Disability and Convenience Voting: Evidence from the 2020 Election Cycle

Abstract

Disability is usually a missing feature of studies of voting behavior. Estimates of structural equation models from an original survey with verified vote demonstrate the importance of including disability in models of voting and taking into account the complex relationship of disability with other prominent influences on the vote. Confirming other studies, persons with disabilities and impairments are shown to have voted less frequently than other Americans in the 2020 national elections. However, it the estimates reported here convenience voting methods adopted by the states shrink or eliminate the difference in propensity to vote between persons with disabilities and others. These findings improve on previous studies by a more appropriate modeling strategy, incorporating voting methods, and using verified vote. The findings call the inclusion of disability and voting methods as a more central features of our studies of political behavior.

Word count:

Our understanding of how life circumstances influence the likelihood of voting is expanding, but only gradually. Beyond gender, race, and education “gaps,” social scientists have long been interested in other circumstances, traits, and conditions produce higher or lower rates of voting. In this paper, we draw attention to the role of disabilities and the barriers to voting.

Remarkably, there is *no* mention of persons with disabilities in recent major studies of who votes, who does not vote, and inequality in participation. This is true for recent books entitled “The Turnout Gap,” “Who Votes Now,” and “The American Nonvoter” (Fraga 2018, Leighley and Nagler 2014, Ragsdale and Rusk 2017). The subject of “inequality,” which features in the subtitles in two of those studies, surely was not intended to exclude persons with disabilities, but the analysis certainly did. And yet many of the economic, psychological, social, and even political explanations for gender, race, ethnicity, and social class inequalities in participation in democratic processes apply to people with disabilities (PWD).

This is not a minor missing variable. There are at least as many Americans with a disability as there are Latinx Americans; there are considerably more Americans with a disability than there are black Americans. In our survey (detailed below), nearly 18 percent of respondents indicated that they are limited in their “ability to work at a job, do housework, or go to school because of an impairment or a physical or mental health problem.” This is nearly identical to U.S. Census Bureau findings for “severe” disability for non-institutionalized civilians (Taylor 2018)—55.2 million people was the census estimate for 2014. As the census analysis points out, these estimates are likely to be low. Most obviously missing is that another 1-2 million people live in assisted living, nearly all with disabilities, and are not included in household surveys. Just after the 2020 election, 25 percent of Americans reported having “a health problem, disability, or handicap” that kept them “from participating fully in work, school, housework, or other activities” and about 12 percent of nonvoters reported that sickness or disability was a reason for not voting (Stewart 2021). Moreover, a survey with a large over-sample of PWD found that about twice as many voters with a disability than other voters had difficulty voting, even with a mail-in ballot (Shur and Kruse 2021).

In this report, we provide a new look at the role of disability by examining voting turnout in the 2020 U.S. general election. We contribute to the few existing studies in several ways. First, we start with the observation that the effects of disability on voting participation are likely to be complex. This leads us to estimate a structural equation model that allows us to examine the effects of disability in the context of other factors that may influence voting and are influenced by disability. Second, we confirm the existence of a sizable “disability gap,” show the direct and indirect effects of disability on voting, and compare our survey estimates with those of the Current Population Survey (CPS). Third, we provide a tentative estimate of the effect of “convenience” voting methods, widely expanded in 2020 during the pandemic, on disability effects. Fourth, unlike the CPS, our survey includes measures of key political attitudes to assess whether they serve as distinctive intermediate factors for disability effects and allows us to use a verified vote to reduce measurement error on the dependent variable. We conclude with the suggestion that future national surveys address the role of disability in political participation.

We would be remiss if we did not mention an important motivation for our study. Efforts in many states to limit convenience voting—entirely for reasons having little to do with accessibility for persons with disabilities—may have serious unanticipated or deliberately ignored consequences for persons with disabilities. The gains in accessibility in 2020, a by-product of the pandemic, may be lost and affect civic participation by people who struggle to vote for years to come (Shur and Kruse 2021). Understanding the role of disability in voting should be a high priority for advocates of broad participation in American democracy.

Previous Studies

There is a long history of neglect, bias, active discrimination, and institutionalization that marginalized persons with disabilities and discouraged or made impossible their participation in civic life (Braddock and Parish 2001; Fleischer and Zames 2001). As a result, disability has not been a salient feature of American election campaigns in the way that race, class, and gender have been in recent decades. Indeed, federal protection in the form of the Americans with Disabilities Act (ADA) came long after the civil rights era and congressional action on the Equal Rights Amendment (Scotch 2001; Switzer 2003). Over the decades, however, a large advocacy community has emerged that represents persons with disabilities in policy debates on employment, education, civic participation, and other issues.

Political scientists have seldom included disability as a demographic or background variable in studies of political attitudes and behavior. By ignoring disability, we may be generating misleading estimates for the influence of factors that are correlated with voting and having a disability. On average, people with disabilities are older, more likely to be unmarried, have less formal education, and less likely to be Latinx (Shur and Kruse 2021). While age and race are exogenous to disability and marital status and education are likely to be endogenous to disability, all have been shown to affect the likelihood of voting. Age, particularly, is known to be a strong correlate of both disability and voting, increasing the probability of having a disability and yet increasing the probability of voting.

Fortunately, a few path-breaking social scientists have given persons with disabilities more attention than the core political science literature on voting turnout would suggest. In general, these studies add disability to a set of demographic variables that are shown to predict voting (for a review, see Shur, Ameri, and Adya 2017). Disability is always shown to have an effect on voting participation, even when other factors, such as resources, are taken into account. The size of the effect is substantial—typically, an effect larger than race or gender.

Understandably, most studies have relied upon the Current Population Survey, which provides a large national sample that allows analysis of many subgroups with at least some statistical power. Missing from CPS-based studies is an effort to break open the black box of economic, psychological, social, and political factors that may be associated with disability and impairment. The CPS does not include measures of political attitudes that have been shown to mediate the effects of life circumstances on voting participation. Because persons with disabilities often suffer from impairments that are barriers to participation, disabilities of some kinds may have both direct effects on participation while disabilities may have an indirect effect on participation by having attitudinal implications that influence decisions to participate.

Shur and Adya (2013) and Shur, Ameri, and Adya (2017) are the most sophisticated studies. They and others have recognized these pathways and, in some cases, have estimated some of the pathways in a set of reduced or nested models. Others simply add many partially-justified variables a controls. In Johnson and Powell (2020), for example, demographic and political attitude variables are treated as control variables and their estimated effects are not reported in the body of the paper. The study reports that having a disability is an influence on political attachment, but the pathways from demographic factors and disability through political attitudes to voting is not considered. These reduced models that do not account for the causal relationships between their independent variables and therefore cannot give us a picture of the direct and indirect effects of having a disability or impairment on the likelihood of voting.

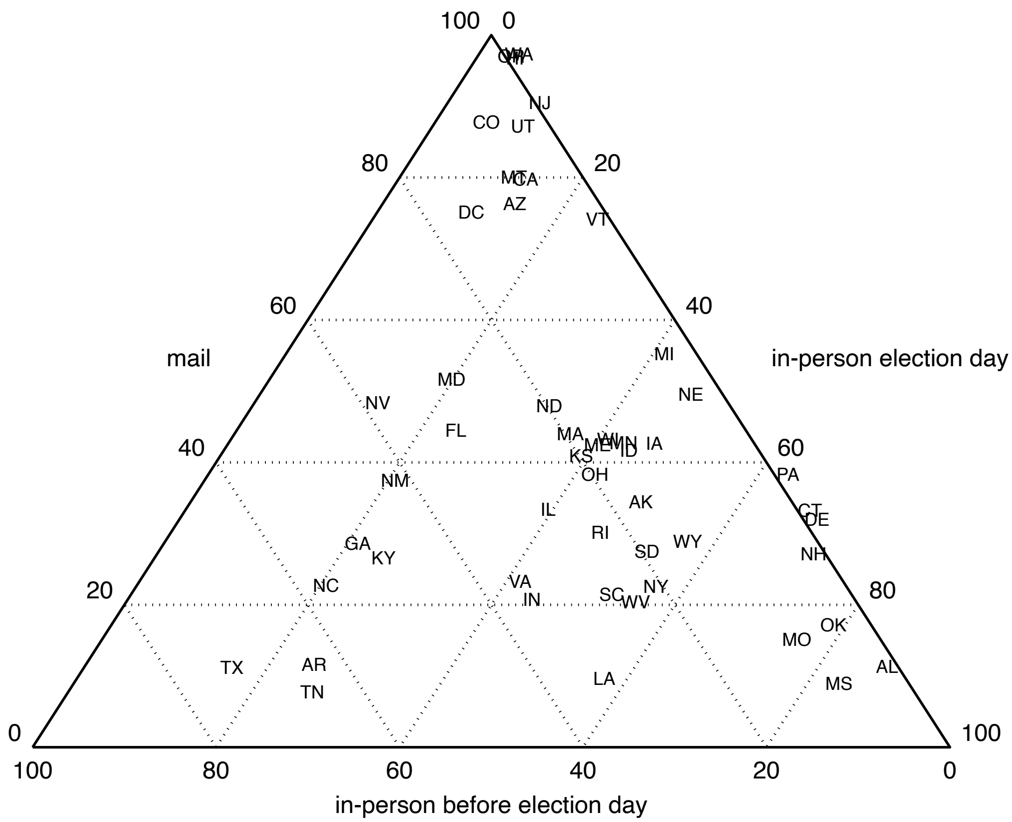
Generally, missing in these studies is serious consideration of attitudinal implications of disability for political participation. A Pew Research Center study found only small differences between people with and without disabilities in giving thought about the 2016 election, following the campaign, and thinking that who wins really matters (Igielnik 2016). A study drawn from the European Social Survey found lower levels of efficacy and political interest, associated with lower political participation, among people with disabilities than among other Europeans (Reher 2020). A study limited to New Mexico found some differences related to disability, but concluded by calling for a national study (Gastil 2000). The unexamined question is how those attitudes contribute to voting once we account for other implications of having a disability or impairtment. We examine that question here.

The 2020 Election Cycle

The 2020 election cycle is of particular interest in the study of voting and disability. In response to the COVID-19 crisis in 2020, most states took steps to facilitate voting by mail and early in-person voting. Early in-person voting and voting by mail or absentee ballot had been increasing gradually in the early 21st century. These “convenience voting” methods were used by about 14 percent of voters in 2000 and 40 percent in 2016. In 2020, primarily as an effort by states and localities to avoid close contact among voters and election workers, that number shot up to 72 percent, with 46 percent voting by mail or absentee ballot and 26 voting early in person (Stewart, no date).

States varied widely on the extent to which they instituted convenience voting methods, usually under authority granted 0r assumed by executive branch officials, or gave localities discretion to facilitate early voting. The result was wide variety across the states in the voting method used by voters. The CPS, which includes a Voting and Registration Supplement in November following federal elections, provides a large-sample look at voting experience in the 50 states and the District of Columbia (Census Bureau 2021). The 2020 supplement’s estimates are shown in Figure 1, a triplot of the three major voting methods.

Figure 1. Voting Methods in 2020, by State,
in Percent Reporting Voting by Each Method



Source: CPS Voting and Registration Supplement, November 2020.

The variation has some regional patterns. Southern states did not expand opportunities to vote by mail in some cases but liberalized early voting. States in the west (California, Colorado, Hawaii, Oregon, Utah) and a few others either already had easy or universal voting by mail or adopted it in 2020. Most other states had expanded early voting and voting by mail systems in place that produced little in-person voting on election day.

These developments reduced obstacles to voting for many people with disabilities and probably increased participation disproportionately in that group. Nevertheless, the “disability gap” remained. The (weighted) difference between PWDs and other Americans, not adjusted for age or other considerations, in the 2020 CPS Supplement (see below) was about 9 percent. Our findings may serve as a warning about the effect of retraction in convenience voting methods, which is a subject now being addressed by activists (Katz 2021).

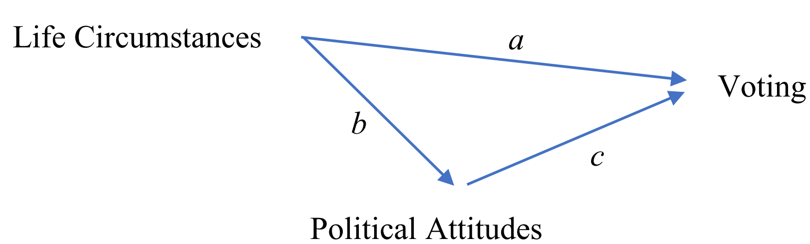
Methods, Measures, and Hypotheses

We report our analysis in two parts. The first part is a structural equation modeling exercise to demonstrate how the addition of disability contributes to our understanding of voting. The second part focuses on the importance of convenience voting for improving turnout for persons with disabilities.

Structural Relationships

We adopt a general theory of the vote common to other studies. Specifically, life circumstances, usually captured by demographic variables, influence political attitudes and experiences, which, along with life circumstances, determine the probability of voting. The causal paths are depicted in Figure 2.

Figure 2. A General Theory of Voting



A typical model of path *a* is one proposed and estimated in Leighley and Nagler (2014). Using the large samples of the CPS, they find that a higher income, higher education, higher age, being Anglo rather than Latino, being female rather than male, and being married have a positive effect on the probability of voting in recent election cycles, even in multivariate estimates. Black vs Anglo is not a difference that produced significant different in voting participation, which we confirm. Like others, they emphasize the importance of education over income as a predictor of voting, which, it is argued, reflects the importance the skills and information in motivations to vote (Wolfinger and Rosenstone 1980). We replicate this analysis as our starting point.

We then add disability to the set of life circumstances variables and posit relationships between disability and the other life circumstance variables. Our hypotheses are straightforward: Age and race are exogenous to disability; marital status, income, and education are endogenous to disability. In addition to the previously identified correlates of votes, we expect disability to have an independent effect on voting. Importantly, however, previous research does not provide clear guidance about the strength of the direct and indirect effects of disability, age, education, and other key variables. That is left for us to explore.

As a third step, we consider how disability and other life circumstances influence political attitudes and traits (path *b*) that, in turn, are likely to factor in the decision to vote (path *c*). In particular, we consider the role of political information, the perception that the election makes a difference, and strength of partisanship and ideology as variables that intervene between disability and voting, factors that are similar or identical to those examined in many studies (see Ragsdale and Rusk 2017, 16-25, 91-125). The key question is whether disability influences voting through intervening political attitudes or has other paths to voting, if they exist.

The Conditional Effect of Ease of Voting

Disability, we confirm below, has an effect on propensity to vote (Shur and Kruse 2021). We also know that convenience voting methods increase turnout (Leighly and Nagler 2014). A remaining issue is whether convenience voting is of disproportionate importance for persons with disabilities. The voting methods vary systematically by state so we pursue of strategy of estimating the cross-sectional interaction effects of state voting practices and disability on voting.

Characterizing state voting practices is a not as straightforward as categorizing state voting laws. Variation among states is in the legal authority behind voting methods and variation within states in the implementation of convenience voting methods makes a simple classification problematic. In Figure 1, we demonstrated state practices on the basis of CPS estimates of the method of voting used by its respondents. For the purpose of characterizing "ease of voting" by state, we use a single index created from state registration and voting requirements as proposed by

Measures: Verified Vote, Disability, and Other Variables

We draw upon a survey conducted for us by NORC-AmeriSpeak (details in the appendix), called The American Social Survey (TASS), and the CPS 2020 Supplement. TASS has about 1,350 respondents, while the CPS post-election supplement has over 69,000 respondents and a correspondingly larger number of PWDs. TASS has the advantage of including questions that tap political attitudes often associated with voting and therefore permits estimates of a more complete model. TASS also uses a better question to determine disability status and allows us to determine verified vote to reduce measurement error. In estimates of the interaction effects of disability and convenience voting on propensity to vote, we report estimates for both surveys.

Whether or not a respondent voted is our primary dependent variable. In our survey, we asked whether the respondent voted in the November 2020 election, but we use verified vote in this report. The verification process used state voter data collected by Aristotle and processed by NORC-AmeriSpeak. The matching process is summarized in the appendix.

Determining disability status relies on self-reports. The CPS asks a battery of questions about whether the respondent "has difficulty" hearing, seeing, walking, and so on. We count a respondent as a PWD when there is "yes" response to any of these questions. In our survey, we ask about type of disability, but first ask, "Are you limited in any way in your ability to work at a job, do housework, or go to school because of an impairment or a physical or mental health problem?" We use responses to this question as our measure, which is similar to those used in other studies (Shur and Kruse 2021).

The other variables and their measurement are summarized in Table 1. All estimates are weighted by CPS benchmarks.

Insert table on measures

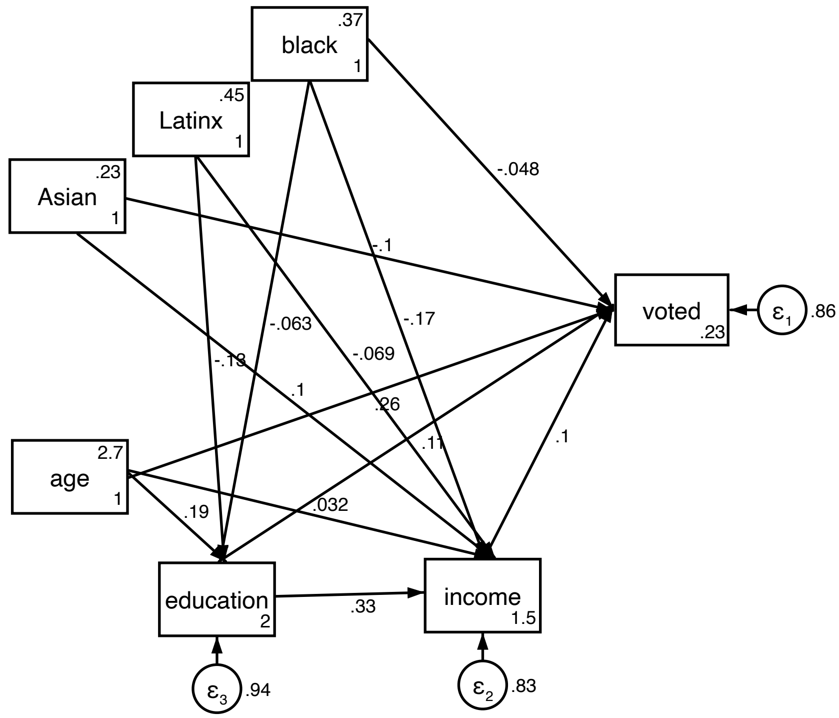
Findings

Structural Equation Model Estimates

To begin, Figure 2 presents estimates for a generalized structural equation model for standard demographic variables and a binary voter/nonvoter outcome variable, with standardized coefficients displayed to facilitate path comparisons. Here and elsewhere, the full estimates are reported in Appendix 2. All relationships are in the expected direction, but only the paths with a p-value < 0.10 are shown in the figure. There is very good overall fit. There are several highlights to note:

- The effects of being black or Latino, in comparison to being non-Latino white, on the vote are heavily through their negative effects on education and income.
- Age has a positive effect on voting, both directly and through education.
- Both education and income have an effect on voting, but the total effect of education is much larger because of both its direct effect and indirect effect through income.

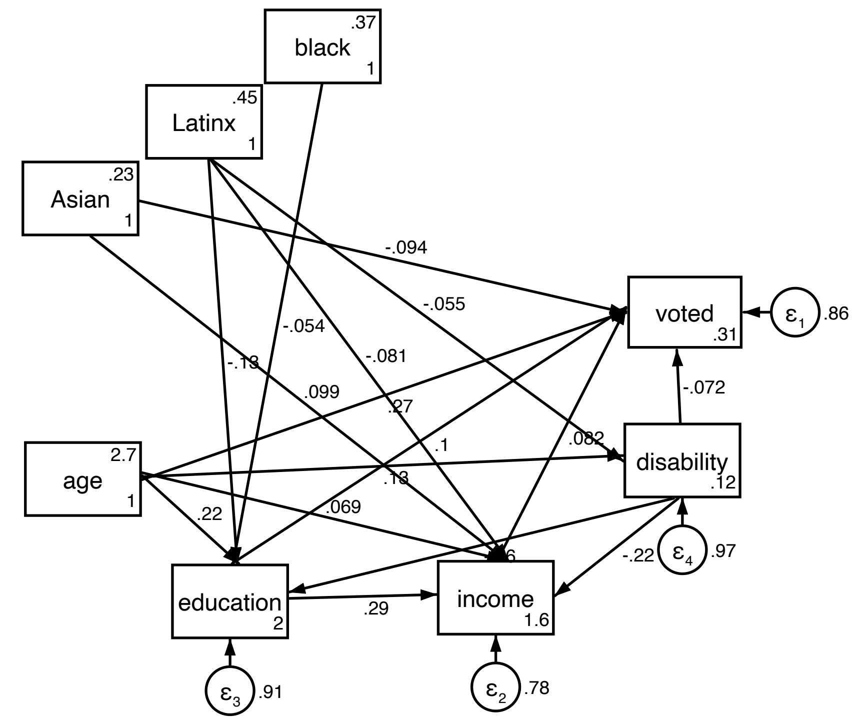
Figure 2. Structural Equation Model 1: Demographic Variables.



Note: SRMR = 0.014. Maximum likelihood estimate; standardized coefficients shown. e is error variance; lower right values are variances; upper left values are standardized means. Weighted.

Remarkably, having a disability reduces the probability of voting in 2020 but not to a statistically significant degree in our survey. Obviously, the small subsample of persons with disability in our national sample of about 1,300 for the post-election November wave undermine finding relationships that are not large. However, taking into account just one strong predictor of having a disability—age—produces a “disability gap” of 7.5 percent. This long-appreciated relationship has led some investigators to routinely report an “age-adjusted” disability gap. We choose to report a more complete model by adding disability to Model 1, as we do in Model 2 (Figure 3).

Figure 3. Structural Equation Model 2: Demographic Variables with Disability.



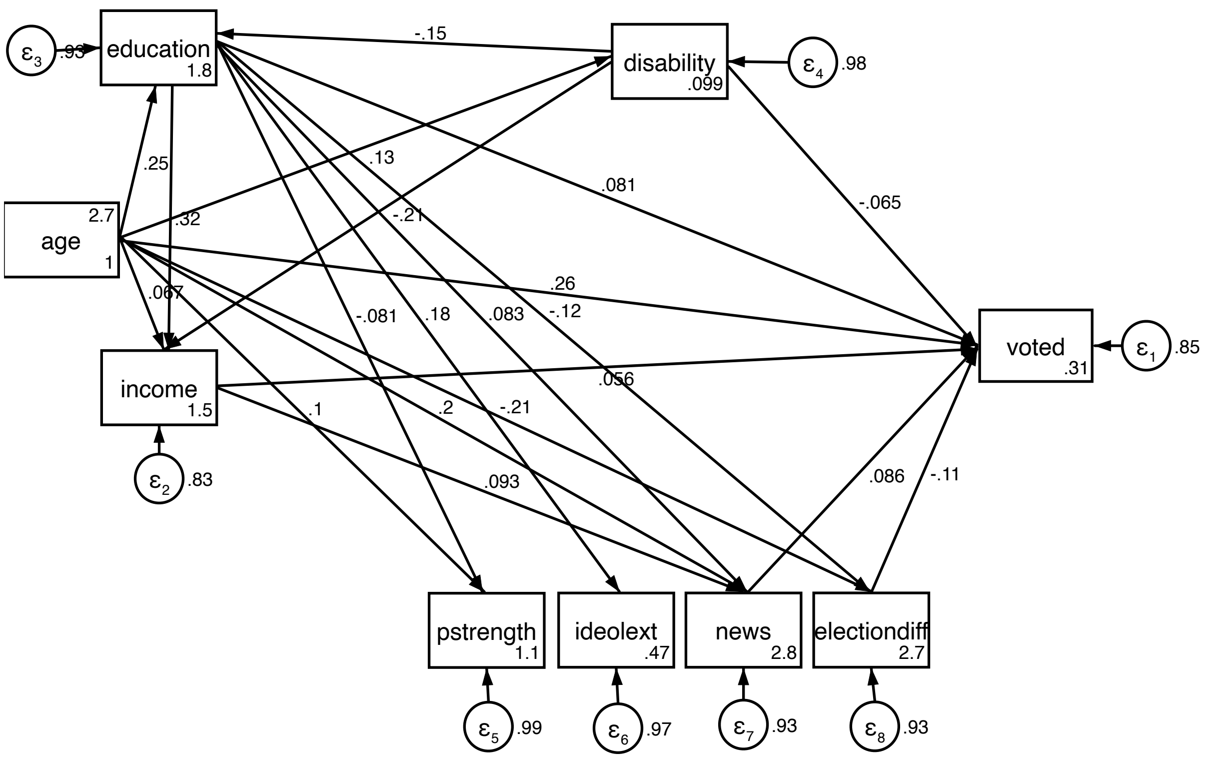
Note: SRMR = 0.010. Maximum likelihood estimate; standardized coefficients shown. e is error variance; lower right values are variances; upper left values are standardized means. Weighted.

Model 2 estimates lead to two important inferences. First, a small but measurable part of the effect of age on voting is due to the effect of age and the probability of having a disability. While age is positively related to voting, which is usually attributed to political awareness and caring about the outcome increasing with age, disability among the elderly tends to suppress voting. Because of the number of people involved, the former effect is much stronger than the latter, but studies that fail to account for disability are likely generating inaccurate estimates of the direct effect of age.

Second, disability has a very modest direct effect on voting. Much of the disability effect on voting is through education and income. In fact, based on the strength of the path coefficients, the 7.5 percent net effect of disability is split about equally between a direct effect and the indirect effects through education and income. By failing to account for disability, the nature of the education and income effects on voting cannot be fully understood.

Model 3 adds variables tapping political attitudes and behavior that are often associated with voting—strength of partisanship, ideological extremism, use of the news, and perception that the election will make a difference. To facilitate the presentation, we exclude the exogenous variables associated with race and ethnicity and others with little or no relationship to disability. Again, all relationships are in the expected direction, but, for ease of view, only the paths with a p-value $<$ 0.10 are shown in Figure 4.

Figure 4. Structural Equation Model 3: Addition of Attitudinal Variables.



Note: SRMR = 0.067. Maximum likelihood estimate; standardized coefficients shown. e is error variance; lower right values are variances; upper left values are standardized means. Weighted.

The estimated model shows that keeping up with the news and perceiving that the election makes a difference are associated with voting. However, the estimates indicate that persons with disabilities are no more likely to be strong or weak partisans, to keep up with the news, be more moderate or extreme than other respondents, or think that the election makes a difference. Overall, the political attitudes of PWDs are *not* much different than other Americans. The effect of disability on voting is primarily through education, income, and a direct effect; it is not found through political attitudes.

This finding is of considerable importance. By accounting for political attitudes that may distinguish persons with disabilities and the indirect effects of disability through education and income, the direct effect we can consider to be an upper bound on the effect of the physical, social, psychological obstacles to voting for persons with disabilities. The obstacles existed even in the 2020 election cycle when convenience voting—mail-in voting and early voting—achieved a level never experienced before in the U.S. In the Model 3 estimates, this disability gap stands at 6.4 percent.

The Interaction Effects of State Voting Methods and Disability on Voting

Having shown a clear and significant disability gap, we now turn to efforts to the question of whether such a gap can be reduced by state election practices. With a linear probability model, we test the interaction between ease of voting and disability status on verified (in our data) or reported (in the CPS) vote. Our measure of ease of voting is the Cost of Voting Index (Li, Pomantell, and Schraufnagel 2018; Schraufnagel, Pomantell, and Li 2020), which is a principal component of measures that capture state laws on automatic voter registration, registration deadlines, voter ID laws, early voting, and mail in voting. The 2020 update of this measure (incorporated changes that were introduced in light of the COVID-19 pandemic that broadly eased restrictions on voting in some states. For interpretability, we multiply the index by negative one, so higher values indicate fewer restrictions on voting.

In Table 2 and Figure 4, we report the comparable multivariate estimates for TASS and CPS surveys. Disability and ease of voting have remarkably similar effects in both surveys, although the larger sample size of the CPS survey allow the ease-of-voting effect to reach statistical significance. In both cases, the estimated disability gap is 7.7 percent, controlling for the other variables in the equations.

In TASS estimates, the interaction effect of disability and state practices on voting is strong. States that greatly eased voting bridged the disability gap. In the CPS estimates, without verified vote and with a weaker measure of disability status, greater ease of voting had about the same positive effect on people with and without a disability.

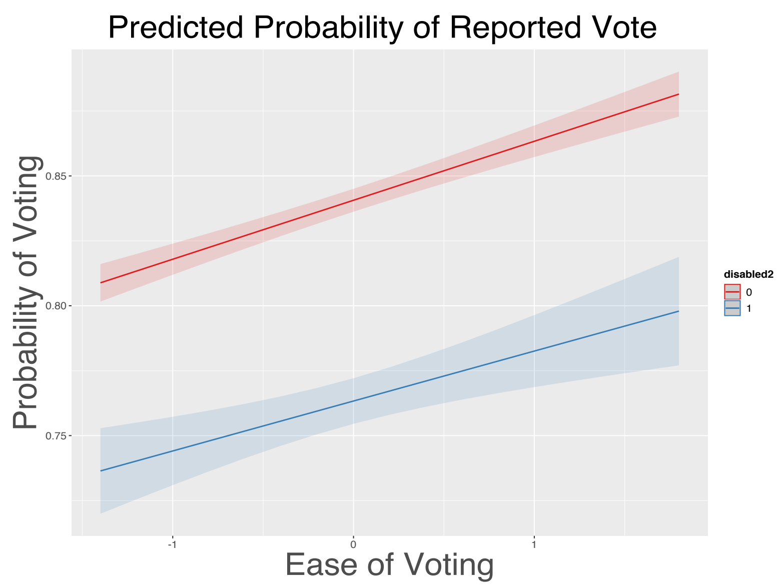
Table 2. Correlates of Voting (Linear Probability Model Estimates).

	CPS (Self-Reported Vote)	TASS (Verified Vote)
Disability x Ease of Voting	-0.003 (0.006)	0.132 (0.043)**
Disability	-0.077 (0.004)**	-0.077 (0.033)*
Ease of Voting	0.023 (0.002)**	0.027 (0.018)
Black	0.045 (0.004)**	-0.078 (0.039)*
Hispanic		-0.067 (0.034)*
Asian	-0.088 (0.007)**	-0.259 (0.056)**
Other Race or Ethnicity	-0.093 (0.012)**	0.034 (0.097)
Two or More Races or Ethnicities	-0.027 (0.010)**	0.034 (0.097)
Education	0.039 (0.001)**	0.037 (0.011)**
Age	0.005 (0.0001)**	0.070 (0.007)**
Income	0.014 (0.0004)**	0.009 (0.003)**
Female	0.027 (0.003)**	0.033 (0.024)

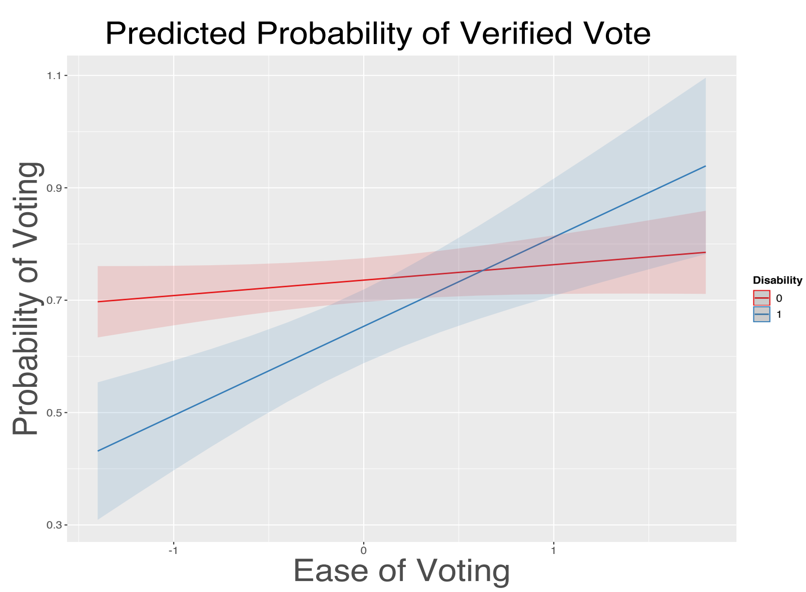
Intercept	-0.031 (0.009)**	0.230 (0.048)**
N	68,122	1,353
R^2	0.146	0.154
Adjusted R^2	0.146	0.147
* *p*<0.05, ** *p*<0.01		

Figure 5. Interactions Effects of Disability and Ease of Voting on Probability of Voting.

CPS Estimates



TAS Estimates



Source: CPS and TASS (see text). From weighted OLS estimates.

The TASS-CPS differences are not readily explained. As we noted, the TASS measure of disability is not the same as the CPS. The TASS measure is more inclusive and may capture some people with disabilities who the CPS instrument missed. Indeed, the TASS shows 17.0 percent of respondents with a disability, while the CPS survey reports 13.6 percent. Thus, it is possible that some people with disabilities

Discussion

We present two main findings. First, we show how disability status can directly and indirectly influence the probability of voting. We demonstrate that the effect of disability is split between a direct effect and the effect that disability has on income and education. Further, age indirectly affects the probability of voting through an increased likelihood of having a disability. Second, we consider how the disability gap can be ameliorated. We find that, under our broad measure of disability, making voting easier, both through fewer barriers to registration and convenience voting methods, eliminates the disability gap.

- Additions to previous studies

In this analysis, we have shown that disability status has an impact on voting, that there are indirect effects of voting that influence standard variables in the voting literature, and provide a tentative understanding of how ease of voting can address the disability gap. There remains, though, much future research to better understand the role of disability status on politics. First, not all disabilities would impact voting in the same way. A person who has difficulty with mobility would disproportionately bear the burden of long lines at the polls, while people with visual impairments might find universal vote by mail particularly onerous. Future work must address type of disability in order to understand the structural barriers that prevent voting.

Voting, to be sure, is not the only form of political participation. Disability gaps likely exist in other areas, such as participation in protests and contacting representatives to assist with problems. The latter is especially precarious, as many people with disabilities rely on government assistance for their livelihood. Even with the advanced age of Members of Congress, there is scant descriptive representation of people with disabilities in legislatures. Representation of people with disabilities requires them to be able to equitably access the political process, and describing current deficiencies is a step toward learning what public policies can address them.

Finally, the intersectionality of disability status with age and race needs more characterization. As we show, people with disabilities are disproportionately old and disproportionately Black. In this study, we find an indirect effect of disability on voting through age and no indirect effect of disability through race. Future research will address other ways that disability may affect other the way in which other demographic characteristics influence access to the political process.

References

Braddock, D.L. & Parish, S. L. (2001). An Institutional history of disability, G. Albreht, K.D Sellman, and M. Bury (Eds.) *Handbook of disability studies*. Thousands Oaks, CA: Sage.

Census Bureau. 2021. “Voting and Registration in the Election of November 2020” (April), https://www.census.gov/data/tables/time-series/demo/voting-and-registration/p20-585.html [accessed May 1, 2021]

Ivan P. Fellegi & Alan B. Sunter (1969) A Theory for Record Linkage, Journal of the American Statistical Association, 64:328, 1183-1210

Fleischer, D., & Zames, F. (2001). *The disability rights movement: From charity to confrontation*. Philadelphia: Temple University Press.

Gastil, John. 2000. “The Political Beliefs and Orientations of People with Disabilities,” *Social Science Quarterly* 81:2, 588-603.

Idielnik, Ruth. “A Political Profile of Disabled Americans,” Pew Resesearch Center, September 22, 2016 {https://www.pewresearch.org/fact-tank/2016/09/22/a-political-profile-of-disabled-americans/]

Katz, Sarah. 2021. “Voting While Disabled is About to Get Harder,” The Atlantic, August 25 [https://www.theatlantic.com/politics/archive/2021/08/people-disabilities-vote/619834/?utm_campaign=the-atlantic&utm_term=2021-08-25T11%3A00%3A56&utm_source=twitter&utm_content=edit-promo&utm_medium=social]

Li, Quan, Michael J. PomanteII, and Scot Schraufnagel. *Election Law Journal: Rules, Politics, and Policy*. Sep 2018.234-247.

Reher, Stephanie. 2020. “Mind This Gap, Too: Political Orientations of People with Disabilities in Europe,” *Political Behavior* 42: 791-818.

Schraufnagel, Scot, Michael J. Pomante II, and Quan Li. Election Law Journal: Rules, Politics, and Policy. Dec 2020.503-509.

Scotch, R. (2001) *From good will to civil rights: Transforming federal disability policy*. Philadelphia: Temple University Press.

Shur, Lisa, and Douglas Kruse. 2021. "Disability and Voting Accessibility in the 2020 Elections: Final Report on Survey Results Submitted to the Election Administration Commission." Election Administration Commission. https://www.eac.gov/sites/default/files/voters/Disability_and_voting_accessibility_in_the_2020_elections_final_report_on_survey_results.pdf [accessed May 1, 2021].

Stewart, Charles. No date. "How We Voted in 2020: A Topical Look at the Survey of the Performance of American Elections," https://electionlab.mit.edu/sites/default/files/2021-03/HowWeVotedIn2020-March2021.pdf [accessed May 1, 2021]

Stewart, Charles, 2021, "2020 Survey of the Performance of American Elections", https://doi.org/10.7910/DVN/FSGX7Z, Harvard Dataverse, V1; Appendix 5. Questionnaire.docx

Switzer, J. (2003) *Disabled rights: American disability policy and the fight for equality*. Washington, D.C.: Georgetown University Press.

Taylor, Danielle M. 2018. "Americans with Disabilities: 2014," Household Economic Studies, United States Census Bureau. Available at: https://www.census.gov/content/dam/Census/library/publications/2018/demo/p70-152.pdf. Accessed May 1, 2021.

Appendix 1: Survey

The survey, The American Social Survey, was sponsored by the Weidenbaum Center of Washington University, and conducted as a three-wave panel in August and November 2020 and February 2021. Most of our questions come from the August and November waves. The November wave was fielded a week after the 2020 election.

Sample. A general population sample of U.S. adults age 18 and older was selected from AmeriSpeak Panel of the NORC at the University of Chicago for this study. The sample for this study is selected from the AmeriSpeak Panel using sampling strata based on age, race/Hispanic ethnicity, education, and gender (48 sampling strata in total). The size of the selected sample per sampling stratum is determined by the population distribution for each stratum. The survey is completed online by most respondents; less the five percent of respondents complete the survey by phone.

AmeriSpeak® is a probability-based panel designed to be representative of the U.S .household population. Randomly selected US households are sampled using area probability and address-based sampling, with a known, non- zero probability of selection from the NORC National Sample Frame. These sampled households are then contacted by US mail, telephone, and field interviewers (face to face). The panel provides sample coverage of approximately 97% of the U.S. household population. Those excluded from the sample include people with P.O. Box only addresses, some addresses not listed in the USPS Delivery Sequence File, and some newly constructed dwellings. While most AmeriSpeak households participate in surveys by web, non-internet households can participate in AmeriSpeak surveys by telephone. Households without conventional internet access but having web access via smartphones are allowed to participate in AmeriSpeak surveys by web. AmeriSpeak panelists participate in NORC studies or studies conducted by NORC on behalf of governmental agencies, academic researchers, and media and commercial organizations.

Weights. Weighting makes only a modest difference because of the stratified sampling. The weighting procedure begins with panel weights, which are computed as the inverse of probability of selection from the NORC National Frame (the sampling frame that is used to sample housing units for AmeriSpeak) or address-based sample. The panel weights are further adjusted to account for unknown eligibility and nonresponse among eligible housing units. The household-level nonresponse adjusted weights are then post-stratified to external counts for number of households obtained from the Current Population Survey. Then, these household-level post-stratified weights are assigned to each eligible adult in every recruited household. Furthermore, a person-level nonresponse adjustment accounts for nonresponding adults within a recruited household. The resulting panel weights are raked to external population totals associated with age, sex, education, race/Hispanic ethnicity, housing tenure, telephone status, and Census Division. The external population totals are obtained from the Current Population Survey. The weights adjusted to the external population totals are the *final panel weights*. The weights for this study are adjusted to account for and adjust for survey nonresponse for those sampled.

Appendix 2. Vote Verification

Matching of state records and our respondents was conducted by Aristotle and reviewed by NORC. To match state records to our respondents, the method of Fellegi and Suner (1969) was used. This method works in the following manner:

- For two files being compared, comparison variables (i.e., from names, date-of-birth fields, and other identification fields) shared on the two files are identified.
- For each comparison variable (birth year, phone number, etc.), the probability of agreement of the two values (i.e., one from each file) being compared are estimates for two distinct sets of pairs.
-
    - Pairs that are matched: i.e. they (truly) represent the same person
    - Pairs that are unmatched: i.e., they do not represent the same person.

  So, for first name, the m-probability is the probability that two records representing the same person have the same value for it. Reasonably this probability is quite high, but disagreement can occur because of misspellings, nicknames, substitution of middle names, etc.
  Again for first name, the u-probability is the probability that two records not representing the same people agree on names. This probability is going to be quite low but dependent on the commonness of the names.
- Using the m- and u- probabilities, agreement and disagreement weights are assigned to each variable based on a formula developed prior to the work of Fellegi and Sunter (based on the statistical model structure):
-
    - The agreement weight (for each of comparison variable) is computed as $AW = log_2\left(\frac{M}{U}\right)$.
        - M – estimated m-probability; U – estimated u-probability
    - The disagreement weight is computes as $DW = log_2\left(\frac{(1-M)}{(1-U)}\right)$
- For each pair being analyzed, we analyze each comparison field (i.e., year of birth, last name, phone number):
    - For a given pair, if the values of the field agree on the two records composing the pair, we assign a variable weight equal to the agreement weight.
    - If the values of the field disagree, we assign a variable weight equal to the disagreement weight.
    - If it is not possible to compare the values (because either are missing or invalid) then a variable weight of 0 is assigned.
- Then for each pair, we sum all the variable weights to generate a pair weight for it.
- In the most basic use of the Fellegi-Sunter technique all pairs scoring above a set cut-off value are linked (i.e., imputed to be true matches). The cuffoff is based on judgment: looking at pairs ranked by pair score and seeing where it seems the

best place to draw the line between what is imputed a match and what is imputed a non-match.

For this analysis, we convert the pair weight into a probability of being a true match. This conversion is based on the assumptions that

- Agreement statuses for the various variables are statistically independent of each other given matched/unmatched status.
- m- and u- probabilities can be accurately estimated
- The proportion of pairs in the set being analyzed can be accurately estimated.

If these assumptions hold then match probability is estimated as follows:

1. An adjustment factor applicable to a set of pairs under analysis is computed as $Adj = log_2\left(\frac{N_M}{N_U}\right)$, where N_M is the number of matched pairs and N_U is the number of unmatched pairs.
2. We compute the adjusted pair weight: $PW_{Adj} = PW + Adj$, where PW is the pair weight computed by summing the variable weights computed for a specific pair and PW_{Adj} is the adjusted pair weight.
3. We compute the pairs' odds of being a match: $Odds_M = 2^{PW_{Adj}}$ where $Odds_M$ is the Odds that the pair is a match. A value of $Odds_M = 3$ says that the odds that a pair is match is 3:1.
4. We compute the match probability as $P(M) = \frac{Odds_M}{1+Odds_M}$

   $P(M)$ is the match probability for a pair. So if the odds were 3 (or 3:1), then $P(M) = \frac{3}{4} = 75\%$.

Again this computation is only accurate to the degree that the assumptions stated above are being met. We will note that some variable agreements are typically correlated, such as sex/gender and first name: when first name agrees, there is a much higher probability that sex/gender will agree (i.e., these agreement statuses are not statistically independent). It is for this reason that we try not to use sex/gender as a comparison variable.
There are several reasons why it is best to use a match probability than a raw pair weight to assess match status:

- Can be used to make error estimates
- Allow pairs developed in different linkage passes to be compared
- Makes understanding accuracy of match status estimates more understandable.

For this analysis, for each AmeriSpeak person record, we keep the pair with the highest match probability assuming it is greater than 97.5%.
To conduct the linkage, we use SAS code (NORCLink) that NORC developed internally to apply the Fellegi-Sunter methodology.

There are several extensions to this methodology that have been implemented with the NORC code:

- Additional pre-processing is performed to clean the data fields being compared

- Nickname substitution are made for commonly used nicknames to generate alternate records to be included in analysis
- For first and last names (but only for exact agreement), we substitute name frequencies computed within the data for the estimated u-probabilities. This means that the u-probability for a common name will be much higher than for an unusual name, and in turn that means that the agreement weight computed for it will be substantially higher. The idea is that agreement on an uncommon name is much more indicative of match status than agreement on a common name.
- For name fields consisting only of an initial, we conduct analysis based on agreement of that initial but the m- and u- probabilities and agreement and disagreement weights are specific to name-initial comparison.
- Name comparisons are made based on measured string similarity. We use the Jaro-Winkler similarity score which ranks similarity from 0 – no letters are in common in the values (for name) being compared to 1 – the values for the name are exactly the same.
- Name comparisons are made according to several levels of similarity:
    - Jaro-Winkler similarity ≥ .85
    - Jaro-Winkler similarity ≥ .90
    - Jaro-Winkler similarity ≥ .95
    - Jaro-Winkler similarity = 1.00 (exactly the same)

The use of multiple levels is meant to enhance the precision of name comparisons. Note that when name values have similarity < .85 they are treated as though they are completely dissimilar. The m- and u- probabilities and agreement and disagreement weights are computed identically to other comparison variable agreements. If the values have similarity ≥ .85 they result in the assigned variable weight being the agreement weight computed for this level based on the corresponding m- and u-probabilities. So they m-probability for the .85 level is the estimated probability that records for the same person have a name similarity of .85 or greater.
For the .90 level, m- and u- probabilities and agreement and disagreement weights are computed contingent on (i.e., conditionally) on having met the .85 level. This is to say that the m-probability for the .90 level is the probability of having a similarity score ≥ .85 or, mathematically, $m = P(JW \geq .90|JW \geq .85)$. If for the pair being analyzed, the similarity score < .85, then no variable weights are assigned for the levels higher than it. The logic used for the .90 is shared for the similarity levels above it. Note analysis for each level above .85 (i.e., .90, .95, 1.00) is applied contingent on having met the immediately lower level.

There were two methods used to pull Aristotle voter registration records to include in the linkage analysis:

- Based on agreement on street-level address (NORC sent Aristotle every address we have ever had on file for each panelist; NORC did not include apartment/unit information; Aristotle returned every record they had for any of these addresses, including every apartment/unit record for a given main address)

- Based on first and last name agreement and several other non-geographic variable agreements (NORC sent Aristotle the name, gender, and age of every AmeriSpeak panelist; Aristotle returned any record with a proprietary fuzzy name match plus gender (match or missing) and age (within 5 year range)

The linkage process was run separately for these two files as they are expected to have different linkage characteristics (particularly matching parameters).

For each of these runs, we used one or more blocking passes. Blocking is a commonly used strategy for record linkage. It indicates that only records falling in the same block are formed into pairs for analysis. For instance for the address-based run, blocking was made by sex and ZIP code. Within these blocks (defined by specific combinations of sex and ZIP code), every AmeriSpeak record included in the analysis was paired with every record in the street-level address-returned Aristotle records: this is known as a Cartesian product. Blocking is used to reduce the computational load that would result from producing and analyzing all possible pairs across the files being linked. Note that records with different values for sex are never compared so a transcription error on this field (or persons undergoing sex change) will not have pairs returned for them.

These were the blocking passes that were conducted by run:

- Aristotle street-level address pull run
  1. Sex and ZIP Code
- Name agreement pull run
  1. Sex, State, ZIP code (first 3 digits)
  2. Sex, Year-of-Birth, Last Name (1st character, only)
  3. Sex, First Name (1st character, only), Month-of-birth, Day of Birth

For each blocking pass, m- and u- probabilities and agreement and disagreement weights are estimated or set separately. However, essentially the same comparison variables were used in each, except the street-level address pull run did not use email address as this was not available:

- First Name (or initial if only this was available)
  - Based on Jaro-Winkler comparison score, compared to levels
- Last Name (or initial if only this was available)
  - Based on Jaro-Winkler comparison score, compared to levels
- Day of Birth (i.e., the numerical day of the month)
- Month of Birth
- Year of Birth
- Phone Number (full 10 digits)
- email Address (only for name agreement pull run)

The results of all these runs were combined, and for each AmeriSpeak person record, we retained only the pairing with the highest estimated match probability if that probability was greater than 97.5%.
Within each pass, parameters (m- and u- probabilities and estimated number of true matches) were estimated used a machine learning methodology known as the

Expectation Maximization Algorithm (EM, for short) which is a well-developed and researched method of obtaining these estimates under the Fellegi-Sunter paradigm.

Appendix 3. SEM Estimates

The models reported in the text exclude paths with a $p < 0.10$. The full models from which the reported models were taken are shown here.

Model 1

```
Structural equation model                                Number of obs = 1,354
Estimation method: ml

Log pseudolikelihood = -14365.74
```

	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
Structural						
income						
black	-2.414912	.4487368	-5.38	0.000	-3.29442	-1.535404
Latinx	-1.28865	.3964177	-3.25	0.001	-2.065614	-.5116852
Asian	2.342608	.842055	2.78	0.005	.6922102	3.993005
gender	-.8584694	.281496	-3.05	0.002	-1.410191	-.3067473
age	.0243731	.008602	2.83	0.005	.0075136	.0412327
_cons	9.56455	.4855634	19.70	0.000	8.612863	10.51624
voted						
income	.0115792	.0041475	2.79	0.005	.0034504	.0197081
education	.0391431	.0151638	2.58	0.010	.0094227	.0688635
black	-.0845324	.0472895	-1.79	0.074	-.1772181	.0081534
Latinx	-.0506023	.0487977	-1.04	0.300	-.146244	.0450393
Asian	-.2273709	.097233	-2.34	0.019	-.4179441	-.0367977
gender	.0350061	.0307042	1.14	0.254	-.0251731	.0951853
age	.0068806	.0009717	7.08	0.000	.0049761	.008785
_cons	.1168211	.0734628	1.59	0.112	-.0271633	.2608056
education						
black	-.2209332	.1260588	-1.75	0.080	-.4680038	.0261375
Latinx	-.3909072	.1136796	-3.44	0.001	-.6137151	-.1680993
Asian	.3371816	.3452509	0.98	0.329	-.3394978	1.013861
gender	.0637034	.088926	0.72	0.474	-.1105884	.2379952
age	.0138704	.0027477	5.05	0.000	.0084851	.0192558
_cons	2.381696	.1574044	15.13	0.000	2.073189	2.690203

Model 2

```
Structural equation model                                Number of obs = 1,345
Estimation method: ml

Log pseudolikelihood = -13737.637
```

	Robust

```
    Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+----------------------------------------------------------------
Structural       |
  income         |
       education |   .2921874   .0329826     8.86   0.000     .2275426    .3568321
      disability |  -.2164154   .0289835    -7.47   0.000    -.2732221   -.1596088
           black |  -.1697984   .0300448    -5.65   0.000    -.2286851   -.1109116
          Latinx |  -.0813744   .0324783    -2.51   0.012    -.1450307   -.0177181
           Asian |   .0994871   .0520185     1.91   0.056    -.0024673    .2014416
             age |   .0690792   .0339123     2.04   0.042     .0026122    .1355461
           _cons |   1.598446   .1465018    10.91   0.000     1.311308    1.885584
  ---------------+----------------------------------------------------------------
  voted          |
          income |   .0815676   .0385807     2.11   0.034     .0059508    .1571844
       education |   .1028952    .039033     2.64   0.008      .026392    .1793984
      disability |  -.0722215   .0347476    -2.08   0.038    -.1403255   -.0041174
           black |  -.0484024   .0316593    -1.53   0.126    -.1104535    .0136487
           Asian |  -.0938369   .0456316    -2.06   0.040    -.1832733   -.0044006
             age |   .2744562   .0350801     7.82   0.000     .2057005    .3432119
           _cons |   .3058631   .1483462     2.06   0.039       .01511    .5966163
  ---------------+----------------------------------------------------------------
  education      |
      disability |  -.1603641   .0357235    -4.49   0.000    -.2303809   -.0903473
           black |  -.0538331   .0336987    -1.60   0.110    -.1198813    .0122152
          Latinx |  -.1303907   .0353388    -3.69   0.000    -.1996534    -.061128
             age |   .2158528   .0390344     5.53   0.000     .1393468    .2923589
           _cons |   2.008174   .1431097    14.03   0.000     1.727684    2.288664
  ---------------+----------------------------------------------------------------
  disability     |
           black |   .0219057   .0356369     0.61   0.539    -.0479413    .0917527
          Latinx |   -.054634   .0321072    -1.70   0.089    -.1175628    .0082949
           Asian |  -.0388495     .02523    -1.54   0.124    -.0882994    .0106005
             age |   .1320626   .0381173     3.46   0.001     .0573541    .2067712
           _cons |   .1235223   .1186677     1.04   0.298    -.1090621    .3561067
```

Model 3

```
Structural equation model                                Number of obs = 1,335
Estimation method: ml

Log pseudolikelihood = -20456.978

----------------------------------------------------------------------------------
                   |                Robust
      Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
Structural         |
  income           |
         education |   .3022019   .0339051     8.91   0.000     .2357491    .3686546
        disability |  -.2142487   .0301355    -7.11   0.000    -.2733132   -.1551841
            Latinx |  -.0568862   .0321712    -1.77   0.077    -.1199405    .0061681
             Asian |   .1184336   .0529072     2.24   0.025     .0147375    .2221298
               age |   .0790087    .034599     2.28   0.022     .0111958    .1468215
             _cons |   1.471016   .1485713     9.90   0.000     1.179821     1.76221
  -----------------+----------------------------------------------------------------
  voted            |
            income |   .0683531   .0383181     1.78   0.074    -.0067489    .1434552
         pstrength |   .0414743   .0329188     1.26   0.208    -.0230453    .1059939
          ideolext |   .0426774   .0333771     1.28   0.201    -.0227405    .1080954
              news |   .0953748   .0372866     2.56   0.011     .0222943    .1684552
      electiondiff |  -.0937561   .0395953    -2.37   0.018    -.1713614   -.0161507
         education |   .0805108    .039424     2.04   0.041     .0032412    .1577804
```

disability	-.0645941	.0322243	-2.00	0.045	-.1277526	-.0014356
Asian	-.1040079	.04679	-2.22	0.026	-.1957147	-.012301
age	.2363548	.0367446	6.43	0.000	.1643366	.3083729
_cons	.2071587	.2255844	0.92	0.358	-.2349785	.6492959
pstrength						
income	.0563385	.0384865	1.46	0.143	-.0190936	.1317707
education	-.0903155	.0394522	-2.29	0.022	-.1676403	-.0129907
disability	.039113	.0359294	1.09	0.276	-.0313074	.1095334
black	.0826424	.0353694	2.34	0.019	.0133196	.1519651
Latinx	.0012921	.0354028	0.04	0.971	-.068096	.0706802
Asian	-.0106385	.0359907	-0.30	0.768	-.081179	.0599021
age	.0949997	.0366073	2.60	0.009	.0232507	.1667487
_cons	.930915	.1541831	6.04	0.000	.6287217	1.233108
ideolext						
income	.0124228	.0386099	0.32	0.748	-.0632512	.0880968
education	.1593197	.0417073	3.82	0.000	.077575	.2410644
disability	-.0240615	.0337057	-0.71	0.475	-.0901234	.0420004
black	-.079838	.0304751	-2.62	0.009	-.1395681	-.0201078
Latinx	-.0213761	.037214	-0.57	0.566	-.0943143	.0515621
Asian	.0254114	.0457564	0.56	0.579	-.0642694	.1150923
age	.0042102	.0367324	0.11	0.909	-.067784	.0762045
_cons	.5123769	.1459699	3.51	0.000	.2262811	.7984728
news						
income	.0747721	.0392562	1.90	0.057	-.0021687	.1517128
education	.0761071	.0406047	1.87	0.061	-.0034765	.1556908
disability	-.0401945	.0349168	-1.15	0.250	-.1086301	.0282412
black	.0483608	.0337094	1.43	0.151	-.0177084	.11443
Latinx	-.0102917	.0342495	-0.30	0.764	-.0774194	.056836
Asian	.1177663	.0647185	1.82	0.069	-.0090796	.2446122
age	.2285044	.0340999	6.70	0.000	.1616698	.295339
_cons	2.806553	.1956398	14.35	0.000	2.423106	3.19
electiondiff						
income	-.0447328	.0407526	-1.10	0.272	-.1246064	.0351409
education	-.1071617	.0386458	-2.77	0.006	-.1829061	-.0314172
disability	-.0206882	.0389632	-0.53	0.595	-.0970547	.0556783
black	.0236116	.0380302	0.62	0.535	-.0509262	.0981493
Latinx	.0266718	.0424121	0.63	0.529	-.0564543	.109798
Asian	-.0127743	.0470537	-0.27	0.786	-.1049978	.0794492
age	-.2076776	.0368234	-5.64	0.000	-.2798503	-.135505
_cons	2.738818	.141653	19.33	0.000	2.461183	3.016452
education						
disability	-.1505125	.0350617	-4.29	0.000	-.2192322	-.0817928
black	-.0601307	.0337342	-1.78	0.075	-.1262485	.0059871
Latinx	-.1291943	.0355095	-3.64	0.000	-.1987917	-.0595968
age	.2322178	.037936	6.12	0.000	.1578646	.3065709
_cons	1.969789	.1422375	13.85	0.000	1.691008	2.248569
disability						
Latinx	-.0501056	.031026	-1.61	0.106	-.1109154	.0107042
age	.1235771	.03595	3.44	0.001	.0531164	.1940377
_cons	.1393025	.1059098	1.32	0.188	-.068277	.3468819