Informal Voting in the 2004 Australian Election: a brief look at the aggregate data

Simon Jackman*

April 25, 2005

In this brief note I document several interesting features of informal voting in the 2004 Australian election for the House of Representatives:

1. As the number of candidates on the ballot paper increases (i.e., ballot length) in a given electoral division, so too does the rate of informality.

2. As the proportion of an electoral division's voters who come from non-English speaking households (NESH) increases, so too does the rate of informal voting in the division.

3. The relationship between ballot length and informality becomes stronger as NESH increases; similarly, the relationship between NESH and informality becomes stronger as ballot length increases.

4. As the proportion of a division's population with tertiary education increases, informality decreases.

5. There are interesting patterns of state-level variation in rates of informality, with divisions in NSW, QLD and ACT recording substantially higher rates of informality than we would expect given the divisional characteristics in the statistical modeling (i.e., number of candidates running and levels NESH and tertiary education). This is consistent with optional preferential voting in state and territory elections operating as a source of confusion for voters in federal elections.

*Associate Professor and Director, Political Science Computational Laboratory, Department of Political Science and (by courtesy) Department of Statistics, Stanford University, Stanford, CA, 94305-6044, USA. e-mail: jackman@stanford.edu web:http://jackman.stanford.edu

I stress that

1. this analysis is *preliminary*; my point here is not to give a comprehensive study of informality in 2004 or more generally, but rather, to highlight some important features of the phenomenon. In particular, I apologize in advance to those who have worked on this question in the past, but whoc3se work I am unable to give sufficient credit and consideration in this brief note. This list includes (but is certainly not limited to) Medew (2003), McAllister, Makkai and Patterson (1992) and Young (2004). In particular, the regression analysis reported below builds on similar analyses reported in work on informality in earlier elections by McAllister and co-authors and Medew.

2. this analysis is *cross-sectional*; I make no claims about changes in rates of informality over time.

3. this analysis uses *aggregate data*; I can not make strong claims about the individual-level processes that produce informal ballots, although the patterns I describe in the division-level data are consistent with informal voting being disproportionately concentrated among particular sets of voters.

1 Data

The data used in the analysis comes from two sources. Electoral data (informality, ballot length, and information about incumbents) comes from the Australian Electoral Commission's website. Data on the proportion of a division's population who speak a language other than English at home and tertiary education comes from the aggregations of 2001 Census data published by the Parliamentary Library (Kopras, 2004).

Figure 1 displays rates of informal voting in the 150 House of Representatives divisions in 2004. The distribution has a pronounced skew to the right: the minimum rate of informality is just 2.8% (in Higgins, Peter Costello's seat in inner-metropolitan Melbourne), and most divisions record rates of informality less than 5%, but the maximum rate of informality is 11.8% (in Greenway, in Sydney's western suburbs). After Greenway, the top ten divisions in terms of informality are (in descending order) Reid, Blaxland, Chifley, Prospect, Fowler, Watson, Parramatta, Kingsford-Smith, and Werriwa. All except Kingsford-Smith are in Western Sydney; all are held by the Labor Party; all have relatively high proportions of voters from non-English speaking households.

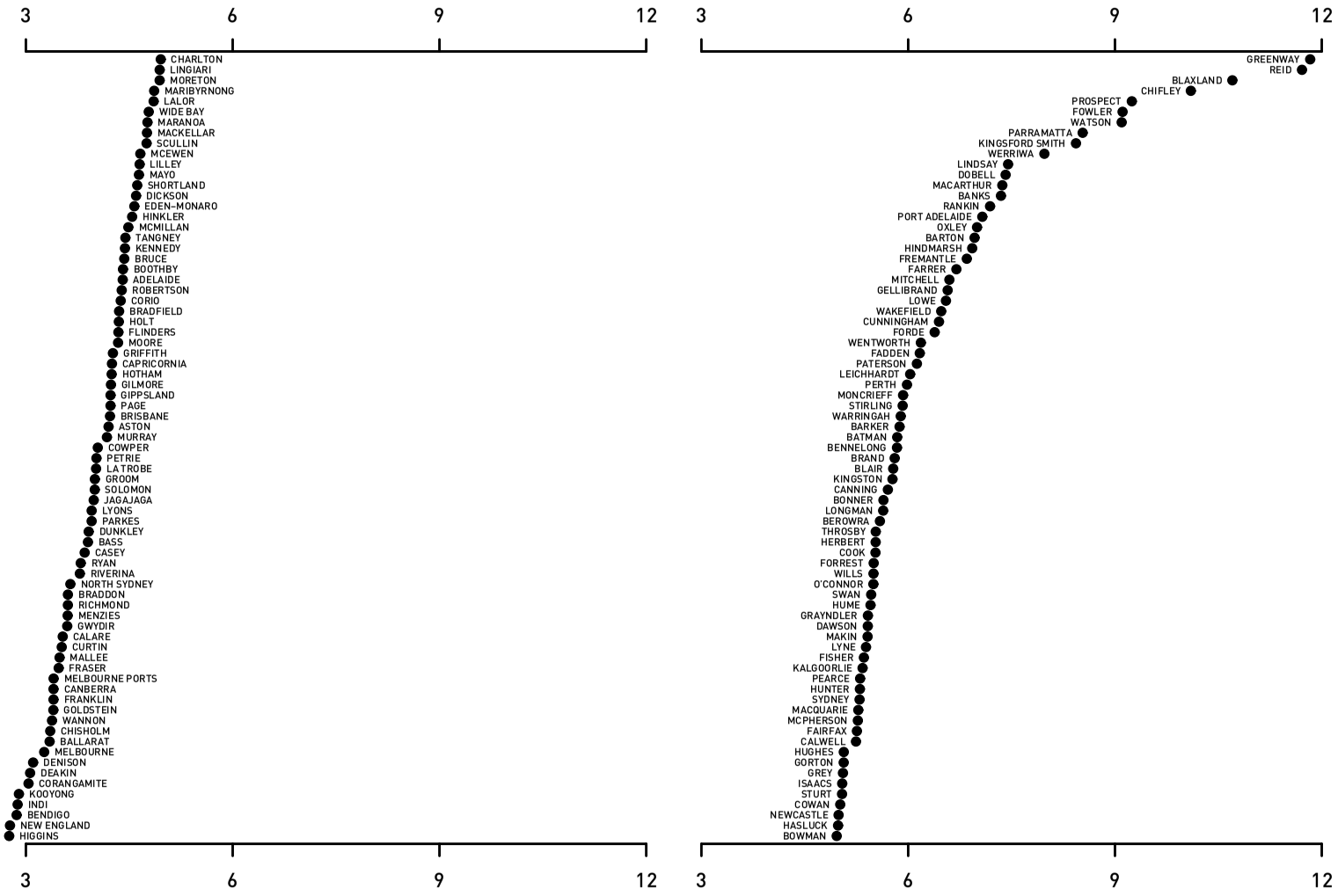

Figure 1: Informality Rate, as Percentage of Turnout.

Figure 2 shows a scatterplot of informality (vertical axis) against ballot length. The modal ballot length in 2004 was 7 (39 divisions), but 14 candidates stood in Greenway (the seat with the highest rate of informality) and 12 in Dobell. The minimum ballot length was 4, seen in 3 divisions (Riverina, Throsby and Braddon). The scatterplot strongly suggests that as ballot length increases, so too does informality; the correlation between the two variables is a rather healthy .43.

The relationship between informality and the prevalence of non-English speakers (NESH) is shown in Figure 3. The NESH variable is quite skewed: most divisions have no more than 10% of their population speaking a language other than English at home, but five divisions are "majority NESH", with values of NESH in excess of 50%: these are Blaxland, Fowler, Reid, Watson (all in Western Sydney) and Gorton (in Melbourne). The relationship between informality and NESH is non-linear, at least in the absence of any statistical controls: a quadratic regression model fits the scatterplot better than a linear regression (r^2 = .29 vs r^2 = .26), but for simplicity, I overlay the linear fit in Figure 3. High values of NESH tend to accompany high values of informality, but the relationship is far from perfect.

The last piece of exploratory analysis I report is a scatterplot of informality and the percentage of a division's population (aged 15 and over) with tertiary qualifications (a degree or higher, or a diploma), in Figure 4. The median level is 15.8%, but 10 divisions record values on the variable in excess of 35%. In general, as the proportion of people with tertiary degrees increases, the rate of informality decreases. Divisions with extremely high rates of informality almost all have relatively low proportions of the division population holding tertiary qualifications.

2 Multivariate Analysis

I use multiple regression analysis to assess the joint effects of the variables introduced above. In particular, I assess *interactions*, and, in particular, the interactions between ballot length and non-English speaking households. The conjecture here is that non-English speakers find the task of validly enumerating preferences on the House of Representatives ballot more challenging than do English speakers, and that as ballot length increases, rates of informality will increase for non-English speakers relative to English speakers.

Table 1 presents ordinary least squares regression coefficients and their standard errors for two models. I discuss the findings of the first model, since the results are virtually identical, save in one important respect to be

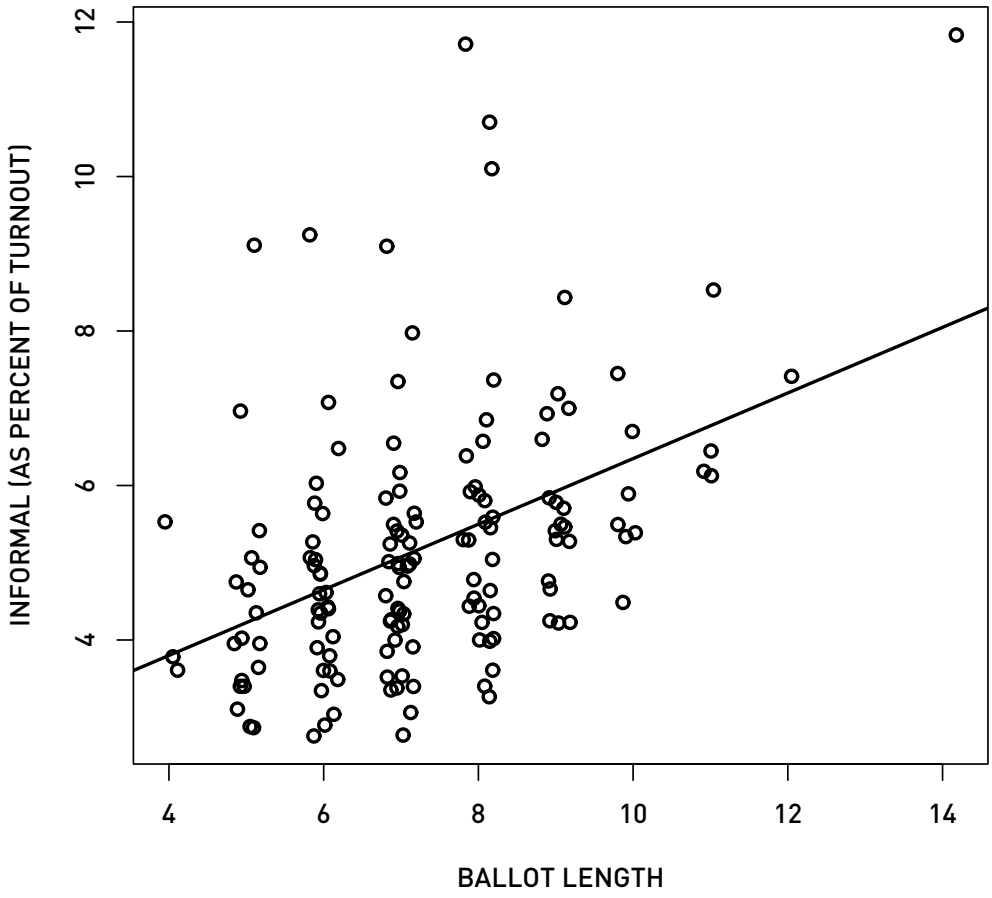

Figure 2: Informality and Ballot Length, 2004 House of Representatives. For graphical clarity, the values of ballot length have been randomly jittered. The upward sloping line is an ordinary least squares (OLS) regression line ($r^2 = .18$).

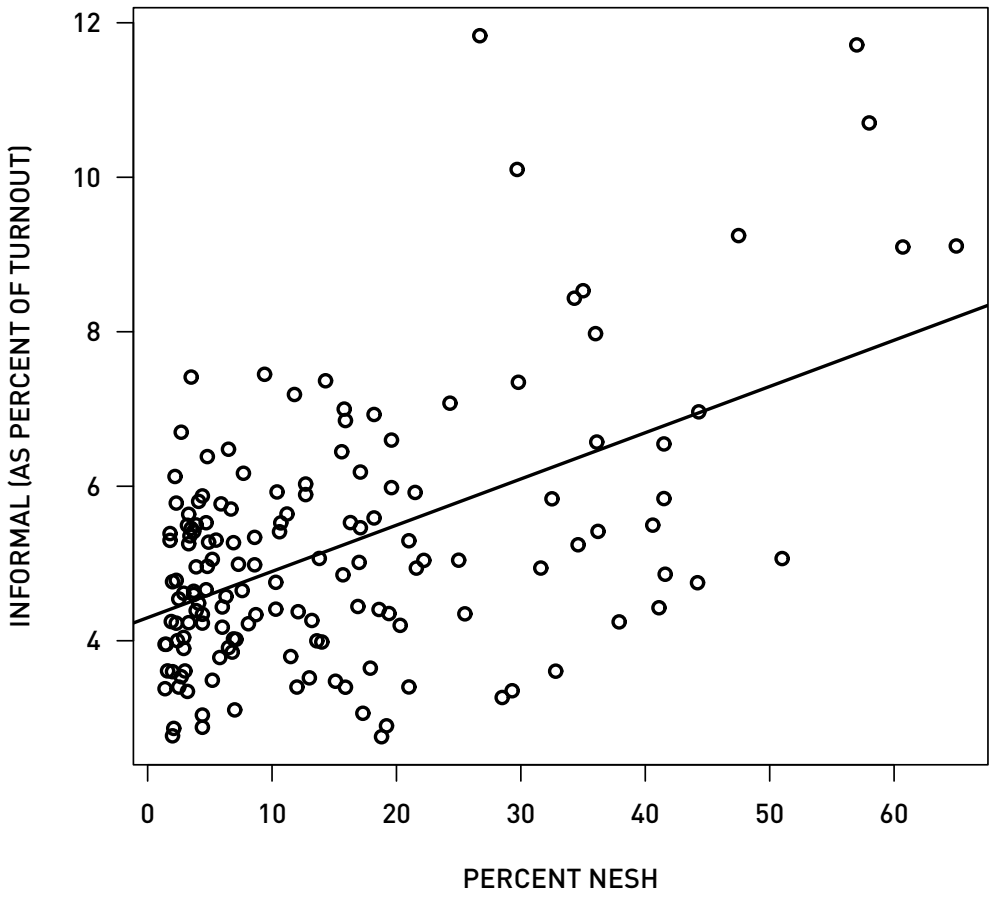

Figure 3: Informality and Percent of Persons speaking a Language Other than English at Home (NESH), 2004 House of Representatives. The upward sloping line is an OLS fit ($r^2 = .26$).

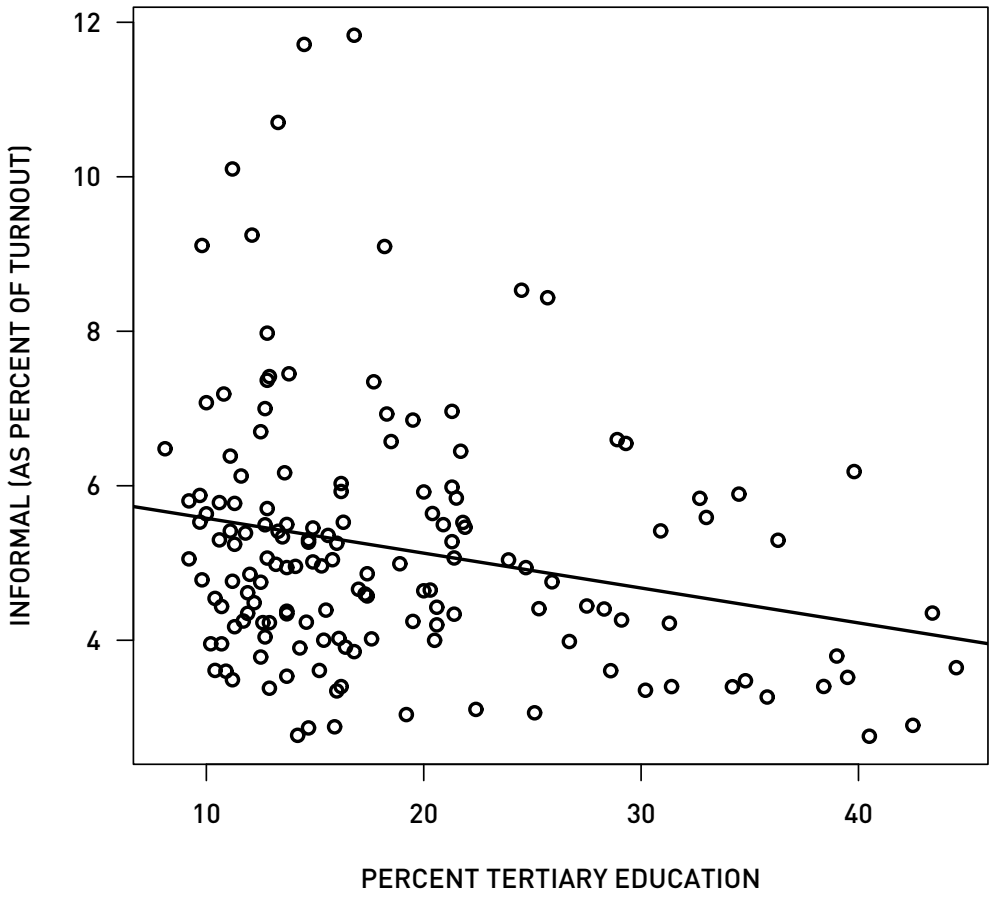

Figure 4: Informality and Percent of Persons with a Tertiary Degree. The downward sloping line is an OLS fit (r^2 = .05).

	Model 1	Model 2
Intercept	3.3	3.0
	(.59)	(.49)
Ballot Length	.30	.26
	(.073)	(.061)
NESH	-.0069	-.0055
	(.026)	(.022)
Ballot Length × NESH	.012	.012
	(.0037)	(.0031)
Tertiary	-.076	-.077
	(.010)	(.0084)
NSW/QLD/ACT		1.1
		(.14)
r^2	.64	.75
$\hat{\sigma}$	1.0	.84

Table 1: Multiple Regression Analysis, Informality in House of Representatives Elections, 2004. Standard errors in parentheses.

discussed in the next section. The model fits the data reasonably well (r^2 = .64), accounting for roughly 2/3 of the observed variation in rates of informality. The message of the previous section's scatterplots is largely replicated, with one important addendum. The interaction between NESH and ballot length is large relative to its standard error (t = 3.3), indicating that (1) as ballot length increases, the effect of NESH on informality gets stronger; (2) as NESH increases, the effect of ballot length also gets larger. Specifically, at the lowest level of ballot length (4), the implied coefficient on NESH is -.0069 + 4 × .012 = .041 (i.e., holding ballot length constant at 4, an increase in NESH of 10 percentage points would increase informality by about four-tenths of a percentage point). But this effect increases with ballot length. By the time, ballot length reaches its modal value of 7, the implied coefficient on NESH is up to .077, meaning that a 10 percentage point increase in NESH would produce an increase in informality of almost 0.8 percentage points. At the maximum value of ballot length, 14, the coefficient on NESH is now .16, and a 10 percentage point increase in NESH is associated with a 1.6 percentage point increase in informality. That is, when the ballot is short, NESH has a tiny impact on rates of informality, but this effects quadruples over the range of ballot lengths observed in the data.

A similar accounting exercise can be performed for the effects of ballot length itself, conditional on different levels of NESH. The estimated coefficient for the direct effect of ballot length is .3, meaning that for each additional candidate on the ballot and in divisions with no non-English speakers, informality increases by 3/10 of a percentage point. As NESH increases from its minimum of 1.4% to its median of 14.9%, the implied effect of ballot length increases to produce about 1/2 of a percentage point increase in informality with the addition of an extra candidate. At, say, the 95-th percentile of NESH (e.g., Scullin, with 44.2%), the implied effect of ballot length is up to .8, and at the maximum of 65% (in Fowler), the ballot length effect is over 1.0, meaning that each additional candidate produces more than a percentage point increase in informality. That is, over the range of NESH observed in the data, the effect of ballot length more than triples.

The combination of increasing levels of NESH and ballot length produces large rates of informality. This helps account for the results in Greenway, where relatively high rates of NESH (123rd out of 150) and a ballot paper with 14 candidates, produced a double-digit rate of informal balloting. Add to the mix the fact that Greenway has roughly only the median level of tertiary education per division, the fact that the incumbent member was not running for re-election, and the fact that NSW has optional preferential voting in state elections, and we have a fairly thorough understanding of what happened in

that particular seat.

3 Institutional Sources of Variation in Informality: optional preferential voting in state elections

NSW, Queensland and the ACT have optional preferential voting in their legislative elections. But does this lead to increase in informal voting in federal elections, which are conducted under compulsory preferential voting? A direct answer to this question would come via longitudinal analysis, comparing rates of informality in federal elections in those jurisdictions before and after the introduction of optional preferential voting. In my cross-sectional analysis, I test this conjecture somewhat indirectly, augmenting the regression analysis with a dummy variable for divisions in NSW, QLD and the ACT. The results are quite provocative, and are reported as "Model 2" in Table 1. The coefficients from Model 1 remain essentially unchanged. The important difference is in goodness-of-fit (an r^2 of .75 versus the r^2 of .64 in Table 1). The coefficient on the NSW/QLD/ACT dummy variable suggests that net of NESH, ballot length, and level of tertiary education, divisions in the "optional preferential" jurisdictions of NSW, QLD and the ACT recorded rates of informality averaging over a percentage point higher than elsewhere. This is a large difference, when one considers that the average level of informality across divisions is about 5.2 percent, and has a standard deviation of 1.7 percentage points. In short, my analysis does provide some support for the notion that voters are confused by the fact that two voting systems are in operation at the two levels of government.

References

Kopras, Andrew. 2004. *Electoral rankings: Census 2001 (2003 boundaries)*. Number 1 in Research Papers 2004-05. Parliament House, Canberra: Department of the Parliamentary Library. http://www.aph.gov.au/library/pubs/rp/2004-05/05RP01.htm.

McAllister, Ian, Toni Makkai and Chris Patterson. 1992. *Informal Voting in the 1987 and 1990 Australian Federal Elections*. Canberra: Australian Government Publishing Service.

Medew, Rod. 2003. "Informal Vote Survey House of Representatives - 2001 Election." Research Report 1. Australian Electoral Commission: Canberra. http://www.aec.gov.au/_content/What/publications/research_papers/paper1/index.htm.

Young, Sally. 2004. "Wasted votes? Informal voting and the 2004 election." Democratic Audit of Australia.