

Rejoinder: The Authors Respond to “Interpreting Ethnicity and Urbanization in Malaysia’s 2013 General Election”

Jason Wei Jian Ng, Gary John Rangel,
Santha Vaithilingam, and Subramaniam S. Pillay

*In this article we respond to Thomas Pepinsky’s commentary on our article “2013 Malaysian Elections: Ethnic Politics or Urban Wave?” (both in this issue). We confirm that both ethnicity and urbanization play important roles in determining the incumbent ruling party’s percentage vote share in the thirteenth general election. In doing so, we address the various econometric issues raised by Pepinsky and clearly explain the advantages of our econometric methodology vis-à-vis the OLS analysis espoused by Pepinsky. Our main results indicate that Barisan National’s (BN) vote share from Bumiputera voters, regardless of urbanization levels of the parliamentary constituency, is below the 50 percent threshold. This result is surprisingly compensated by the more than 50 percent support for BN when Chinese voters are a small minority of the electorate. We also argue that Pepinsky’s statement that Malay voters are predominantly rural voters is inaccurate and provide evidence to the contrary. **KEYWORDS:** Malaysia, thirteenth general election, ethnic politics, fractional logit response model, urbanization, rural*

PEPINSKY’S COMMENTS AND THE UNDERLYING ARTICLES ON WHICH THEY ARE based have become part of an important debate on the role of ethnicity in Malaysian elections. The issues raised are both substantive and technical. In this response, we revisit and reconfirm our core empirical findings that Bumiputera support for Barisan National (BN) is below 50 percent regardless of parliamentary seat classification, and Chinese voters’ behavior is surprisingly not homogeneous but dependent on whether they are small in numbers or make up a substantial proportion of the electorate in a particular seat. Chinese voters help tilt the balance in BN’s favor, especially in rural areas where they make up a small minority of the electorate.

On the technical front, we do not attempt to answer every issue that Pepinsky has raised, but we have done additional analyses to show why the fractional logit methodology is superior to the OLS method. Pepinsky also commented that our use of the ethnic population total has the effect of changing the research question at hand from the analysis of the effect of *ethnic composition* to *ethnic population totals*. However, we argue that the use of ethnic population totals, *in the context of our model specification*, allows the interpretation of the results to be in terms of relatives (or proportions). While it may not be the best way to model ethnicity, it is a better option to model the data. In addition, we will demonstrate how the proportion of ethnic voters as used in Pepinsky's model is not able to identify the subtleties in the results, unlike our model, which uses total ethnic population. We follow up with an in-depth explanation of the key contributions of our article and address some of the claims made by Pepinsky.

Summary of Results and Methodology

Our aim in the article published in this volume on the 2013 Malaysian general election (GE13) as to identify which of the two factors, ethnicity or urbanization, provides a stronger explanation for the erosion of BN's popular vote. We do not assume that ethnicity and urbanization are mutually exclusive, as argued by Pepinsky, but instead, our analysis allows for the interaction of both factors. Our findings suggest that although the results of Malaysia's GE13 displayed an ethnic effect, complementing Pepinsky's main finding in his commentary, rapid urbanization of the country also played a role in determining the outcome of the election. Malaysians across ethnic lines voted overwhelmingly for Pakatan Rakyat (PR) in urban areas while support for BN remained reasonably strong in rural areas. Our results therefore do not rule out the ethnicity effect. Pepinsky finds that "1. Both district-level ethnic structure and district land area (a proxy for urbanization) predict BN vote shares at the district level. 2. Neither the effect of ethnicity nor that of urbanization can be reduced to the other. 3. There is no interactive effect between ethnicity and urbanization."

In short, the differences in results can be attributed to differences in the econometric modeling of the data, namely, model specification and choice of scaling of ethnic variables. Table 1 summarizes the differences in methodology between the two studies.

When confronted with different models reporting starkly contradicting results, how should one determine which of the two models is "cor-

Table 1 Differences in Econometric Methodology Between Ng et al. (2015) and Pepinsky (2015)

	Ng, Rangel, Vaithilingam, Pillay	Pepinsky
Econometric model	Fractional response model	Linear OLS
Independent variables	<ol style="list-style-type: none"> 1. Ethnic population totals 2. Area 3. Interaction terms between ethnic population total and Area 	<ol style="list-style-type: none"> 1. Ethnic composition 2. Area 3. State fixed effects
Model specification	A single econometric model containing all four ethnic population totals, Area, and interaction terms	Several econometric models with each model containing only one ethnic population composition, state fixed effects, subsequently augmented by Area

rect”? It is helpful to note the popularly cited maxim, “Essentially, all models are wrong, but some are useful” (Box and Draper 1987, 424). All models are wrong largely because no one ever knows the true model specification. The specified econometric model is a simplified representation of reality. Therefore, the practical question to ask is not which model is wrong, but rather, “how wrong do [the models] have to be to not be useful?” (Box and Draper 1987, 74). Against this backdrop, we use this guiding principle to further justify the choice of our model specification and variables.

Fractional Logit or Linear OLS Model?

Pepinsky casts doubt on the appropriateness and effectiveness of the use of a fractional response logit model in our article, demonstrating that “simple ordinary least squares (OLS) regression performs extremely well in modeling the relationships between ethnicity, urbanization, and vote share, such that employing the fractional logit approach makes no substantive difference to the inferences we draw from the analysis.” Pepinsky subsequently performs empirical analysis to show that the predicted vote share, based on his model specifications, is almost identical for both OLS and the fractional logit approach.

The argument above that Pepinsky makes against our use of the fractional logit model is an example of the difference in disposition between taking either a theory-driven or data-driven approach. While largely similar, econometrics is predominantly theory driven while statistics tend to

be data driven. Therefore, an econometrician develops a model based on economic (and other relevant) theories while a statistician may build a model after looking at datasets. The econometrician subsequently confronts the model with datasets to test the theory. The interested reader can refer to Rob Hyndman's blog post¹ for interesting insights into the differences between the two. In this context, it can be said that our econometric model is theory driven while Pepinsky's model is data driven.

It should be noted that Pepinsky makes the observation that the OLS model performs as well as the fractional logit model *ex post* (i.e., after the data have been observed and modeled). On the contrary, the choice of the fractional logit model is theory driven—the dependent variable to be modeled (vote share) is a proportion quantity known to be restricted to an interval between 0 and 1. In other words, the choice of our model was dependent on the known *nature* of the data, and not based on what the data reveal. It is along this line of reasoning that probit (and logit) models were developed to model binary dependent variables (i.e., variables that take on values of either 0 or 1), and tobit models were developed to model corner solution dependent variables (i.e., variables that have a population distribution that is spread out over a large range of positive values, but has a pileup at the value 0). These models were developed to account for the theoretical nature of the data.

From a theoretical perspective, the econometric model specified to model the proportion of vote share to BN must account for all possibilities, including the possibility of observing either a 0 or 1. However, in the event that the dependent variable does not realize values of either 0 or 1, as per the dataset observed for Malaysia's general election, the fractional logit model specified assigns 0 weights to the probability of observing these two value bounds in the log-likelihood function. Furthermore, by specifying a fractional logit model, we have not made any a priori assumptions on the restricted range of values that the dependent variable can take, except that it must be between 0 and 1. However, the use of OLS to model vote share makes an a priori assumption that the dependent variable *cannot* take on (possible) values of 0 and 1. Therefore, we are of the opinion that it is more prudent to use the fractional logit model to model proportion data, as compared to the OLS.

Pepinsky further questions the fractional response model by stating that "most political scientists use OLS to model vote shares [because] fractional regression methods rarely change substantive conclusions unless vote shares of zero appear frequently in the data," suggesting that the fractional logit model is only useful if a huge number of zero observations is present in the data. Far from it, Papke and Wooldridge (1996,

619–632) highlight that the fractional response model can help avoid the use of ad hoc transformations to handle data at the extreme values of 0 or 1—an added benefit of the model. Hence, the fractional response model was not designed to merely handle the extreme values of 0 or 1, but it has an advantage over earlier methods/models when handling such extreme values. Therefore, there is no restriction for the use of the fractional response model to be only for cases where there are 0s or 1s observed in the dependent variable.

Nevertheless, when judging the fractional logit model against the guiding principle of how wrong the model has to be for it to be not useful, the above explanations that we have put forth show that the use of the fractional logit model is not wrong at all. In fact, Papke and Wooldridge (1996, 619–632) argue that the linear regression model is not a good model specification if the dependent variable is bounded between 0 and 1, primarily because the effect of any particular explanatory variable cannot be constant through the range of the explanatory variable. We also reiterate the point in our article that from the theoretical perspective, the predicted values from an OLS regression are not guaranteed to lie in the unit interval, although we note from Pepinsky's empirical analysis that none of his predicted values exceeded the unit interval. However, our earlier explorations and considerations of other model specifications that applied OLS did produce predictions that exceeded the unit interval. Without going into the details of a particular model specification that we considered in early stages of this research, Figures 1 and 2 show the predicted values of the vote share to BN using OLS and fractional logit, respectively. It is clear that the fractional logit model can help constrain predicted values to be between the unit interval, but the OLS cannot. This example has also shown that producing sensible predictions is conditional on the model specification, to which we now turn our attention.

Model Specification and Ethnic Variables

In Pepinsky's article and comment, strong arguments and empirical evidence were put forth to question our model specification and the use of ethnic population total. In particular, Pepinsky puts forth a strong argument to use ethnic proportions, and to use a model specification that considers each of the ethnic variables in turn. Instead of tackling all the issues raised from the beginning to the end, we will demonstrate how from the outset, Pepinsky's model specification has limitations, therefore casting doubt on his subsequent analyses.

Figure 1 Predicted Vote Share to BN via OLS

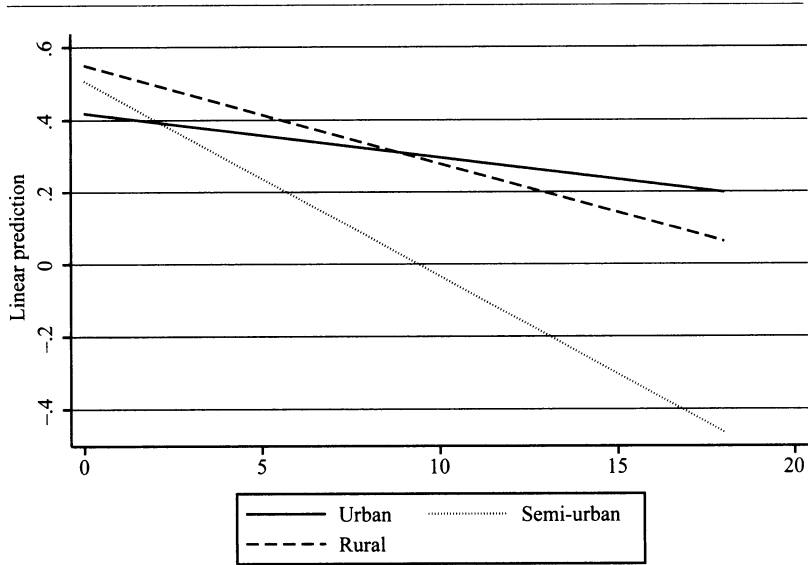
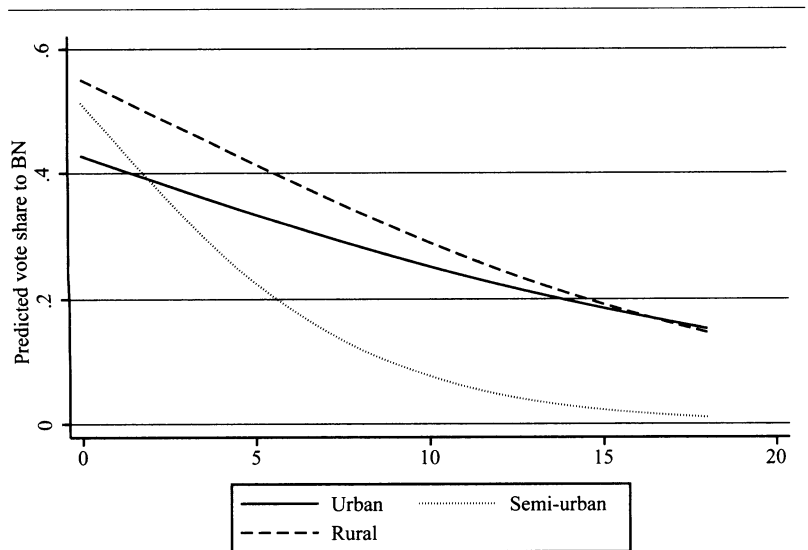


Figure 2 Predicted Vote Share to BN via Fractional Logit Model



Pepinsky has eloquently described the challenges in accommodating for the nature of the ethnic structure in statistical modeling. In short, statistical challenges arise from the fact that this set of compositional data faces a constraint whereby the ethnicity population shares must sum to one, that is, $F_{\text{Bumi}} + F_{\text{Chinese}} + F_{\text{Indian}} + F_{\text{Others}} = 1$. As an alternative to our approach of circumventing this challenge by using ethnic population total, Pepinsky proposes, and subsequently applies, “a simple, theoretically appropriate, and statistically sound modeling strategy for testing the effects of ethnic population shares on BN vote shares,” whereby he estimates four separate baseline regressions, with each regression including the ethnic proportion of one ethnic group only. The baseline model he considers is therefore represented as such:

$$BN \text{ Share} = \beta_0 + \beta_1 \% \text{ Ethnicity}_i + \delta \mathbf{D} + \varepsilon$$

where \mathbf{D} is a vector of state fixed effects, and ε is an error term. Pepinsky then claims that by doing so, it “preserves the substantive hypothesis about the predictive effects of ethnicity on BN votes, violates no assumptions about coefficient interpretability due to compositional data problems, and can be extended in a straightforward manner to interaction models.” Pepinsky’s claims are true—only if the sample space is in the real Euclidean space, which in this case, it is not.

The Case Against Using Proportions in Regression Modeling

Ethnic proportions are compositional data that are *constrained*, and the *components* of the composition must sum to a given constant, which in our case would be either 1 or 100 percent. Accordingly, this data structure is radically different from that of unconstrained data; statistical methods designed for unconstrained data are therefore inappropriate for application to (constrained) compositional data.² Therefore, the correct way to model compositional data is to *remove the constraints* of the compositional data via a transformation,³ perform traditional statistical methods (e.g., OLS) on the *transformed vectors*, and then transform the results back into the original space (Wang et al. 2013).

Pepinsky sought to offer a simple modeling alternative to remove the challenges in dealing with compositional data by including only one ethnic proportion variable in his regression model while deliberately excluding the other ethnic proportion variables. However, this approach does not remove the constraint on compositional data at all.⁴ In addition, Aitchison (1986), who made huge advances in this area in the 1980s,

warns against adopting this approach.⁵ Subsequently, any regression technique that is applied to the original untransformed compositional variable may give rise to misleading inferences (Hron, Filzmoser, and Thompson 2012, 1115–1128).

The Use of Ethnic Population Total

In recognizing the correct sample space of compositional data, our earlier works attempted to perform the isometric log-ratio transformation as per Hron, Filzmoser, and Thompson (2012). However, we decided against it because of the following:

- It would make the article too technical, distracting the reader from the political issues at hand.
- Interpretation of isometric log-ratio transformed variables is difficult, even in linear regression models, thereby making it hard to make useful inferences.
- No work has been done on how the isometric log-ratio transformation can be performed on quadratic variables and for interaction variables.

In lieu of the above, we decided to go with ethnic population totals as our measure of ethnicity, as the sum constraint would at least somewhat be removed. However, we acknowledge that this is not the best way to model ethnicity, which Pepinsky has correctly and strongly pointed out. Nevertheless, in our opinion, it is the better choice to model the data.

In referring to the guiding principle, again, as to how incorrect the model has to be for it to be not useful, we are of the opinion that leaving the original compositional variables as they are, while leaving out some parts from the regression model, is more incorrect than our approach of using ethnic population totals, which at least attempts to remove the sum constraint. Moreover, Pepinsky's inclusion of one ethnic variable in the regression, while leaving the rest of the ethnic variables out of the regression and effectively moving them into the regression error term, raises questions about potential endogeneity and omitted variable biases.

To address the point raised by Pepinsky—that our use of ethnic population totals has the effect of changing the research question at hand from the analysis of the effect of *ethnic composition* to *ethnic population totals*—let us consider the linear population regression function:

$$E(y|\mathbf{x}) = \beta_0 + \beta_1 \text{Bumiputera}_i + \beta_2 \text{Chinese}_i + \beta_3 \text{Indians}_i + \beta_4 \text{Others}_i$$

where *Bumiputera*, *Chinese*, *Indians*, and *Others* represent the respective number of voters in each of those ethnic groups. If we consider, for ex-

ample, the coefficient β_1 , it is interpreted as the expected change in vote share to BN from an increase in the number of Bumiputera voters, while holding the number of voters in all the other ethnic groups constant. Therefore, the interpretation is still, to some extent, in terms of relatives (or proportions). It is only if the population regression function is presented as

$$E(y|\mathbf{x}) = \beta_0 + \beta_1 \text{ Ethnicity}_i$$

where $i = \text{Bumiputera, Chinese, Indians and Others}$, that Pepinsky will then be correct to say that we would have only been examining the effect of ethnic population totals. In this case, β_1 would be interpreted as the expected change in vote share to BN from an increase in the number of Bumiputera voters. The latter case would then have been examining the effect of ethnic population totals instead. This is, therefore, one of the merits of including all four ethnic groups in our model specification.

To conclude the matter on technical specification issues, the estimation of the fractional logit model and the use of ethnic population totals as variables in our model specification are fully justified. The results are therefore credible and are subsequently useful to draw insightful inferences.

Urbanization and Ethnicity in GE13 Outcome

Turning to his other comments, Pepinsky asserts that “even after decades of urbanization, Malay voters still tend to be rural voters.” While it is true that rural voters in Peninsular Malaysia tend to be Malay voters, the converse is not necessarily true. With the rapid rural-urban migration and a higher Malay population growth, an increasing proportion of urban voters are Malay voters (Tey 2012). Many urban constituencies, especially in the east coast states, Kuala Lumpur and Selangor, have either Malay majority or plurality. As this trend continues in the next few decades, the Malay/rural versus non-Malay/urban paradigm that has underpinned much discussion on Malaysian politics in the past may need to be reviewed. In fact, a major motivation for our article is to explore the ramifications of this trend on voting patterns.

Table 2 shows the electoral outcome in parliamentary seats that are classified as urban by Politweet, and where Malays make up more than 50 percent of the electorate.

Of the fourteen parliamentary seats shown, ten were won by PR and only four by BN. If we follow Pepinsky’s argument to disregard the urbanization effect and assume that a higher proportion of Bumiputera voters entails a higher proportion of vote share for BN, most of the

Table 2 Voting Results from Urban Seats with Bumiputera Majority

Voting District No. and Name	Majority	Total Number of Voters	Bumiputera Voters		Chinese Voters		Indian Voters		Other Voters		Winning Party
			% of Voters	No. of Voters	% of Voters	No. of Voters	% of Voters	No. of Voters	% of Voters	No. of Voters	
P.125 Putrajaya	5,541	15,791	95.5	15,080	1.1	174	3.1	490	0.3	47	BN
P.036 Kuala Terengganu	10,785	94,406	89.2	84,210	10.1	9,535	0.6	566	0.1	94	PR
P.021 Kota Bharu	15,970	81,268	82.9	67,371	15.8	12,840	0.9	731	0.4	325	PR
P.108 Shah Alam	10,939	99,957	70.5	70,470	14.9	14,894	14.2	14,194	0.4	400	PR
P.119 Titiwangsa	866	55,282	68.7	37,979	19.9	11,001	10.4	5,749	1.0	553	BN
P.083 Kuantan	4,515	56,280	62.9	35,400	33.1	18,629	3.8	2,139	0.2	113	PR
P.009 Alor Star	1,873	69,009	61.3	42,303	33.6	23,187	4.7	3,243	0.4	276	PR
P.118 Setiawangsa	1,390	62,309	57.8	36,015	30.4	18,942	11.2	6,979	0.6	374	BN
P.121 Lembah Pantai	1,847	72,396	55.9	40,469	22.8	16,506	20.0	14,479	1.3	941	PR
P.116 Wangsa Maju	5,511	67,775	54.5	36,937	36.2	24,535	8.8	5,964	0.5	339	PR
P.124 Bandar Tun Razak	11,832	90,993	53.5	48,681	37.4	34,031	8.7	7,916	0.4	364	PR
P.137 Bukit Katil	5,447	99,438	53.1	52,802	40.8	40,571	5.7	5,668	0.4	398	PR
P.160 Johor Bahru	10,134	96,321	52.1	50,183	42.6	41,033	5.0	4,816	0.3	289	BN
P.109 Kapar	23,790	144,159	51.8	74,674	33.6	48,437	14.4	20,759	0.2	288	PR

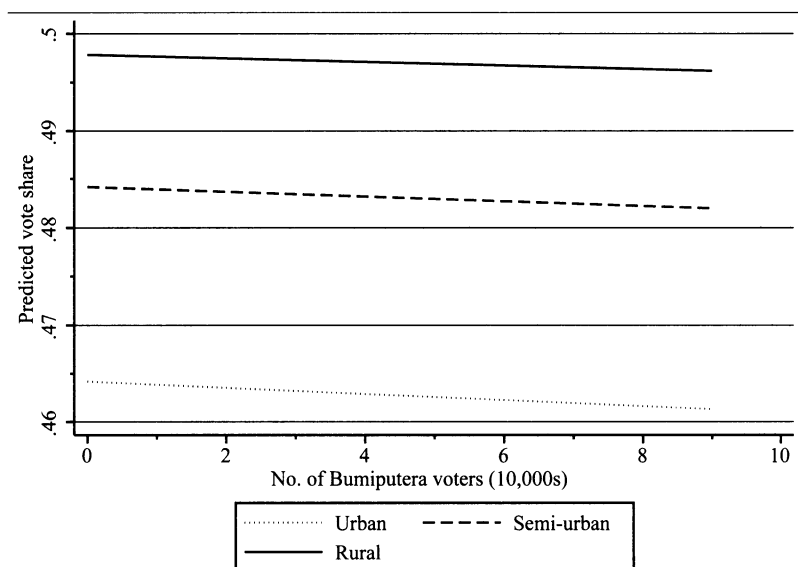
Note: Urban seats as identified by Politweet.

parliamentary seats in Table 2 should have been won by BN. It is also interesting to note that of the four parliamentary seats won by BN, two of them (Titiwangsa and Setiawangsa) were won with majorities of less than 4 percent of the total number of registered voters. In the case of PR, only two (Alor Star and Lembah Pantai) out of the ten parliamentary seats won do not exceed the 4 percent threshold. This analysis suggests that the urbanization effect in determining election outcomes should not be easily discounted. However, as we are using aggregate data in our analysis, there are possible outliers that go against the underlying trend. Specifically, the other two urban parliamentary seats won by BN with large majorities listed in Table 2 are Putrajaya (also highlighted in Pepinsky's commentary) and Johor Bahru. The Putrajaya parliamentary seat encompasses the new administrative capital of Malaysia and Malay voters registered in that constituency are predominantly government civil servants. As for the case of Johor Bahru, BN candidate Datuk Shahrir Samad's personal popularity may have played an important role in ensuring BN's large majority in this urban constituency.

Bumiputera Support for BN

In contrast to Pepinsky's findings whereby a higher proportion of Bumiputera voters corresponds to a higher predicted mean vote share to BN (Pepinsky 2015, Figure 6), our results indicate that Bumiputera support for BN ranges from 46 percent for an urban seat to just below the 50 percent threshold for a rural seat. Bumiputera support remains stable at these percentage levels across increasing numbers of total Bumiputera voters, irrespective of the seat urbanization classification. We reproduce Figure 3 in our article as Figure 3 in this reply.

What is revealing about these results is the fact that even for rural seats, the predicted average BN percentage vote share is less than 50 percent. This finding is surprising, as we had expected rural Bumiputera support for the BN to be significantly higher than 50 percent. There are two possible factors that may have contributed to the lower than expected level of support. First, since we use aggregated data, the support that Parti Islam Se-Malaysia (PAS) receives in its traditional stronghold states of Kelantan, Terengganu, and Kedah may mask a higher level of support that the BN receives in the rest of the rural areas of Peninsular Malaysia. It also can be arguably inferred that such levels of support are due to the presence of out-of-town Bumiputera voters who return to their respective rural constituencies to cast their ballots. These Bumiputera voters work in cities located in urban areas and their mindsets are attuned to issues

Figure 3 Predicted Vote Share for BN in Urban, Semi-Urban, and Rural Constituencies for Varying Numbers of Bumiputera Voters

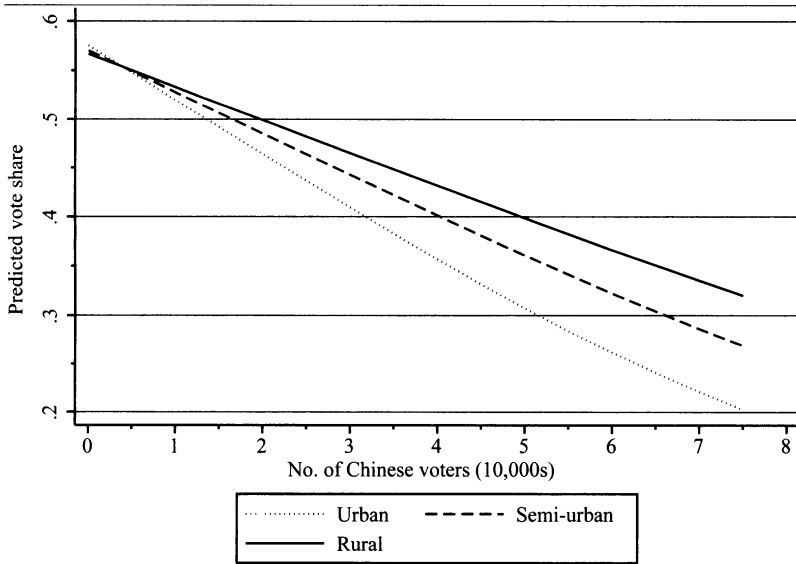
that affect urbanites. They may elicit the mentality of “urban chauvinism” as espoused by Thompson (2013). In conclusion, we agree with Pepinsky’s view that BN has firm support of the Bumiputera electorate in the rural belt. However, we find this support to be less than 50 percent, even for rural seats. This means that some level of support has to come from the Chinese voters, which constitutes the second part of our findings. We now elaborate on our findings on rural Chinese voters and also illustrate an added advantage of our econometric model specification.

Chinese Voters Helped BN Cross the Finish Line

In our article, we found that Chinese voters voted overwhelmingly for PR in urban seats. However, the significant results we would like to reiterate are depicted in Figure 4, which is a reproduction of Figure 4 in our article. It also depicts an added advantage of using total ethnic population rather than ethnic population proportions, as put forth by Pepinsky.

Figure 4 clearly indicates that Chinese support for BN declines drastically when the total Chinese voter population rises for all constituency

Figure 4 Predicted Vote Share for BN in Urban, Semi-Urban, and Rural Constituencies for Varying Numbers of Chinese Voters



types, with urban constituencies showing the steepest decline. The surprising result is the level of Chinese voter support for BN when their numbers are small. Support for BN is around 54 to 57 percent when their numbers are around 5,000. This support is crucial for BN because Bumiputera support is slightly less than 50 percent, even in rural areas, as we have shown. These Chinese voters essentially helped BN cross the finish line. If we had used the proportion of ethnic voters, as argued by Pepinsky, we would not have uncovered this new contribution to the Malaysian politics literature.

Conclusion

We have clarified that although our initial intent was to find which variable, ethnicity or urbanization, was the dominant factor in explaining the thirteenth Malaysian general election, our results have shown that both variables are important in determining BN vote share. We do not take a “horserace” perspective, as put forth by Pepinsky, but rather have shown that both variables operate in unison, with the *Chinese-Urbanization* factor being a dominant influence on vote share to BN. We have also ar-

gued that Malay voters need not be predominantly rural. The econometric methodology we have used in our article breaks new ground toward contributing to the vast literature on Malaysian politics. The surprising results on Bumiputera and Chinese voters' support for BN in rural areas shed new insights on voter behavior that could never have been uncovered by OLS methodology. Much more can be done to improve our modeling of Malaysian voters' behavior. Future research should incorporate control state variables, as per Pepinsky's analysis.

Future work at the micro level can be directed at explaining the observed behavior of Chinese voters when they make up a small minority of the electorate in any particular seat. Reliance on the government may be an important contributing factor. However, more research needs to be done to understand the motivations for this behavior.

Our results present a working hypothesis that can be answered only in the definitive if we look at micro-level data: looking at the voting patterns in individual voting streams across voting districts that are categorized as urban, semi-urban, or rural. However, comparing those who actually vote against the electoral roll to determine ethnicity is indeed an impossible task for any social scientist, given that ethnicity is not listed for each voter on the electoral roll. The approach taken in our original article seems to be the better option. Other data on the electoral roll can open up more avenues for research within the context of Malaysian politics. Electoral rolls provide information on the age and gender of the voter. Future research can tap these data to examine the voting patterns of women as well as the voting patterns of various age cohorts (see, for example, Khor 2014, 89–121).

Notes

1. <http://robjhyndman.com/hyndsight/statistics-vs-econometrics/>.
2. The sample space that compositional data occupy is referred to as the "Aitchison geometry on the simplex" (Hron, Filzmoser, and Thompson 2012, 1116). In contrast, unconstrained data are associated with the real Euclidean sample space. In particular, there is a nonlinear relation between the Euclidean space and the Aitchison geometry, therefore making it inappropriate for standard statistical methods designed for unconstrained data to be applied *directly* to constrained compositional data.
3. Some of the transformations developed over the decades include the additive log-ratio transformation, the centered log-ratio transformation, and the isometric log-ratio transformation.
4. The one ethnic proportion variable that remains in the model is still *not* in the Euclidean space.

5. Aitchison classifies those who opt for the nontreatment of compositional variables as “wishful thinkers.” In particular, Aitchison (n.d., 111) says, “No problem exists (Gower 1987) or, at worst, it is some esoteric mathematical statistical curiosity which has not worried our predecessors and so should not worry us. Let us continue to calculate and interpret correlations of raw components. After all if we omit one of the parts the constant-sum constraint no longer applies. Someday, somehow, what we are doing will be shown by someone to have been correct all the time.”

References

- Aitchison, John. “A Concise Guide to Compositional Data Analysis.” n.d. www.leg.ufpr.br/lib/exe/fetch.php/pessoais:abtmartins:a_concise_guide_to_compositional_data_analysis.pdf.
- . 1986. *The Statistical Analysis of Compositional Data*. New York: Chapman and Hall.
- Box, George Edward Pelham, and Norman R. Draper. 1987. *Empirical Model-Building and Response Surfaces*, Wiley Series in Probability and Statistics. New York: John Wiley and Sons.
- Gower, John C. 1987. “Introduction to Ordination Techniques.” In *Developments in Numerical Ecology*, ed. Pierre Legendre and Louis Legendre, NATO ASI Series. Berlin: Springer.
- Hron, K., P. Filzmoser, and K. Thompson. 2012. “Linear Regression with Compositional Explanatory Variables.” *Journal of Applied Statistics* 39, 5: 1115–1128.
- Khor, Yu Leng. 2014. “The Political Tussle over Felda Land Schemes—UMNO Strengthens Its Malay Rural Fortress in 13th General Elections.” *Kajian Malaysia* 32, Supp. 2: 89–121.
- Ng, Jason Wei Jian, Gary John Rangel, Santha Vaithilingam, Subramaniam S. Pillay. 2015. “2013 Malaysian Elections: Ethnic Politics or Urban Wave?” *Journal of East Asian Studies* (this issue).
- Papke, Leslie E., and Jeffrey M. Wooldridge. 1996. “Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates.” *Journal of Applied Econometrics* 11, 6: 619–632.
- Pepinsky, Thomas B. 2015. “Interpreting Ethnicity and Urbanization in Malaysia’s 2013 General Election.” *Journal of East Asian Studies* (this issue).
- Tey, Nai Peng. 2012. “Internal Migration in the Klang Valley of Malaysia: Issues and Implications.” *Malaysian Journal of Chinese Studies* 1: 40–59.
- Thompson, Eric C. 2013. “Urban Cosmopolitan Chauvinism and the Politics of Rural Identity.” In *Cleavage, Connection and Conflict in Rural, Urban and Contemporary Asia*, ed. Tim Bunnell, D. Parthasarathy, and Eric C. Thompson, ARI—Springer Asia Series. Dordrecht, Netherlands: Springer.
- Wang, Huiwen, Shanguan Liying, Wu Junjie, and Guan Rong. 2013. “Multiple Linear Regression Modeling for Compositional Data.” *Neurocomputing* 122: 490–500.

US Taiwan Strait Policy: The Origins of Strategic Ambiguity

DEAN P. CHEN

“Chen’s probing exploration of the origins of the American commitment to Taiwan deepens our understanding of key determinants that have continued to influence US policy in relations with Taiwan and China.”

—Robert Sutter, George Washington University

Why did the Truman administration reject a pragmatic approach to the Taiwan Strait conflict—recognizing Beijing and severing ties with Taipei—and instead choose the path of strategic ambiguity? Dean Chen sheds light on current US policy by exploring the thoughts and deliberations of President Truman and his top advisers. Chen also highlights the very unambiguous, and continuing, liberal aims of US Taiwan policy.

CONTENTS

- US Interests in Taiwan.
- US-China-Taiwan Relations from Nixon to Obama.
- Wilson’s Vision for an Open China.
- Freeing Taiwan from Communist Domination.
- The Inception of Strategic Ambiguity.
- The Future of US Policy in the Taiwan Strait.

2012/298 pages • hc \$68.50

1800 30th Street, Suite 314 • Boulder, CO 80301 • Tel: 303-444-6684

www.rienner.com

**US
Taiwan
Strait
Policy**

The Origins of
Strategic Ambiguity

Dean P. Chen