responses to its daily SLOTS interviews. This differs from the usual methodology of typical public opinion sample surveys, which generate estimates by interviewing representative samples of the public. MRP remains relatively novel and has been used widely to date only in the UK. The method involves using survey data to statistically model the relationship between voter demographics and vote choice, and then using these statistical model estimates to project voting intent in a separate data file that is designed to represent the demographics and other characteristics of the overall electorate.

IHP chose MRP to overcome the challenge of tracking vote intent using relatively small monthly samples, which have averaged just 500–600 in the case of SLOTs. MRP conventionally has been used to estimate vote choice in small spatial units such as British parliamentary constituencies, where typical national surveys typically lack adequate sample size to provide reliable estimates. In IHP's case, we adapted it to estimate voting intent in small temporal units, *i.e.*, months or days. This adaptation of MRP to estimate differences in vote intent over time as opposed to over spatial units is relatively novel, although the underlying theory is the same.

A key variable used in IHP's MRP model through August 2024 was past voting history. The SLOTS survey asked all respondents how they voted in the 2019 and 2020 national elections. Globally, how people voted in previous elections is one of the best predictors of current vote intent, and this is also true in the Sri Lankan data. In doing this, a problem that IHP faced was that a significant number of SLOTS respondents (33%) did not provide their past voting history, and those who refused to answer differed substantially from those who did. To overcome this, their past voting history was imputed through a statistical process called multiple imputation, which generates multiple estimates of past vote choice that reflect the uncertainty in any imputations. This imputed past voting history was then used in combination with other variables, such as age, gender, and income level, to model vote intent.

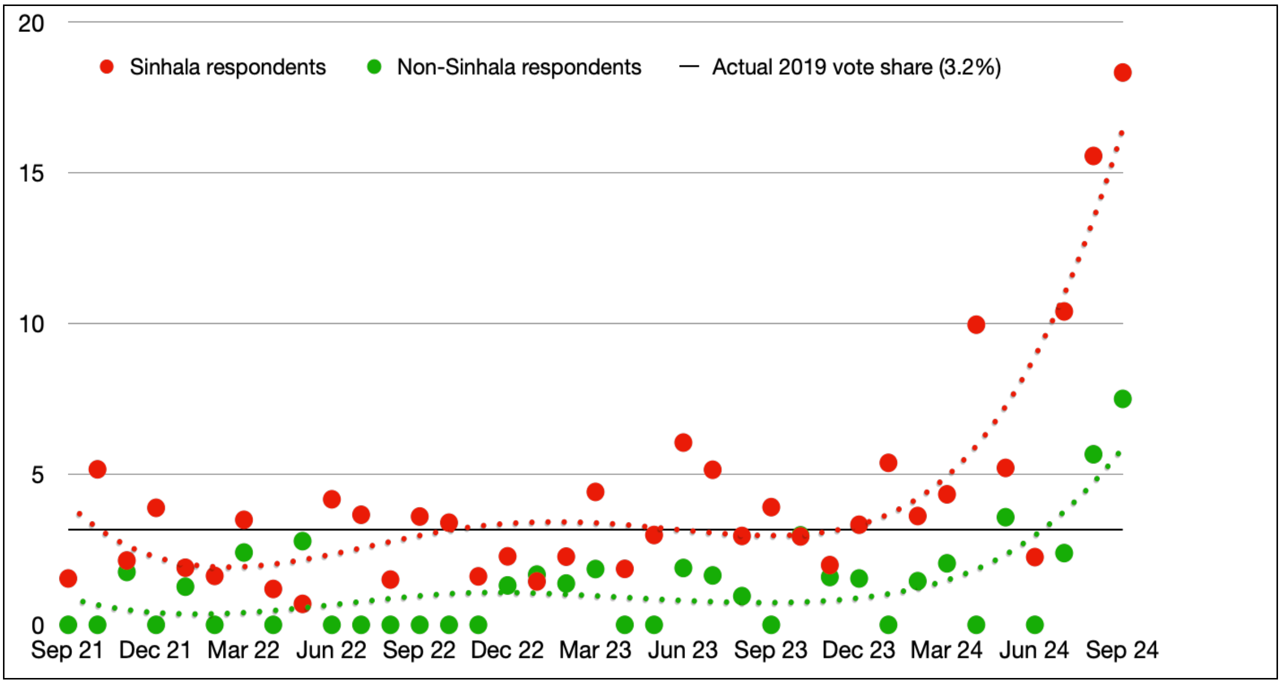
To our knowledge, this imputation of past voting history has not been used in other MRP analyses, and the problem of imputing past voting history remains a topic of ongoing research in the academic literature, so our approach could be considered novel. To implement it, we adapted a method from the epidemiological literature, which has been used to estimate national disease prevalence in a scenario where detection of the disease requires a test, and where there are non-random differences between individuals who do or do not participate in the test. This corresponded to the SLOTS context, where those refusing to divulge past voting history were different to those who did, with Gotabaya Rajapaksa voters much more likely to disclose this than other voters during 2021–2023. As implemented, our method assumed there was differential reporting by supporters of different candidates, but that this was due largely to differences in willingness to respond and not differences in truthfulness if they did respond. This assumption was supported by various data analyses through 2023.

This assumption failed just prior to the election, when the percentage of respondents responding that

that they had voted for AKD in 2019 increased from 2% (of all responses including refusals) during Jan. 2023–Jun. 2024 to 5%, 8% and 9% during the months of July to September 2024. As only 2.6% of the electorate voted for AKD in 2019, the immediate impact of this in the model was that these respondents, almost all of whom responded that their current vote intent was for AKD, were weighted down in the final estimates. This led the model to increasingly underestimate the AKD vote.

This change in responses was most likely due to social desirability bias. This involves respondents giving responses that they believe are socially desirable or will satisfy interviewers expectations. Although the IHP MRP model was designed to cope with some degree of social desirability bias, this extreme level of bias was outside its design parameters. Nevertheless, the emergence of this social desirability bias from early 2024 suggests that respondents were becoming more aware of overall public support for AKD (Exhibit 2).

Exhibit 2: Over-reporting of past vote for AKD in SLOTS survey, Sep. 2021–Sep. 2024



Note: Author's analysis of IHP SLOTS data. AKD 2019 vote share is computed after excluding "Did not vote" and respondents who would not have been eligible because of age.

Unfortunately, we did not detect this issue until one week before the election when the MRP model produced estimates that were in complete variance with the underlying raw data. Having identified the problem, we faced two choices. One was to not publish our final estimates, since we knew they were erroneous and flawed. The other was to not publish the flawed estimates, and to revise the model to mitigate the problem, and to communicate that we had changed the model. Neither option was a pleasant one. We decided in the interests of transparency to pursue the second option, but under significant time pressures owing to the Election Commission's embargo on media reporting that was to start from 19 September.

The temporary fix that was implemented was to treat the change in reporting as a real social phenomenon, and to model it before using it as input into the existing MRP model. This revised model was used to publish our final estimates on 18 September, with a disclosure that we had changed the model.

This fix did improve the accuracy of the final estimates in comparison to the flawed output from the established model, so it represented an improvement. However, the overall accuracy was not satisfactory, even if it was the best that could have been done in the circumstances.

Evaluation of alternative model improvements

The core problem that our IHP MRP modelling faced is how to exploit past voting behaviour in predicting current vote choice, in a context where there is substantial non-response to the question about past vote choice, and/or where social desirability bias influences responses and this bias changes over time. Researchers are certainly aware of this problem–as evidenced by recent research into the problem of the "shy Trump voter", but no adequate or accepted solutions currently exist. Nevertheless, for IHP MRP purposes, we still need to find a solution to relying on self-reported past vote choice.

Methodology

We evaluated four approaches to modelling current vote intent:

(i) **Model A**–Not considering past vote choice, equivalent to considering only demographic characteristics such as age, gender, ethnicity, income level, *etc*, when modelling current vote choice.

(ii) **Model B**–Using administrative data on the percentage of votes cast for candidates in the Presidential Election 2019 at the district level as additional controls. The administrative data sourced from the Elections Commission can be considered unbiased, but the district level controls will be weakly correlated with the past voting choice of individual survey respondents.

(iii) **Model C**–Using the SLOTS self-reported data to model the likelihoods that a respondent voted for specific candidates in 2019, and then using these likelihoods as controls when predicting current vote choice. The SLOTS data suffer from bias, but the modelled likelihoods will be more strongly correlated with the past voting choice of individual survey respondents.

(iv) **Model D**–Using both the administrative district level controls and the modelled likelihoods of individual level past vote choice when predicting current vote choice.

To minimize bias when modelling past vote choice using self-reported survey data (Methods B, C and D), the analysis included two additional features. First, only self-reported data from 1 December 2021 to 31 March 2024 was used for modelling, excluding the April–September 2024 data where substantial social desirability bias was observed. Second, the modelling process was constrained to ensure that its aggregate predictions matched the actual 2019 election results at the district and national levels. This last part was intended to eliminate aggregate bias, whilst retaining the information provided by the survey on how demographic characteristics influenced vote choice at the individual level. We reserve specific methodological details to a potential future academic paper.

Implementing an MRP-type approach over time involves allowing the influence of predictive variables to change over time. We do this by using cubic splines of time (measured in days). When implementing cubic spline smoothing, a key trade-off or choice is between using a smaller number of knots, which smooths the results more and reduces the impact of sample noise, and a larger number of knots, which makes the model more sensitive to changes over time but at the expense of more sensitivity to sample noise. In practice, this trade-off is also determined by computational constraints as more knots are more likely to lead to the computer failing to estimate the model or for the computation to take too long (hours or days), and partly made subjectively, since it is not possible to know what the underlying trend in vote intent is. For Model A, we used a cubic spline with 10 knots in interaction with the date, and 4 and 5 knots as interactions with other variables based on the combination used in the previous MRP model. For the other models, we took an agnostic approach, repeating all analyses using a range of knots ranging from four to 11 for the cubic spline interaction with time, and the same or fewer knots for other variables.

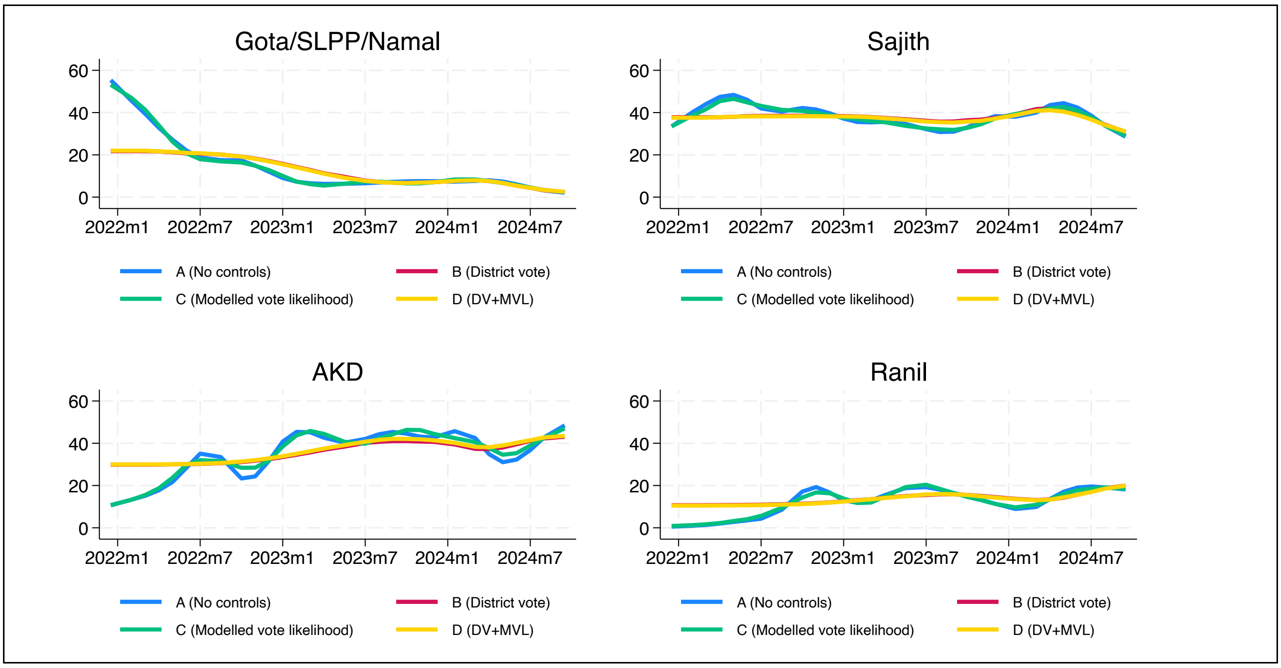
All analyses used only the SLOTS data collected from 1 Dec. 2021–20 Sep. 2024 (N=19,797 interviews).

Of these, we used only the interviews from Dec. 2021–Mar. 2024 for modelling past voting choice (N=11,384).

Results

In general, Models B and D, both of which included district vote controls, over-smoothed predictions over time, and both model sets failed to capture the big changes in vote intent in the first half of 2022 (Exhibit 3). Specifically, close inspection of SLOTS data does not support the possibility that support for AKD was more than 15–20% in early 2022, as estimated by Models B and D. There was surprisingly little difference in the trends generated by Models A (demographic controls only) and C (demographic controls plus modelled past vote choice in 2019), and both picked up changes in vote intent during 2023–2024, which we saw in the raw data and which we believe did occur. However, model A demonstrated bigger fluctuations, which probably reflects more sensitivity to sample noise.

Exhibit 3: Median estimates of current vote intent during Jan. 2021–Sep. 2024 by model type

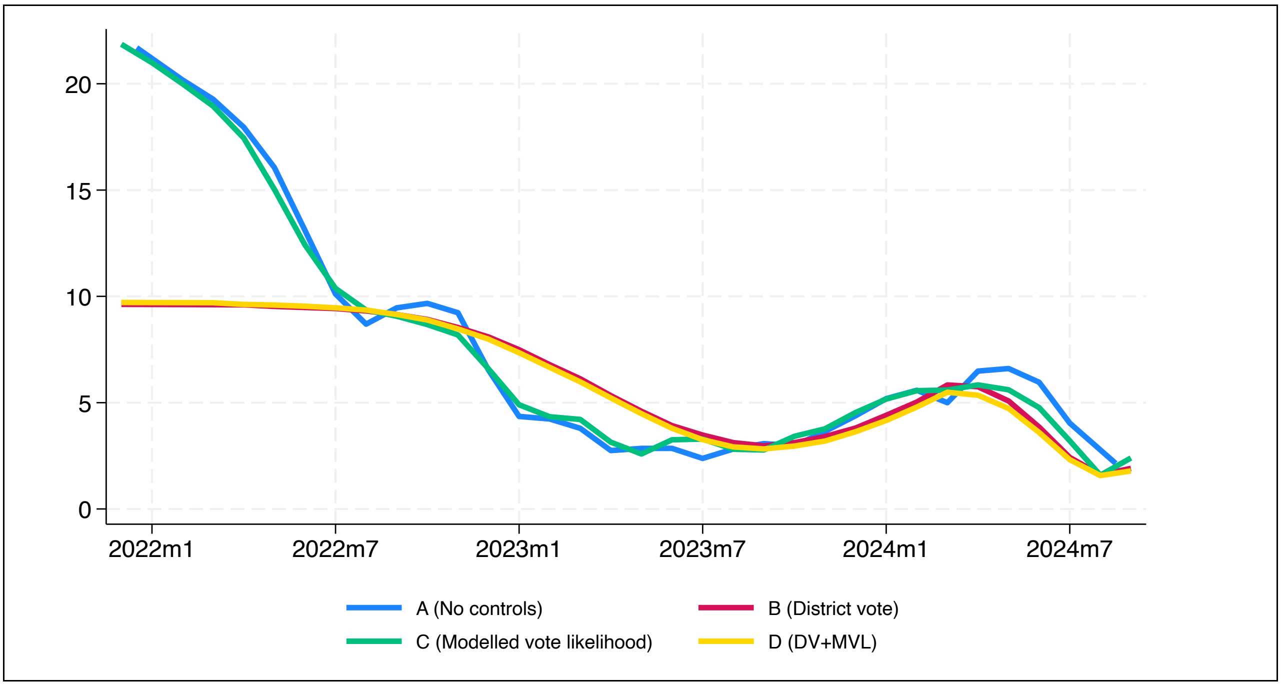


Note: Respondents were originally asked whether they would vote for Gotabaya Rajapaksa, and subsequently this was updated first to a generic SLPP candidate, and in August 2024 to Namal Rajapaksa. Chart lines refer to the median estimate of each set of models.

Although this could not have been computed before the election and so could not have been a selection criterion, it is worth noting that all four sets of models generated September 2024 estimates of vote intent that were very close to the actual election result. Average RMSEs ranged from 1.9 with Models B and D to 2.7 and 2.5 with Models A and C, respectively. This implies that all four models would have produced better estimates than the results IHP published just prior to the actual election.

When considering how historical estimates of voting intent predicted the result, the biggest differences are observed in Models A and C during 2022 through mid-2023 (Exhibit 4). But this reflects how voting choices substantially changed between 2021 and 2023. The August 2024 estimates would have been the best predictor of the final election result, with an average RMSE ranging from 1.6 (Models A and D) to 1.7 (Models B and C). This is interesting, since in the long-running US Gallup surveys, the estimates two months prior perform best in predicting the result of the US Presidential Election.

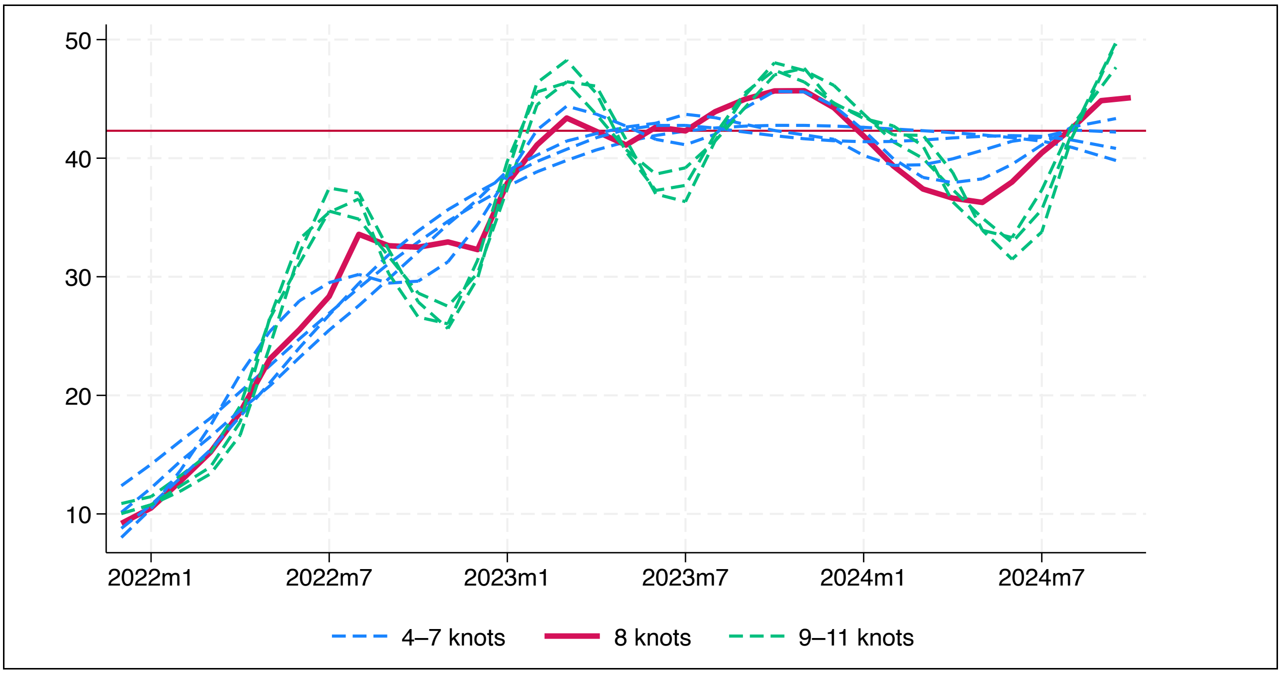
Exhibit 4: Median RMSE of monthly estimates of vote intent during Jan. 2021–Sep. 2024 as predictors of Presidential Election 2024 actual results by model type



Note: RMSE= Root Mean Squared Error. RMSE computed over five vote outcomes which included "Other".

On balance, given overall plausibility of the estimated time trends, assessed sensitivity to sample noise (something that MRP aims to reduce), and better model fit, Model C appears to perform best.

Exhibit 5: Median estimates of vote intent for AKD during Jan. 2022–Sep. 2024 using Model C by number of cubic spline knots interacted with time

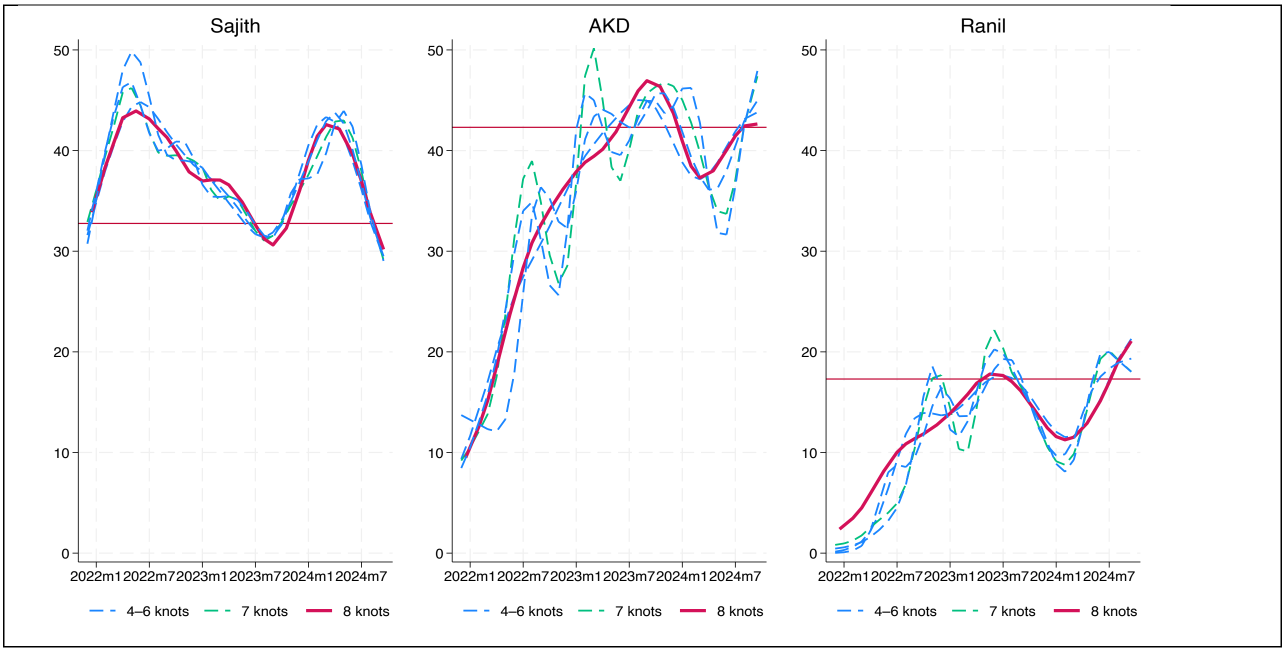


Note: Lines represent median estimates for each model combination. Horizontal crimson line indicates AKD's actual vote share in the Presidential Election 2024.

The next question is what combination of splines to use when estimating how the influence of past voting choice and other demographic characteristics on vote intent changes over time. For this we visually inspected the estimates generated using varying combinations of knot numbers. In the case of spline smoothing of time, a cubic spline with 8 knots appears to provide the best balance between smoothing and sensitivity to short-term changes in vote intent (Exhibit 5).

If using an 8-knot cubic spline to model the direct impact of time, then using 8-knot splines for interactions with other variables produces the best balance of smoothness and plausible fluctuations with time (Exhibit 6).

Exhibit 6: Median estimates of vote intent for the three leading candidates during Jan. 2022–Sep. 2024 using Model C with time smoothed using a cubic spline with 8 knots, and other variables smoothed with a varying number of 4–8 knots



Note: Lines represent median estimates for each model combination. Horizontal crimson lines indicate each candidate's actual vote share in the Presidential Election 2024.

This identifies a final model choice of Model C using cubic splines with 8 knots. This model provides the best balance of plausible vote intent trends during 2021–2024, sensitivity to fluctuations in support, and smoothing of sample noise, based on analysis of the survey data to 20 September 2024. If it had been used at that time to predict the Presidential Election 2024 results, it would have yielded estimates of AKD 42.7%, Sajith 30.5%, Ranil 20.9%, Namal 3.2%, and Others 2.8%. This would have been associated with an average error (RMSE) of 1.8%.

Additionally, this model would have estimated time trends very similar to those previously obtained and published by IHP using its earlier MRP model until about April 2024, after which growing social desirability bias affected model performance (Exhibit 7).

Overall, these final estimates of vote intent trends suggest the following:

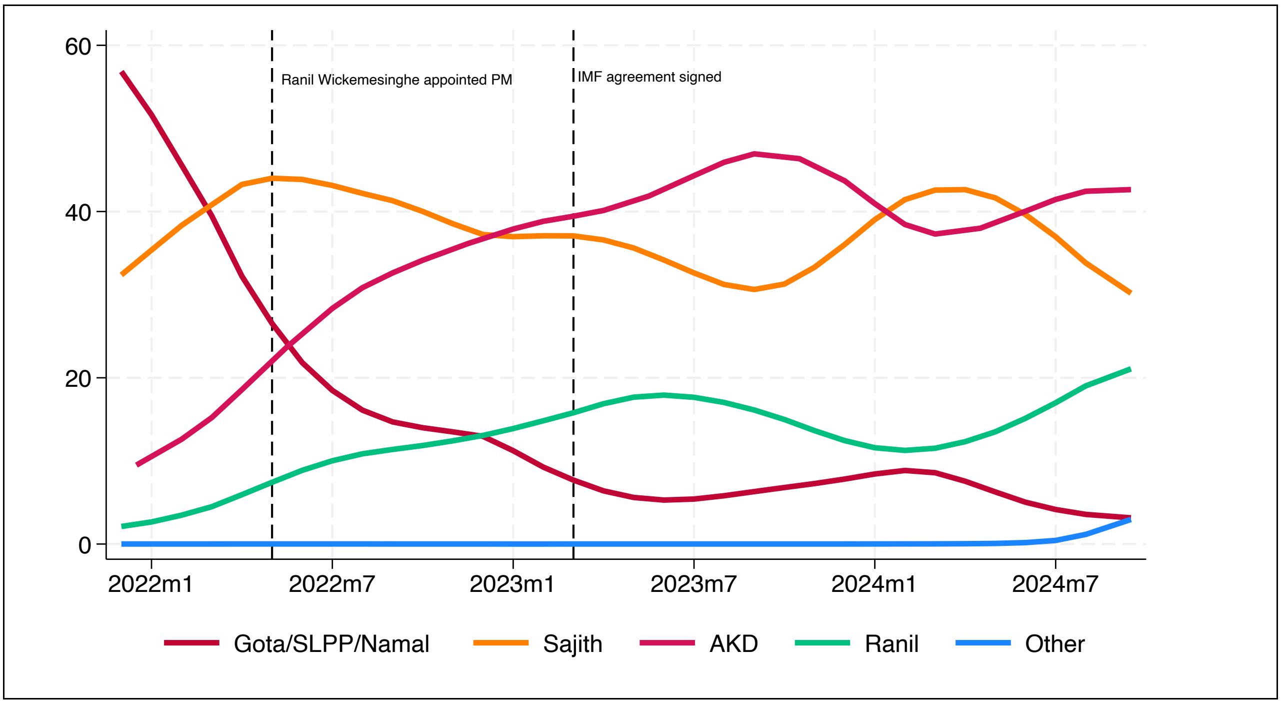
(i) Both Sajith Premadasa and AKD benefited in increased support from the collapse in support for President Gotabaya Rajapaksa in early 2022.

(ii) Support for Sajith increased until May 2022 when Ranil Wickremesinghe was appointed prime minister, after which it gradually fell.

(iii) Sajith took a brief lead in vote intent during early 2024 but was not able to sustain this once the Presidential Election campaigns started in mid-2024.

(iv) After campaigning started in earnest in early 2024, vote intent for the SLPP candidate fell again, with both AKD and Ranil gaining most.

Exhibit 7: Final best model estimates of vote intent Jan. 2022–Sep. 2024



Note: SLOTS did not allow respondents to select “Other” candidates until Jul. 2024, so estimated support for “Other” candidates is effectively zero until that time. Vertical dashed lines indicate key events.

Conclusions

This analysis indicates that using individually modelled past voting choice instead of self-reported past vote choice would have provided a satisfactory solution to the problem of social desirability bias as encountered by the SLOTS survey prior to the 2024 Presidential Election. Indeed, its estimates would have been associated with far smaller errors than the original IHP MRP model and the September 2024 model revision, with an RMSE of about 1.8 (range of 1.2–3.5 depending on the specific parameters used).

This result confirms that the MRP approach is a feasible approach to track vote intent in the Sri Lankan context, and to do so with relatively small monthly samples (N=400–700) contacted by telephone and with high levels of non-response and reporting bias. It also confirms that it is possible to accurately track voting intent in Sri Lanka using well-designed surveys. The estimates generated by the final selected method (Model C using cubic splines with 8 knots) track closely the previously published IHP vote intent estimates until about May 2024.

The analysis presented here also suggests a novel approach to tackling the general problem of handling reporting bias in past vote choice. It confirms some recent work that found that using area level voting aggregates does not offer significant improvements, and points to the potential of using modelled individual level vote choice.