Do Higher Housing Values Make Communities More Conservative? Evidence from the Introduction of E-ZPass

Connor Jerzak[1] and Brian Libgober[2]

[1] Department of Government and Institute for Quantitative Social Science, Harvard University, 1737 Cambridge Street, Cambridge MA 02138
e-mail: cjerzak@g.harvard.edu
[2] Department of Government and Institute for Quantitative Social Science, Harvard University, 1737 Cambridge Street, Cambridge MA 02138
e-mail: blibgober@g.harvard.edu

June 22, 2016

Abstract

A rich literature exists on the extent to which homeownership is central to American political attitudes. This paper uses the introduction of E-ZPass in Pennsylvania and New Jersey to identify the effect of traffic-reducing transportation infrastructure on property values and, in turn, political behaviors. We develop a model showing that faster travel times results in individuals preferring a lower tax rate, as those who face the lower travel times are made effectively wealthier. Next, we present empirical evidence consistent with this theoretical result. We show that voting precincts near newly introduced E-ZPass toll plazas experienced a sharp increase in property values relative to similar precincts near non-E-ZPass exists, giving us leverage to identify the causal effect of property value changes on voting. We find that positive shocks in property values are associated with an increase in Republican vote share.

1 Introduction

2 Relationship to the Literature

At least since the 1940s, social scientists have relied on survey research to understand the factors affecting voting behavior, yet only recently have scholars begun to use experimental and observational evidence to gain "causal leverage for analyses of voting behavior" (Bartels 2010). Early studies of voting behavior established that although individual vote choice is susceptible to election-year specific shocks (such as the presence of an unusually popular presidential candidate or war-time discontent), the way an individual votes in most elections is for the most part a function of their personal identity characteristics (Converse 1966). According to this line of research, the most significant determinants of long-term voting behavior include party identification, ethnicity, gender, age, religion, education, and occupation (Lazarsfeld et al. 1948; Berelson et al. 1954; Campbell et al. 1960; Stanley and Niemi 2006). Noteworthy too are the factors these studies did *not* find significant for long-term voting behavior: preferences about political issues, for example, and economic self-interest. This is not to say that financial well-being was ever considered irrelevant for voting behavior; the literature is replete with studies demonstrating that the success of incumbents is tied to that of the economy (Tufte 1975; Meltzer and Vellrath 1975; Hibbs 1987). Yet this effect appears to be based largely on "sociotropic" evaluations of national economic performance, not personal economic experience,

Preliminary. The authors thank Jeffry Frieden and Kenneth Shepsle for insightful comments. The authors names' are listed alphabetically, and the authors contributed equally to all aspects of the analysis. The usual disclaimers apply.

and evidence for the latter's effect on voting behavior is at best mixed, even in the short term(Linn et al. 2010). For early survey research, long-term voting behavior was for the most part reducible to an individual's demographic characteristics, and would not predict that increasing an individual's wealth or income would ipso facto change their vote choice. This view is in considerable tension with formal models whose comparative statics expect such a relationship between an indivdual's economic situation and their voting behavior (e.g., Romer 1975; Roberts 1977; Meltzer and Richard 1981).

Although the conclusions of survey research just described have been replicated numerous times (Nie et al. 1979; Smith 1989; Lewis-Beck et al. 2008), in the past decade these results have nonetheless faced significant challenges from scholars bringing to bear new data sources and applying new data analysis techniques. In particular, ideology and economic interest are gaining renewed recognition as determinants of life-long voting behavior. Ansolabehere et al. (2008) show that the apparent incoherence of individual issue attitudes may largely be a result of measurement error, and that by averaging multiple question responses one gets a more stable estimate of individual issue attitudes, which become nearly as predictive of voting behavior as partisan identification. Gelman et al. (2007) use exit poll data to establish that income can be an important determiner of individual vote behavior, but its predictive power depends greatly on geography: in poor states like Alabama those who make little vote very differently from those who make a lot, while in rich states like Connecticut the difference is barely present. The primary innovation of both these papers is that they make compelling arguments for the use of data analysis techniques not normally used within the strand of survey research concerned with voting behavior. Other recent papers have proposed to use entirely different data sources to explore questions about voting behavior that would have previously been answered with survey data alone. Hersh and Nall (2015), building on the results of Gelman et al. (2007), use highly disaggregated registration, census, and election returns to show that income is a significant determiner of voting behavior only in Congressional districts with large minority populations. Gerber et al. (2011) found a significant but ephemeral effect for televsion advertising by conducting a randomized experiment involving a $2,000,000 ad-buy in a Texas gubernatorial race. Arunchalam and Watson use height as an instrument for studying the effect of income on support for the Conservative Party in the United Kingdom. Ansell (2014) looks at ANES survey data from individuals living in Metropolitan Statistical Areas for which the Federal Housing Adminstration collects data on home values to calculate the effect of home value appreciation on attitudes toward a redistributive program. Ansell's substantive results, that increasing property values makes individuals less likely to support redistribution, nicely compliment ours, especially since we do not rely on any of the same data sources and uses a different identification strategy. This article continues in the vein of these most recent papers, using data sources and methods not frequently used in the voting behavior literature to support the claim that economic experience can have a significant effect on voting behavior.

3 Research Design

The self-interested voter hypothesis suggests that support for conservative policies should increase following house price appreciation. In what follows, we show that reductions in traffic plausibly increase housing values and, in turn, conservative support. For further details on why traffic reductions should, all else equal, increase housing values, see the appendix. The basic intuition is that reductions in traffic make nearby communities more desirable, increasing the bargaining power of residents and driving up housing values. Our empirical approach to addressing this relationship exploits the fact that the introduction of E-ZPass in Pennsylvania and New Jersey resulted in decreases in traffic in some (but not all) parts of the states. In addition, E-ZPass plazas were not selected endogenously at the time of introduction, but replaced already existing toll structures.[1] We propose IV and conditional difference-in-difference estimators to evaluate the change in political attitudes between precincts that were exposed to E-ZPass (and that thus experienced a sharp rise in property values) and those that were not (Donald and Lang 2007).

For this study, we rely on a combination of low-level voting, political contribution, census and geographic data. All geographic analysis, in particular data about the relative distance of two locations, were conducted using ArcGIS with a national highway map provided by the developer of ArcGIS. The location of EZ-Pass tolling booths were taken from the replication dataset to **Currie2011** and this data was replicated and

[1]We later perform a series of robustness checks to ensure that the initial (endogenous) placement of the tolls does not confound our inferences.

supplemented with data collected from Department of Transportation websites. Precinct-level data on the number of registered voters and the number of votes received by each party in Presidential elections, as well as shape files detailing the geography location of each precinct, were taken from Ansolabehere et al. (2014). Contribution data collected by the Federal Election Commission (FEC) was reported at the individual level, however this data is reported by zipcode. Census data used for matching and for constructing our instrumental variable were taken from the 2000 decenial census and the 2005-2009 American Community Survey (ACS).

Because these data are reported at different levels of aggregation, substantial effort was required to create a dataset suitable for analysis. Formally, we consider an observational unit in our study to be the voting precinct. Voting precincts are contiguous areas, typically about 5 square miles in area, and are roughly the same size as a census block group, the smallest area at which census data are reported. Zipcodes are usually much larger areas containing multiple precincts or census block groups. Since the boundaries of precincts, census block groups, and zipcodes are generally not identical, we use areal interpolation to impute the data from these other sources to precincts. ArcPython replication code is provided for those interested in seeing exactly how each column in our dataset was constructed, however the basic idea is no more complicated than a weighted average based on the amount of overlapping area between the precinct and the geographic area containing relevant data. Also crucial to our analysis is the distance of each precinct to the highway. For this purpose we consider the polling place as coded in Ansolabehere et al. (2014) to be the precinct's location, and we operationalize distance to the highway to be the network or "over-road distance" to the nearest highway entrance.

This analysis must address a key issue around internal validity. That is, to obtain any causal estimate, it is first necessary to identify treatment and control groups. However, studies that use geographic boundaries as identification mechanisms are open to criticism around how the treatment and control groups are specified. If the treatment and control groups can be defined arbitrarily, then a study may be vulnerable to data mining problems. As a consequence, this analysis uses two definitions of treated and control groups in order to illustrate that our results are robust to these two specifications.

Our analysis first uses a conditional difference-in-differences model with matching. This method takes as the treated group those precincts living close to an E-ZPass exit toll plaza, who therefore are more likely to have received an exogenous increase in their housing price. Whether "close" should mean 5, 10, or 15 miles is unclear *a priori*, and we provide estimates under a plausible range of values. For the distance calculations, we use a network distance optimized distance measure (i.e. the minimum distance one would have to drive from the precinct center to the nearest E-ZPass plaza). We must also define a reasonable control group that could have received a reduction in traffic via E-ZPass but did not. To construct this control group, we examine precincts close to exits on major highways without E-ZPass tolls. However, some precincts are both within 10 miles of an E-ZPass and 10 miles of an exit to a highway without E-ZPass. Thus, in order to get genuine separation of treatment and control groups, we create a rule excluding those precincts that are in one testing group but are on the cusp of being on the other. Clearly, the exclusion rule should have a radius at least as big as the inclusion rule to get true separation. But one should also be concerned that citizens may be willing to drive further to take a non-toll road than one with tolls (i.e. in citizens' everyday experience, perceived "closeness" to an E-ZPass exit may differ from perceived "closeness" to a non-E-Z-Pass exit). As a result, the radius of the exclusion rule should, according to this line of reasoning, be larger than the radius of the inclusion rule. At the same time, if it is too large, one will then exclude too many units, particularly those in suburban areas where many different highways intersect. While, in principle, treatment and control groups could each have their own inclusion and exclusion rules, we shall assume that treatment and control each have the same rule for inclusion and the same rule for exclusion. This approach has the benefit of consistency. Moreover, rather than presenting a single, arbitrarily chosen difference-and-difference estimate, we provide estimates for our main effect under a plausible grid of inclusion and exclusion values.

Our analysis next uses an instrumental variables (IV) approach. This method does not posit the existence of a separate treated and control group. Rather, we use distance to an E-ZPass toll plaza as an instrument for housing price appreciation for those precincts within some fixed distance from the toll plaza. This framework depends on two assumptions. First, distance to E-ZPass toll plaza must be correlated with the endogenous explanatory variable (i.e. gain in housing price). This correlation is strong (with a t-statistic of over 20). Next, the instrument cannot be correlated with the error term of our explanatory equation predicting change in Democratic 2-party vote share. This assumption would be violated if distance to

E-ZPass toll plazas affected Democratic 2-party vote share even when housing prices are kept constant.

In what follows, we show that our results are robust to the specification of the treated and control groups, as well as to the method of inference. We have a slight preference for the conditional difference-in-difference approach because this specification of treated and control groups seems more natural. Areas 2 vs. 10 km from an E-ZPass toll plaza may be more systematically different from each other than precincts near E-ZPass toll plazas and those near non-E-ZPass traffic exits. Even so, it is important to give careful thought to the specification of the treated and control groups in geographically-based analyses, and it is best to show the robustness of a result across a range of specifications. We also take care to show that other important predictors of Democratic support (such as race and income) do not undergo significant changes following the introduction of E-ZPass.

4 Robustness Check

5 Conclusion