IZA

DISCUSSION PAPER SERIES

IZA DP No. 5854

Time to Vote?

John Gibson
Bonggeun Kim
Steven Stillman
Geua Boe-Gibson

July 2011

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Time to Vote?

John Gibson
University of Waikato
and Motu Economic Research

Bonggeun Kim
Seoul National University

Steven Stillman
University of Otago
and IZA

Geua Boe-Gibson
University of Waikato

Discussion Paper No. 5854
July 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5854
July 2011

ABSTRACT

Time to Vote?*

Despite the centrality of voting costs to the paradox of voting, little effort has been made to accurately measure these costs outside of a few spatially limited case studies. In this paper, we apply Geographic Information Systems (GIS) tools to validated national election survey data from New Zealand. We calculate distance and travel time by road from the place of residence to the nearest polling place and combine our time estimate with imputed wages for all sample members. Using this new measure of the opportunity cost of voting to predict turnout at the individual level, we find that small increases in the opportunity costs of time can have large effects in reducing voter turnout.

JEL Classification: D7, R4

Keywords: opportunity cost, paradox of voting, travel time

Corresponding author:

John Gibson
Department of Economics
University of Waikato
Private Bag 3105
Hamilton
New Zealand
E-mail: jkgibson@waikato.ac.nz

* We are grateful for the financial support from Marsden Fund grant 07-MEP-003. Stillman was employed at Motu Economic and Public Policy Research while this paper was written. All remaining errors in this paper are those of the authors.

I. Introduction

For a rational, self-interested voter, the cost of voting normally exceeds any plausible value of expected benefits from the preferred candidate winning.[1] Yet despite these calculations, many people vote. One solution to this paradox of voting, used at least since Riker and Ordeshook (1968), is to assume some additional (consumption) benefit from voting. Thus, a person will choose to vote if the costs (C) are outweighed by the benefits (B) that flow from the preferred candidate winning, discounted by the probability of casting the deciding vote (P), and allowing for the consumption benefit (D) that results from the feeling of satisfaction in performing one's civic duty:

$$C < PB + D.$$

Hence, efforts in this literature have turned to developing complex 'group-based' models that explain why voters might get a consumption benefit from voting (Feddersen, 2004).

But without clear information on the costs, C, it is difficult to know whether consumption benefits need to be large or small to salvage the rational voter model (Niemi, 1976). Accurate estimates also are needed because if costs and benefits of voting are small, the decision to vote is sensitive to small variations in either term (Aldrich, 1993); if these small effects are hard to measure, individual voter turnout will seem largely random (Matsusaka and Palda, 1999). Yet most studies struggle to measure the costs of voting at the individual level because of inadequate data and resort to using imperfect proxies. For example, Cebula and Toma (2006) use state-level median family income as their sole proxy for the opportunity cost of voting. In addition to being a crude measure, this ignores the early argument by Frey (1971) that even if people with high income face a high opportunity cost, they may participate more

[1] Feddersen (2004) considers a two-candidate election with five million voters and candidate 1's expected vote share of 50.1 percent; the benefit to a voter who prefers candidate 2 must be more than eight billion times greater than the cost of voting, in order for expected benefits to outweigh costs. Such a ratio of benefits is inconceivable for any voter, with typical stakes in an election (in terms of the amount of compensation needed for indifference between who wins) of only a few thousand dollars (Tollison and Willett, 1973).

in politics – including voting – because of greater productivity in their use of time when performing political activities.[2]

In this paper, we use validated national election unit record survey data from the New Zealand Election Study (NZES) to examine the impact of the opportunity cost of time on voting behaviour. We improve on previous work by exploiting both individual level socioeconomic information that allows us to impute wage rates for all sample members and address information that allows us to use Geographic Information Systems (GIS) tools to calculate distance and travel time by road from each individual's place of residence to the nearest polling place. Combined, this information gives a new, more accurate, measure of the opportunity cost of voting than is used in previous papers in the literature.

The NZES data have several advantages over the data used in previous studies. First, the survey data on whether each respondent voted is *validated* by comparing the data with the lists of votes cast that are held by the registrar of electors in each electorate. In contrast, many surveys used in the literature do not validate self-reports even though voting is prone to over-reporting since it is seen as a socially desirable activity.[3] Systemic overstatement of voting generates a non-random misclassification error, so when voting is the dependent variable it induces bias in all of the regression coefficients (Hausman et al, 1998).[4] Second, voter registration is compulsory in New Zealand so variation in individual turnout is not driven by differences in voter registration, as happens in the US.[5] Third, under New Zealand's Mixed Member Proportional (MMP) voting system, each voter has the same weight in determining overall proportionality of parties in the Parliament, irrespective of the marginality of the

[2] Tollison and Willett (1973) point out that predictions about whether higher or lower income individuals have a greater incentive to vote depend on both income and substitution effects, and so will be *a priori* unclear.
[3] For example, Matsusaka and Palda (1999) use surveys from four national elections in Canada where the rate of self-reported voting in their samples was 16 percentage points higher, on average, than the actual turnout in each election. Similarly, 22 percent of non-voters in local government elections in Sweden claimed to have voted when they were surveyed (Karp and Brockington, 2005).
[4] Even more non-randomness may come from the tendency for false reports of voting to be more likely for some demographic groups than others (Silver et al, 1986).
[5] Even though registration is compulsory, there is no legal compulsion to vote, unlike the case in neighbouring Australia.

electorate they live in.[6] Finally, at a general election in New Zealand, all voters have just two votes – one for a Party and one for a Member of Parliament for their electorate; referenda are extremely rare and there are never other jurisdictional positions to fill. Thus, the cognitive costs of voting should nor vary between precincts, unlike in the United States where some voters face many more choices than others.

New Zealand is also an ideal case study for using the accessibility of polling places as a measure of the opportunity cost of time facing potential voters. First, almost all resident voters cast their vote in person on election day, rather than beforehand.[7] In contrast, studies of how accessibility affects turnout in the United States have to deal with the complication that a high proportion of voters cast an early vote or vote by mail (Dyck and Gimpel, 2005). Second, Election Day is always a Saturday, so it is reasonable to assume that the average voter travels from their home to the polls, rather than from their workplace.[8] This lets us use the polling place closest to the place of residence when estimating the opportunity cost of time. Third, we can assume that little time is spent waiting at the polling booth due to congestion; relative to other countries, New Zealand appears to have ample polling places per voter. The 69 electorates in New Zealand each have an average of about 50 polling places, and at each an average of just 600 votes are cast in the 10 hours that it is open on election day.

The rest of the paper proceeds as follows. In Section 2, we review the previous literature on using accessibility to measure the cost of voting. In Section 3, we describe the survey data

[6] The exception is that a party that would not be represented in the Parliament if it got less than the threshold five percent of the party vote can still have Members of Parliament to match its vote share, if it wins at least one electorate seat. While voters in that particular electorate could have more than average influence, this situation only rarely occurs. In the empirical analysis below, the results are robust to including electorate fixed effects.

[7] In the 2005 national election studied here, less than seven percent of votes from resident voters were cast before Election Day.

[8] Evidence for this assumption also comes from the low share of Special Votes, which are for people voting outside of their home electorate on Election Day. In the 2005 election studied here, these were just 6.7 percent of total votes cast. Moreover, even for those who work on Saturdays, they must be allowed to leave their workplace no later than 3pm for the purpose of voting (or leave for at least two hours if they have essential work that goes after 3pm) and their employer cannot make deductions from the employee's remuneration for the time taken off. See Section 162 of the *Electoral Act*, 1993.

we use and how we construct the measures of distance and time costs of voting. In Section 4, the estimation strategies and empirical results are presented. The conclusions are in Section 5.

2. Previous Literature

Outside of economics there is a small but growing literature that uses Geographic Information Systems (GIS) tools to create measures of voting costs. In one of the first such studies, Gimpel and Schuknecht (2003) investigated the impact that distance to polling stations and impedance (the time and effort of the commute) had on voter turnout in three suburban Maryland counties in the 2000 Presidential election, finding a positive impact of ease of access on turnout. In related work, Dyck and Gimpel (2005) studied the choices that voters in Clark county, Nevada (where Las Vegas is located) made between voting in person on Election Day, voting early (but in person), mailing in an absentee ballot, or not voting. This research was extended to a comparison with Bernalillo county, New Mexico (wherein lies Albuquerque) by Gimpel et al. (2006).

One drawback of these early studies is that they did not use an actual road network to calculate distance (or time). Instead they used "Manhattan distance" which assumes a regular grid layout of streets between origin and destination, with no allowance for sinuosity.[9] This assumption improves upon straight-line ('crow-fly') distance but it still induces non-random measurement error that may bias regression coefficients. The non-randomness is because actual roads will sometimes follow irregular features such as rivers and broken topography, so Manhattan distance always understates actual distance. In contrast, Haspel and Knotts (2005) use GIS-based road distances to study the impact of accessibility on voting in the 2001 mayoral race in Atlanta. Specifically, they calculate the shortest path over city streets between each individual's home address and the polling place to which he or she was assigned by their optimization software (with a criterion of distance minimization). These

[9] Manhattan distance is defined as: $d_i=|(x_1-x_2)+(y_1-y_2)|$ where x, y are the longitude and latitude coordinates for the origin (1) and the destination (2).

distances were quite short, with a median value of 0.69 miles (1.1 kilometres), but they still mattered to the odds of voting. For householders with cars, the predicted odds of voting dropped from 0.464 at 0.01 miles (20 metres) to 0.385 at the median distance of 0.69 miles.[10]

More recently, political scientists have exploited natural experiments from changes in the accessibility of polling places following the consolidation of voting precincts. In the California recall election of 2003 that replaced Gray Davis with Arnold Schwarzenegger, Los Angeles County reduced the number of polling places from 5,231 used in 2002 to just 1,885. Brady and McNulty (2011) show that this increased the average distance to the nearest polling place from 0.35 to 0.50 miles, and resulted in a three percentage point decline in in-person voting in consolidated voting precincts compared to unconsolidated ones. Most of this decline is attributed to search costs of finding a new polling place rather than to the transportation cost of increased distance, and is partly offset by a rise in absentee voting. At a much finer spatial scale, McNulty et al (2009) examine the effects of consolidating eight polling places down to five in the 2006 school budget referendum in Vestal Central School District of New York state. Even among the habitual voters who participate in these local referenda, consolidation and the increase in average distance (from 0.97 to 1.30 miles) to the polls decreased voter turnout by about seven percentage points.

In summary, this literature from political science suggests that small changes in voting costs may have significant impacts on turnout. But the context for each study is spatially limited, usually considering only a single city or a few counties, so there are reasons to be concerned about the external validity of the findings. Moreover, in almost all of the studies the proxy for cost is just a measure of distance – and sometimes a potentially error-prone measure such as Manhattan distance. In no case has there been an attempt to combine the estimates of travelling time with individual-specific time valuations, so as to create an overall

[10] A drawback of their study is that the data are from voterfiles, so are restricted to those citizens actually registered to vote, which creates a potential sample selection bias, and also limits the availability of predictor variables compared with what would be available in surveys.

dollar cost for the amount of time that would be spent going to vote. There also has yet to be a nationwide study using GIS tools to measure voting costs and their impact on turnout, which is the aim of the current study.

3. Data

Since 1990, the New Zealand Election Study (NZES) has surveyed individuals of voting age (18 and over) after each general election (held every three years), using self-completion postal questionnaires and telephone survey top-ups (Vowles, 2010). As noted above, the survey self-reports of voting are validated by checking against the lists of votes cast that are held by the registrar of electors. The surveys also provide a variety of information on respondents' demographic details (age, gender, ethnicity), marital, employment and migration status, level of education, household income (in eight brackets), location and home ownership status. In addition, they include factors found in other studies to influence voting, such as participation in religious observance and prior voting history.

In this study, we use data from the 2005 general election, for which the NZES sampled just over 3,000 respondents.[11] Since the NZES is a postal survey there is no clustering used but sampling weights are formed to ensure national representativeness (hence, dealing with any uneven sampling probabilities and non-response bias). All of the analyses reported below use these weights, along with heteroskedasticity-robust measures of sampling error.

Although the NZES is a postal survey, the exact addresses of respondents are not included in the public use data, so as to maintain confidentiality. But a close approximation is provided by the survey reporting the census meshblock that each respondent's residence lies within. There are almost 40,000 mesh blocks in New Zealand, each having an average of just

[11] In the raw data files, there are almost 3,700 observations but with missing values for some of the covariates used in our models we have a sample of n=3,005. The voting rate and the average distance from the polls are the same in the full sample and the estimation sample, supporting the assumption that the observations are missing at random.

110 residents, so this is a very finely scaled spatial unit.[12] Moreover, a population-weighted centroid for every meshblock is provided by Statistics New Zealand to give even more accuracy, and that is what we use here as the best estimate of the location of the respondent's dwelling when estimating the road distance and travel time to the nearest polling place.

The names of all polling places used in the 2005 general election were obtained from the New Zealand Electoral Commission website. In most cases, these were located in schools, and civic and church halls. All locations were then geocoded using ESRI ArcGIS Geocoder with the Core Records System street address data and Google Maps to provide exact latitude and longitude coordinates. Occasionally, the Street View feature of Google Maps was used to virtually drive along roads and locate polling places that had no street numbers. For several rural locations, we also contacted key informants (such as local government officers and staff at visitor information centres) to verify the correct locations.

The algorithm within Google Maps for choosing shortest road distances was used to measure the distance between the population-weighted centroid of the meshblock that each respondent lived in and the nearest polling booth.[13] When calculating these travel times, we assume that the mode of transport is by motor vehicle. New Zealand has the second highest car ownership rate in the world (behind Luxembourg), at almost 600 cars per 1000 population. This ownership rate is much higher than would be predicted by income levels (IMF, 2008), due to the low population density and underdeveloped public transport network. It is therefore likely that the typical journey to the polls was made by motor vehicle.

In order to convert the travel time estimates into a measure of opportunity costs of time, we follow an approach used in Skinner (1987) to impute estimates for a variable that is absent from one survey, using a regression estimated from another survey that contains a richer set

[12] In terms of physical area, a typical urban mesh block would, if perfectly square, have dimensions of just over 200 metres (0.13 miles).

[13] These were batch processed using the *traveltime* command written for *Stata* by Ozimek and Miles (2011).

of variables.[14] Since the NZES has no information available on wages and earnings, we use data from the 2005 New Zealand Income Survey (NZIS),[15] which collects very detailed labour market data including wages for a sample of n=13,182 working individuals of voting age, to impute wages in NZES. While the NZES does not collect wage data, it does collect information on personal and household income in brackets, which are strong predictors of individual wage rates in the NZIS. Following Skinner, we estimate a wage regression using the NZIS controlling for all variables common to both surveys and the coefficients from this regression are then used to predict hourly wage rates for all respondents in the NZES.[16] Hence, we obtain an estimated value of time for everyone, regardless of whether they are wage workers; this is appropriate since leisure and other non-market time uses also have an implicit price that must be accounted for when studying the decision of whether to vote.

4. Empirical Analysis

4.1. Descriptive Statistics

Table 1 present means and standard deviations of the variables we use in our analysis. Because many rural areas in New Zealand are quite isolated, we present estimates throughout that both include and exclude them. The voting rate was 81.5 percent in both the national sample and the urban sub-sample (all estimates use population sampling weights). The average distance by road to the nearest polling place is 1.5 kilometres, or just under one mile, while in urban areas it is one kilometre (0.6 miles). This is comparable to the distances found in previous studies from the US. Based on this distance, the average return trip by car to the nearest polling place would take just five minutes of travelling time, while for urban areas it

[14] Skinner used a regression estimated on Consumer Expenditure Survey data to impute consumption estimates for households in the Panel Study on Income Dynamics.
[15] This survey is run as a supplement to New Zealand's main labour market survey (the Household Labour Force Survey) and hence is equivalent to the March CPS in the United States
[16] The following controls are included: age, gender, the personal and household income bracket, employment status, 2-digit industry and occupation, highest level of education, ethnicity, urbanity and region. The R-squared from this regression is 0.386 indicating that bracketed income along with other socioeconomic characteristics are strong predictors of wage rates.

would take only four minutes. Using the imputed wages, which average just under $28 per hour (US$19.50 at the time of the election), this is equivalent to a time cost of $2.30 (US$1.60) nationally and $1.80 (US$1.25) in urban areas. Obviously, allowing for parking (albeit, often free on Saturdays) and any waiting at the polls would increase this slightly, but there is no information to estimate these costs, unlike the time costs of travel.

The means of the respondent characteristics used in the model predicting turnout are also presented in Table 1. These characteristics include age, gender, and marital status, whether the respondent is an immigrant, and dummy variables for three ethnic groups (Asian and 'other' are the excluded category). Also included in the models are indicators for highest education level, for lack of religious observance, for whether employed, whether a home owner, whether in an urban area, and whether household income is in the bottom three or top three of the eight income brackets used by the survey. Finally, there is an indicator for whether respondents had voted in the local government elections from the previous year, to proxy for their attitude towards the civic duty of voting. All these variables have previously appeared in models predicting individual turnout (see e.g. Matsusaka and Palda, 1999).

4.2. Empirical Specification

We use maximum likelihood probit estimation:

$$\Pr\left(p_j = 1 \middle| \mathbf{x}_j\right) = \Phi\left(\mathbf{x}_j \beta\right) \tag{1}$$

where p_j is an indicator variable for whether the jth individual voted in the 2005 general election, Φ is the standard cumulative normal, $\mathbf{x}_j$ is the vector of explanatory variables for individual j and β is the vector of coefficients to be estimated. These probit coefficients are not directly interpretable, but marginal effects for continuous variables can be calculated (at the mean) as:

$$\left.\frac{\partial \Phi(\mathbf{xb})}{\partial x_i}\right|_{\mathbf{x} = \bar{\mathbf{x}}} = \phi(\bar{\mathbf{x}}\mathbf{b})b_i \tag{2}$$

where **b** is the vector of estimated coefficients and ϕ is the normal density. For dummy variables, the discrete change in probability when the dummy variable switches from zero to one is calculated as $\Phi(\bar{\mathbf{x}}_1\mathbf{b}) - \Phi(\bar{\mathbf{x}}_0\mathbf{b})$ where $\bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_0 = \bar{\mathbf{x}}$ except that the *i*th elements of $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_0$ are set to one and zero, respectively. Only the marginal effects (and *z*-statistics based on heteroskedasticity-robust standard errors) are reported in the tables that follow.

4.3 Estimation Results

The probit models are estimated with either distance, or time, or cost as the main independent variable, along with 16 control variables to account for potentially relevant determinants of voting that may be correlated with the cost of going to the polls. The reason for not including distance, time and cost in the same equation is because of the high correlation between these variables, which is highest for distance and time (r=0.90) but is also statistically significant for distance and cost (r=0.69). Hence, the results for each probit model illustrate different aspects of the opportunity cost of voting, while reporting all three models allows readers to interpret the role that the modelling assumptions (such as the speed of road travel and the imputation of wages) have on the results.

The results in Table 2 show that distance, time and cost are all statistically significant determinants of individual turnout. The *p*-values for the null hypotheses that these variables have no effect on the decision to vote range from 0.02 (time) to 0.04 (distance), indicating that accurately measuring aspects of voting costs can give precisely estimated effects on turnout. In terms of magnitudes, each one kilometre (0.6 mile) increase in distance to the nearest polling place reduces turnout by one percentage point; each minute of travelling time reduces turnout by one-half of a percentage point, and each one-dollar of opportunity cost reduces it by 0.8 percentage points. A standard deviation increase in any of these three variables would reduce turnout by three percentage points.

In contrast to the importance of the cost terms, most of the individual characteristics are not statistically significant determinants of the voting decision. Amongst the demographic factors, only age and ethnicity matter; voting rates go up by two percentage points for every ten years of age (the affect appears linear, since a quadratic in age was statistically insignificant), and are approximately 15 percentage points lower for the Maori and Pacific ethnic groups, although only the effect for Maori is precisely estimated. Voting rates are higher for respondents from higher income households, but the effect is imprecisely measured and only attains borderline statistical significance (p=0.07) in model (3).

The other variable that appears to matter to voting in the general election is whether the individual voted in the local government election of the previous year. The turnout for those who did vote previously is almost seven percentage points higher. One concern with including voting in the local government elections as a covariate might be that this voting also reflects the accessibility of the polling places. However, in New Zealand, all local government elections are postal ballots, so the correlation between distance to the general election polling places and whether the person voted in the local government election is zero (r=-0.005). Thus, there is no change in the magnitude of the distance, time or cost variables (and a slight increase in their statistical significance) if the previous voting behaviour is excluded from the model.

The distance to the nearest polling place varies between electorates, since the electorate boundaries are redistricted after every Census to ensure roughly equal populations but the population density is uneven over space. Since turnout also varies between electorates, one concern with the results reported in Table 2 may be that they reflect omitted variable bias. To address this, the models are re-estimated with fixed effects for each electorate (Table 3). While these electorate fixed effects are (jointly) statistically significant, and cause the pseudo-R^2 to double, they have no effect at all on the magnitude and statistical significance of the distance, time and cost variables.

Another possible concern with the results reported in Table 2 could be that the findings of significant distance effects are driven by a small proportion of rural voters who may face a longer drive to the polls.[17] In fact, the opposite is true. The impact of distance, time and cost in reducing voter turnout is stronger in urban areas (Table 4). Specifically, for the urban population, each one kilometre increase in distance to the nearest polling place reduces turnout by two percentage points; each minute of travelling time reduces turnout by 0.8 of a percentage point, and each one-dollar of opportunity cost reduces it by two percentage points. The effect of the opportunity cost of time in lowering urban turnout is very precisely estimated, with p=0.001.[18]

To illustrate how individual turnout can be affected by the opportunity costs of travelling to the polls, the models in column (3) of Table 2 and Table 4 were used to simulate turnout probability as the opportunity cost increased from $1 to $20 (US$0.70 to US$14). In these simulations all of the other variables are held at their mean, so the confidence intervals quickly expand as the simulations approach the extremes of the observed costs (Figure 1). Nevertheless, it is clear that even small increases in the opportunity cost of travelling to vote may have large effects in reducing voter turnout. For example, at an opportunity cost of $10 (two standard deviations above the national mean, and equivalent to US$7), the predicted national turnout would be just 75 percent, which is down seven percentage points from the mean. In urban areas, the decline is even larger, with a predicted turnout of just 62 percent when opportunity costs are $10.

---

[17] However, even in the rural sector the maximum distances are not especially large, with the 99th percentile of road distances to the closest polling place being 22 kilometres (14 miles). Because of this lack of geographical outliers, there are no differences in the results if logarithms of the distance and cost measures are used. Since there is an easy interpretation for units of kilometres, minutes and dollars, we only report the unlogged results.

[18] These results are also robust to including electorate fixed effects.

4.4 Non-nested tests

The two models with road distance and opportunity costs of travel time as the main independent variable of interest are non-nested, in the sense that it is not possible to impose a set of linear restrictions to derive one model from the other. But a formal comparison of these two models may help guide other researchers deciding whether the effort of imputing a wage for all sample respondents and calculating their opportunity cost of travel time is worthwhile. Forming a ‘compound’ model with both distance and costs is not advisable, both because of the potential multicollinearity problem described above and also because such an approach is a less powerful statistical procedure than are formal non-nested tests (Pesaran, 1982).

Instead, we use the Pesaran (1974) version of a Cox likelihood ratio test of the validity of one linear model, H_0 as opposed to its non-nested alternative H_1. The test can be described in general terms, as follows:

$$H_0\text{:}\ \ y = x_0 b_0 + u_0$$

$$H_1\text{:}\ \ y = x_1 b_1 + u_1$$

where x_0 and x_1 are matrices of n observations on explanatory variables that are not linear combinations of one and other, b_0 and b_1 are corresponding parameter vectors, u_0 and u_1 are random errors with zero mean and variance-covariance matrices $\sigma_0^2 I$ and $\sigma_1^2 I$ (I is an identity matrix of order n). Constructing the test statistic involves six steps, where in what follows we use the notation $M_i = I - x_i (x_i' x_i)^{-1} x_i',\ i = 0,1$:

(i) Regress y on x_0 to form $\hat{y} = x_0 \hat{b}_0$

(ii) Regress the fitted values from (i), $x_0 \hat{b}_0$, on x_1 to form residuals: $M_1 x_0 \hat{b}_0$

(iii) Calculate the sum of squared residuals from (ii), $\hat{b}_0' x_0' M_1 x_0 \hat{b}_0$

(iv) Regress the residuals from (ii) on x_0

(v) Calculate the sum of squared residuals from (iv), $\hat{b}_0' x_0' M_1 M_0 M_1 x_0 \hat{b}_0$

(vi) Calculate the test statistic $N = \sqrt{s/\hat{v}}$ which is N(0,1) under H_0, where:

$$s = (n/2)\ln\{\hat{\sigma}_1^2 / [\hat{\sigma}_0^2 + (1/n)(\hat{b}_0' x_0' M_1 x_0 \hat{b}_0)]\}$$

$$\hat{v} = (\hat{\sigma}_0^2 \hat{b}_0' x_0' M_1 M_0 M_1 x_0 \hat{b}_0) / (\hat{\sigma}_0^2 + \hat{b}_0' x_0' M_1 x_0 \hat{b}_0)^2$$

and where $\hat{\sigma}_0^2, \hat{\sigma}_1^2$ are SSE_0/n and SSE_1/n, where SSE is sum of squared errors.

The decision procedure for the test is to reject H_0 for negative (since this is a lower tail test) values of $\sqrt{s/\hat{v}}$ exceeding the critical value from the standard Normal tables. After testing H_0 versus H_1, the procedure is reversed with H_1 replacing H_0 in the above steps.

To implement the Pesaran non-nested test, the models in columns (1) and (3) of Table 4 were re-run, using Ordinary Least Squares (OLS) so that the interpretation is as a Linear Probability Model (LPM). This is because the likelihood ratio theory for the Pesaran (1974) non-nested test is developed for OLS rather than for Probit models. This change in the estimator to using an LPM rather than Probit makes almost no difference compared with what is reported in Table 4, with the point estimates on distance and cost both remaining at -0.02 and the t-statistics being approximately 0.2 lower than the *z*-statistics reported in Table 4.

When the distance-based model is H_0 it is soundly rejected against H_1, the cost-based model. The test statistic is -2.59, which is statistically significant at the $p<0.01$ level. However, when the test procedure is reversed, with H_1 replacing H_0 as the model under test, there is no rejection of the cost-based model in favour of the distance-based model. Specifically, the test statistic is only 0.34, which is not statistically significant ($p<0.37$). In other words, this direct confrontation of the two models favours the use of an opportunity cost of time variable, rather than just using road distance as a proxy. Thus, the effort to exploit individual-level socioeconomic information so as to impute wages for all sample respondents and then calculate the dollar value of their opportunity cost of travel time is supported by these results. The non-nested testing results imply that this new, more accurate measure of the opportunity cost of voting should be used instead of the simpler, distance-based proxies that have previously been used in the US literature.

5. Conclusions

In this paper, we use validated national election unit record survey data from the New Zealand Election Study (NZES) to examine the impact of the opportunity cost of time on voting behaviour. We improve on previous work by exploiting both individual level socioeconomic information that allows us to impute wage rates for all sample members and address information that allows us to use Geographic Information Systems (GIS) tools to calculate distance and travel time by road from each individual's place of residence to the nearest polling place. We demonstrate that combining this information to measure the opportunity cost of voting for each individual provides a more accurate measure of voting costs than is used in previous papers in the literature.

Our analysis finds that small increases in the opportunity costs of time have large effects in reducing voter turnout for a national election in New Zealand. For example, at an opportunity cost of $10, the predicted national turnout would be just 75 percent, which is down seven percentage points from the mean. Our findings thus confirm the conjecture first made by Niemi (1976, p.117) that "if the *B* (benefits) or *PB* (benefits weighted by the probability that a person's vote matters) term is indeed quite small, then a small increase in the cost of voting – such as driving a mile instead of a half-mile to the polls – would significantly reduce turnout." This responsiveness of turnout to small costs is also suggestive that the consumption benefits of voting need not be very large in order for it to be rational for many people to vote.

Overall, our findings are quite similar to those from Haspel and Knotts (2005), who examine the 2001 mayoral race in Atlanta, and McNulty et al. (2009), who examine a school budget referendum in Vestal Central School District of New York state. Given that we are examining a general election in a country with ample polling places per voter, weekend voting, and little road congestion, it is perhaps surprising that we find that small increases in the opportunity cost of time can discourage individuals from voting. Hence, it seems that initiatives to allow online voting and postal voting – as in more common in the US – may lead to greater voter turnout even in this quite different context to the settings that have previously been studied in the literature on the impact of voting costs on turnout.

References

Aldrich, J. (1993) "Rational choice and turnout" *American Journal of Political Science* 37: 246 -278.

Brady, H. and McNulty, J. (2011) "Turning out to vote: The costs of finding and getting to the polling place" *American Political Science Review* 105: 115-134.

Cebula, R. and Toma, M. (2006) "Determinants of geographical differentials in the voter participation rate" *Atlantic Economic Journal* 34: 33-40.

Downs, A. (1957) *An Economic Theory of Democracy*, Harper and Row, New York.

Dyck, J. and Gimpel, J. (2005) "Distance, turnout, and the convenience of voting" *Social Science Quarterly* 86: 531-548.

Feddersen, T. (2004) "Rational choice theory and the paradox of not voting" *Journal of Economic Perspectives* 18: 99-112.

Frey, B. (1971) "Why do high income people participate more in politics" *Public Choice* 11: 101-105.

Gimpel, J., Dyck, J. and Shaw, D. (2006) "Location, knowledge and time pressures in the spatial structure of convenience voting" *Electoral Studies* 25: 35-58.

Gimpel, J. and Schuknecht, J. (2003) "Political participation and the accessibility of the ballot box" *Political Geography,* 22: 471-488.

Haspel, M. and Knotts, H. (2005) "Location, location, location: Precinct placement and the costs of voting" *The Journal of Politics* 67: 560-573.

Hausman, J., Abrevaya, J., and Scott-Morton, F. (1998) "Misclassification of the dependent variable in a discrete response setting" *Journal of Econometrics* 87:239-269.

International Monetary Fund (2008) "Climate change and the global economy", Chapter 4, *World Economic Outlook: Housing and the Business Cycle* Washington DC.

Karp, J. and Brockington, D. (2005) "Social desirability and response validity: A comparative analysis of over-reporting voter turnout in five countries" *The Journal of Politics* 67: 825-840.

Matsusaka, J. and Palda, F. (1999) "Voter turnout: How much can we explain?" *Public Choice* 98: 431-446.

McNulty, J., Dowling, C., and Ariotti, M. (2009) "Driving saints to sin: How increasing the difficulty of voting dissuades even the most motivated voters" *Political Analysis* 17: 435-455.

Niemi, R. (1976) "Costs of voting and nonvoting" *Public Choice* 27: 115-119.

Ozimek, A. and Miles, D. (2011) "Stata utilities for geocoding and generating travel time and travel distance information" *The Stata Journal* 11: 106-119.

Pesaran, M. (1974) "On the general problem of model selection" *Review of Economic Studies* 41: 153-171.

Pesaran, M. (1982) "Comparison of the local power of alternative tests of non-nested regression models" *Econometrica* 50:1287-1305

Riker, W. and Ordeshook, P. (1968) "A theory of the calculus of voting" *American Political Science Review* 62: 25-42.

Silver, B., Anderson, B., and Abramson, P. (1986) "Who over-reports voting" *American Political Science Review* 80: 613-624.

Skinner, J. (1987) "A superior measure of consumption from the Panel Study of Income Dynamics" *Economics Letters* 23: 213-16.

Tollison, R. and Willett, T. (1973) "Some simple economics of voting and not voting" *Public Choice* 16: 59-71.

Vowles, J. (2010) "Electoral system change, generations, competitiveness and turnout in New Zealand, 1963–2005" *British Journal of Political Science* 40: 875-895.

Figure 1. The Effect of the Opportunity Cost of Travel Time on Predicted Turnout

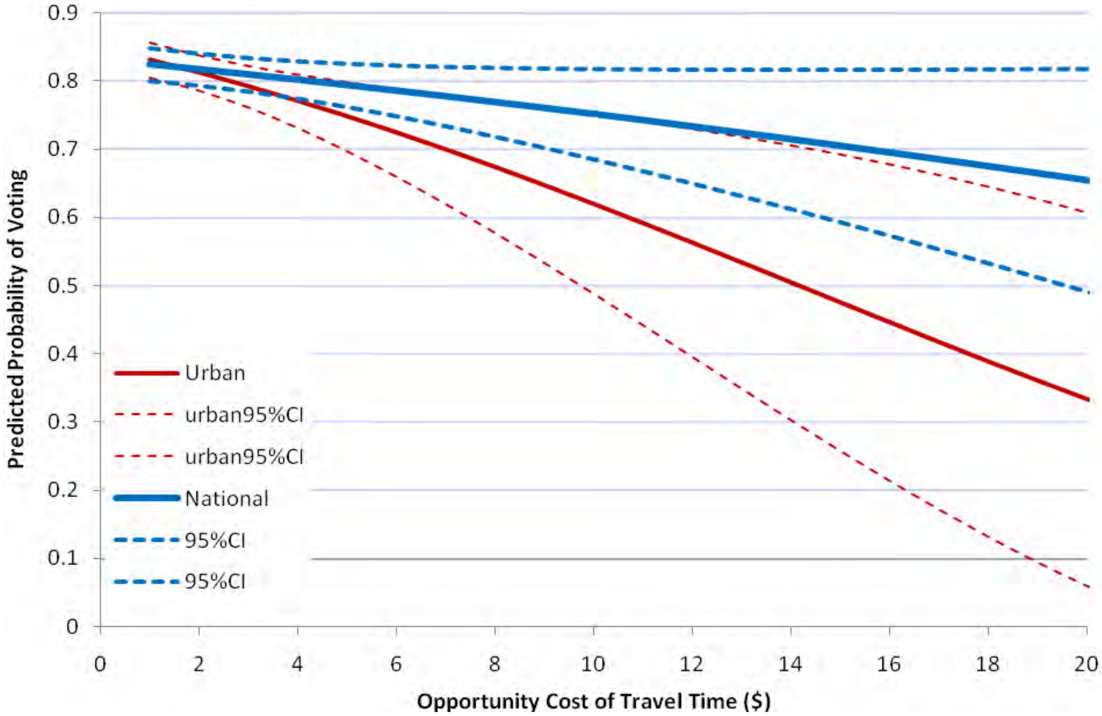

Table 1. Descriptive Statistics

	National Sample		Urban Sub-sample	
	Mean	Std Dev	Mean	Std Dev
Dependent variable = 1 if voted, 0 if not	0.815	0.389	0.815	0.389
Road distance (km) to closest polling booth	1.530	2.500	1.017	1.183
Travel time (mins) to closest polling booth	5.105	6.216	3.908	3.641
Time cost ($) of reaching polling booth	2.322	3.406	1.837	2.190
Imputed wage ($ per hour)	27.759	17.358	27.968	17.377
Age	45.547	16.155	44.804	16.216
Dummy = 1 if Female	0.483	0.500	0.476	0.500
Dummy = 1 if Married	0.682	0.466	0.662	0.473
Dummy = 1 if not born in New Zealand	0.194	0.396	0.200	0.400
Dummy = 1 if < 3 years at current address	0.113	0.316	0.118	0.322
Ethnic group is European/Pakeha	0.845	0.362	0.839	0.367
Ethnic group is Maori	0.128	0.334	0.123	0.329
Ethnic group is Pacific	0.027	0.162	0.030	0.172
Dummy = 1 if no religious observance	0.121	0.326	0.123	0.329
Highest education level is Tertiary	0.516	0.500	0.535	0.499
Employed full- or part-time	0.702	0.457	0.713	0.452
Urban resident	0.832	0.374	1.000	0.000
Homeowner	0.697	0.459	0.682	0.466
Lower household income (< $16,000)	0.063	0.242	0.063	0.243
Higher household income (> $59,000)	0.489	0.500	0.501	0.500
Voted in 2004 local govt elections	0.503	0.500	0.502	0.500
Sample size	3,005		2,452	

Notes:
Means and standard deviations are weighted by the population sampling weights.

Table 2. Marginal Effects from Probit Models of the Decision to Vote – National Sample

	(1)	(2)	(3)
Road distance (km) to closest polling booth	-0.010		
	(2.09)*		
Travel time (mins) to closest polling booth		-0.005	
		(2.37)*	
Time cost ($) of reaching polling booth			-0.008
			(2.25)*
Age	0.002	0.002	0.002
	(2.01)*	(2.01)*	(2.16)*
Dummy = 1 if Female	-0.031	-0.032	-0.024
	(1.29)	(1.33)	(0.99)
Dummy = 1 if Married	-0.012	-0.012	-0.013
	(0.41)	(0.40)	(0.43)
Dummy = 1 if not born in New Zealand	-0.054	-0.055	-0.053
	(1.49)	(1.49)	(1.46)
Dummy = 1 if < 3 years at current address	-0.034	-0.035	-0.031
	(0.83)	(0.86)	(0.78)
Ethnic group is European/Pakeha	0.034	0.036	0.035
	(0.84)	(0.90)	(0.87)
Ethnic group is Maori	-0.152	-0.150	-0.150
	(3.60)**	(3.61)**	(3.58)**
Ethnic group is Pacific	-0.154	-0.159	-0.154
	(1.76)+	(1.80)+	(1.76)+
Dummy = 1 if no religious observance	0.029	0.029	0.031
	(0.87)	(0.85)	(0.93)
Highest education level is Tertiary	0.021	0.022	0.023
	(0.79)	(0.82)	(0.89)
Employed full- or part-time	0.011	0.010	0.019
	(0.39)	(0.34)	(0.64)
Urban resident	-0.027	-0.031	-0.017
	(0.81)	(0.92)	(0.51)
Homeowner	-0.016	-0.014	-0.015
	(0.50)	(0.45)	(0.47)
Lower household income (< $16,000)	-0.010	-0.011	-0.015
	(0.20)	(0.24)	(0.31)
Higher household income (> $59,000)	0.041	0.041	0.050
	(1.50)	(1.50)	(1.80)+
Voted in 2004 local govt elections	0.065	0.065	0.064
	(2.61)**	(2.63)**	(2.55)*
Pseudo-R^2	0.061	0.063	0.061
Wald test: All slopes = 0	80.27**	80.42**	79.08**

Notes:
The coefficients give the change in the probability of voting for a one-unit change in each explanatory variable. Numbers in ( ) are z-statistics, calculated from heteroskedasticity-robust standard errors. The sample size is 3,005 and the estimates are weighted by the population sampling weights. The models all include an intercept.
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 3. Marginal Effects from Probit Models of the Decision to Vote – National Sample With Electorate Fixed Effects

	(1)	(2)	(3)
Road distance (km) to closest polling booth	-0.010		
	(2.33)*		
Travel time (mins) to closest polling booth		-0.005	
		(2.56)*	
Time cost ($) of reaching polling booth			-0.007
			(2.46)*
Age	0.002	0.002	0.002
	(1.85)+	(1.86)+	(2.03)*
Dummy = 1 if Female	-0.029	-0.030	-0.022
	(1.36)	(1.40)	(1.03)
Dummy = 1 if Married	-0.018	-0.018	-0.019
	(0.74)	(0.74)	(0.77)
Dummy = 1 if not born in New Zealand	-0.068	-0.069	-0.068
	(2.06)*	(2.07)*	(2.04)*
Dummy = 1 if < 3 years at current address	-0.037	-0.038	-0.035
	(1.05)	(1.07)	(0.99)
Ethnic group is European/Pakeha	0.035	0.036	0.036
	(0.97)	(0.98)	(0.98)
Ethnic group is Maori	-0.130	-0.126	-0.126
	(2.26)*	(2.21)*	(2.20)*
Ethnic group is Pacific	-0.121	-0.125	-0.122
	(1.55)	(1.59)	(1.55)
Dummy = 1 if no religious observance	0.031	0.030	0.033
	(1.07)	(1.03)	(1.14)
Highest education level is Tertiary	0.011	0.013	0.014
	(0.47)	(0.54)	(0.60)
Employed full- or part-time	0.014	0.012	0.021
	(0.54)	(0.49)	(0.82)
Urban resident	-0.043	-0.045	-0.033
	(1.46)	(1.54)	(1.08)
Homeowner	-0.010	-0.009	-0.009
	(0.35)	(0.31)	(0.31)
Lower household income (< $16,000)	-0.037	-0.039	-0.043
	(0.81)	(0.87)	(0.95)
Higher household income (> $59,000)	0.043	0.043	0.052
	(1.74)+	(1.76)+	(2.07)*
Voted in 2004 local govt elections	0.075	0.076	0.073
	(3.39)**	(3.43)**	(3.29)**
Electorate fixed effects	Yes	Yes	Yes
Pseudo-R^2	0.126	0.127	0.126
Wald test: All slopes = 0	196.32**	200.67**	193.28**

Notes:
The coefficients give the change in the probability of voting for a one-unit change in each explanatory variable. Numbers in ( ) are z-statistics, calculated from heteroskedasticity-robust standard errors. The sample size is 3,005 and the estimates are weighted by the population sampling weights. The models all include an intercept and 64 electorate fixed effects.
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 4. Marginal Effects from Probit Models of the Decision to Vote – Urban Sample

	(1)	(2)	(3)
Road distance (km) to closest polling booth	-0.019		
	(2.17)*		
Travel time (mins) to closest polling booth		-0.008	
		(2.57)*	
Time cost ($) of reaching polling booth			-0.020
			(3.47)**
Age	0.002	0.002	0.002
	(1.85)+	(1.89)+	(2.01)*
Dummy = 1 if Female	-0.038	-0.038	-0.023
	(1.45)	(1.44)	(0.88)
Dummy = 1 if Married	0.009	0.008	0.009
	(0.26)	(0.24)	(0.29)
Dummy = 1 if not born in New Zealand	-0.045	-0.044	-0.047
	(1.12)	(1.11)	(1.16)
Dummy = 1 if < 3 years at current address	-0.039	-0.038	-0.034
	(0.87)	(0.87)	(0.78)
Ethnic group is European/Pakeha	0.051	0.054	0.054
	(1.16)	(1.22)	(1.23)
Ethnic group is Maori	-0.155	-0.152	-0.154
	(3.43)**	(3.44)**	(3.47)**
Ethnic group is Pacific	-0.171	-0.180	-0.175
	(1.90)+	(1.98)*	(1.93)+
Dummy = 1 if no religious observance	0.017	0.018	0.019
	(0.45)	(0.48)	(0.51)
Highest education level is Tertiary	0.030	0.029	0.035
	(1.02)	(1.00)	(1.22)
Employed full- or part-time	0.001	0.003	0.016
	(0.02)	(0.09)	(0.51)
Homeowner	-0.011	-0.011	-0.007
	(0.32)	(0.31)	(0.20)
Lower household income (< $16,000)	-0.036	-0.037	-0.039
	(0.68)	(0.70)	(0.75)
Higher household income (> $59,000)	0.043	0.042	0.062
	(1.42)	(1.39)	(1.96)+
Voted in 2004 local govt elections	0.058	0.059	0.058
	(2.13)*	(2.18)*	(2.16)*
Pseudo-R^2	0.074	0.076	0.083
Wald test: All slopes = 0	81.92**	82.09**	84.79**

Notes:
The coefficients give the change in the probability of voting for a one-unit change in each explanatory variable. Numbers in ( ) are z-statistics, calculated from heteroskedasticity-robust standard errors. The sample size is 2,452 and the estimates are weighted by the population sampling weights. The models all include an intercept.
+ significant at 10%; * significant at 5%; ** significant at 1%