PA

Quantifying Bias from Measurable and Unmeasurable Confounders Across Three Domains of Individual Determinants of Political Preferences

Rafael Ahlskog[ID] and Sven Oskarsson

Department of Government, Uppsala University, Box 514, 75120 Uppsala, Sweden E-mail: rafael.ahlskog@statsvet.uu.se

Abstract

A core part of political research is to identify how political preferences are shaped. The nature of these questions is such that robust causal identification is often difficult to achieve, and we are not seldom stuck with observational methods that we know have limited causal validity. The purpose of this paper is to measure the magnitude of bias stemming from both measurable and unmeasurable confounders across three broad domains of individual determinants of political preferences: socio-economic factors, moral values, and psychological constructs. We leverage a unique combination of rich Swedish registry data for a large sample of identical twins, with a comprehensive battery of 34 political preference measures, and build a meta-analytical model comparing our most conservative observational (naive) estimates with discordant twin estimates. This allows us to infer the amount of bias from unobserved genetic and shared environmental factors that remains in the naive models for our predictors, while avoiding precision issues common in family-based designs. The results are sobering: in most cases, substantial bias remains in naive models. A rough heuristic is that about half of the effect size even in conservative observational estimates is composed of confounding.

Keywords: policy preferences, causal inference, twin, family fixed effects, genetic confounding

1 Introduction

The last decades have seen a steady rise, across the social sciences, in the interest in methods for robust causal inference (Angrist and Pischke 2010; Clark and Golder 2015). This movement, sometimes referred to as the causal identification revolution (Huber 2013), has been spurred by the growing realization that conventional observational methods, utilizing various statistical adjustments for possible confounding factors, often fall short of identifying credible causal effects. An explosion in the use of alternative observational designs, such as instrumental variables, regression discontinuities, or natural experiments, as well as actual experimental designs both in the lab and in the field (Druckman *et al.* 2006), has followed.

A core part of public opinion research, and arguably of political science as a discipline, is to identify how political preferences (such as attitudes about taxation, redistribution, family planning, foreign and environmental policy, etc.) are shaped by things like economic factors, social context, personality traits, education and skills. Scholarly work ranging from classical political economy and sociology (Lipset 1960; Marx 1977), via overarching paradigms like rational choice theory or early life socialization (Jennings and Niemi 1968; 1981) to modern work on the psychological "endophenotypes" underpinning our politics (Oskarsson *et al.* 2015; Smith *et al.* 2011) has argued about various ways, in which individual circumstances and traits can affect placement on ideological issues.

The nature of these questions, however, is such that robust causal identification is often difficult to achieve: experimental interventions into the most important determinants of political preferences are often either unethical, impossible, or prohibitively expensive to implement,[1] while

---

1 Unethical, because interventions required to change, for example, someones' psychological disposition, would be unethical in themselves, or because they may change political attitudes (and therefore outcomes of political processes) that

credible instruments and discontinuities for many of these factors are rare. Where such designs are viable, they may also be of limited value for the substantive questions we are after. As has been argued elsewhere (e.g., Deaton 2009), we often end up saying something credible about the effect among a very narrow group of people (e.g., "compliers" in the case of experiments or IV estimation, or the specific set of people around a regression discontinuity (RDD) threshold) and very little about anyone else, in effect trading external for internal validity (see McDermott 2012 for an overview). Using individual fixed effects estimation with panel data also excludes all predictors that are stable over time. For a fairly large set of important research questions regarding political preference formation, it therefore seems like we are stuck with observational methods whose validity we now know is limited. This raises a crucial question: just how much bias should we expect to find in such estimates? In other words, how wrong will our best guess be?

The purpose of this paper is to attempt to measure the magnitude of bias stemming from both measurable and unmeasurable confounders for three broad and well-established domains of individual determinants of political preferences: socioeconomic factors, moral values, and psychological constructs. To accomplish this, we leverage a unique combination of rich Swedish registry data and a large sample of identical twins, with a comprehensive survey battery of political preference measures. Departing from the registry sources, we attempt to construct the best possible, conservative observational (naive) models, incorporating not only individual, but also family and contextual level controls. We then contrast this to a discordant twin design which also factors out unmeasured genetic and shared familial factors. Doing this across the full range of political preference measures allows us to meta-analyze the average effect size for each model and independent variable. The differences between the naive and discordant twin models can then be used to infer the average degree of confounding stemming from genetic and shared familial factors in the naive models, for each predictor separately. This meta-analytical procedure solves a number of precision problems associated with family-based designs.

The results are sobering: for a large set of important determinants, a substantial bias seems to remain even in conservative naive models. In a majority of cases, half or more of the naive effect size appears to be composed of confounding, and in no cases are the naive effect sizes *under*estimated. The implications of this are important. First of all, it provides a reasonable bound on effect estimates stemming from observational methods without similar adjustments for unobserved confounders. While the degree of bias will vary depending on both predictors and outcomes, a rough but useful heuristic derived from the results of this paper is that effect sizes are often about half as big as they appear. Second, future research will have to consider more carefully the confounding effects of genetic factors and elements of the rearing environment that are not easily captured and controlled for.

2 Background

Statistical controls often go a long way in removing spurious, or noncausal, covariation between two variables of interest. However, the degree to which it is possible to remove *all* bias in this way is crucially dependent on whether or not one can actually measure and correctly specify the variables causing this spurious covariation.

As a salient example, a growing number of studies have documented that just like other human traits (Polderman *et al.* 2015) individual variation in political behavior is also to some degree influenced by genetics (Alford, Funk, and Hibbing 2005; Hatemi *et al.* 2014). This raises the spectre of genetic confounding: traits might be correlated because they are influenced by the same genetic architecture.

---

are not normatively neutral. Impossible, because many things are not sensitive to manipulation (e.g., cognitive capacity). Prohibitively expensive, for example, where income and wealth are concerned.

Controlling for genetic factors requires genetically informative data. This might mean actual genetic sequencing data. However, modern genomic research tells us that complex social phenotypes are influenced by a very large number—possibly millions—of genetic variants (Chabris *et al.* 2015), making required sample sizes for analyses controlling for all appropriate genetic variants, *if* they were known, impossibly large. Moreover, aggregative methods (like adding up all previously identified genetic variants in a so-called polygenic index) have not yet reached a predictive capacity that matches the known magnitude of genetic influences, and will therefore remove only part of the genetic confounding (Young 2019).

Similarly, there might be a number of environmental variables shared in families that are difficult to measure accurately and therefore difficult to control for (parenting style, culture, etc). There is no lack of potential confounders that we can think of, but perhaps we should worry most about the things we *cannot* think of.

Arguably the most powerful way of controlling for both genetic and other familial factors simultaneously is to use known family relationships to partial out these influences. Specifically, the existence of identical twins gives us access to a type of natural experiment that allows us to completely rule out genetic effects *as well as* family environment. This is often called the "discordant MZ-twin" model (Vitaro, Brendgen, and Arseneault 2009), and boils down to comparing individuals *within* identical twin pairs—if the twin with, say, higher education also prefers more stringent environmental policies, this association at least cannot be attributed to the confounding effect of genetics or shared family environment. This approach differs from traditional twin methods in behavior genetics (like variance decomposition) in that it does not seek to map the extent of genetic influence, but instead attempts to find causal relationships between environmental variables free from familial confounding.

Our aim is to quantify the degree of bias both captured by, and remaining in, well-specified observational models of political preference formation. To accomplish this, we will contrast meta-analyzed results for naive models using a comprehensive and conservative set of statistical controls to the results from discordant twin models, for a wide range of political preference measures.

We have tried to cover three general and well-established domains of predictors. The first domain is socioeconomic factors—here, the predictors education, income, and wealth are included. The idea that these types of factors are important for political preference formation is arguably as old as political economy itself. The connection between an individuals' level of education and their politics is well established, with results tending to show higher education associated with more liberal or left-leaning preferences (Dunn 2011; Weakliem 2002, although see Marshall 2016). Similarly, the idea that wealth and income are important determinants of political preferences is central to both the patrimonial voting literature (Lewis-Beck, Nadeau, and Foucault 2013; Quinlan and Okolikj 2019; Ahlskog and Brännlund 2021) as well as public choice theory more broadly (e.g., Meltzer and Richard 1981).

The second domain is moral and social attitudes. In this domain, we have included social trust, altruism and antisocial attitudes, and utilitarian judgement. Social trust—the tendency to think people in general can be trusted (Van Lange 2015)—has been linked to a variety of political preferences, such as support for right-wing populists (Berning and Ziller 2016; Koivula, Saarinen, and Räsänen 2017), attitudes on immigration (e.g., Herreros and Criado 2009) and the size of the welfare state (Bjornskov and Svendsen 2013). Altruism or other-regarding preferences have been proposed to be connected to redistributive politics (Epper, Fehr, and Senn 2020) as well as the general left–right continuum (Zettler and Hilbig 2010). Finally, utilitarian judgement has recently been connected to ideological dimensions such as Right-Wing Authoritarianism and Social Dominance Orientation (Bostyn, Roets, and Van Hiel 2016). Both altruism and utilitarian judgment are also related to the care/harm dimension of Moral Foundations Theory (Graham *et al.*

2013), which has been suggested to be more prevalent among individuals with liberal political attitudes (Graham *et al.* 2011).

The third domain is psychological constructs, and for this domain we have included risk preferences, extraversion, locus of control, and IQ. Risk aversion has been suggested to be a crucial determinant of certain political preferences (Dimick and Stegmueller 2015; Cai, Liu, and Wang 2020). Various personality domains have also been proposed to be important. Although it would be preferable to have data for all of the Big Five domains, we can only include extraversion due to data limitations. There is some evidence that extraversion is connected to social conservatism (Carney *et al.* 2008), although this has been disputed (Gerber *et al.* 2010). The construct locus of control, a measure on to what extent individuals feel responsible for their own life outcomes, has been shown to vary with political affiliation, with conservatives generally having a stronger internal locus of control (Gootnick 1974; Sweetser 2014). Finally, research on the connection between cognitive capacity and political orientation is diverse, with results generally indicating that intelligence predicts more liberal attitudes (Deary, Barry, and Gale 2008; Schoon *et al.* 2010) but also right-wing economic attitudes (Morton, Tyran, and Wengström 2011).

3 Data

The main data come from a large sample of identical twins in the Swedish Twin Registry (STR). The STR is a near-complete nation-wide register of twins established in the 1950s, now containing more than 200,000 individuals (Zagai *et al.* 2019). Apart from being possible to connect to other public registers, the STR also frequently conducts their own surveys, making it not only the largest, but also one of the richest twin data sources available.

Political preference measures are taken from the SALTY survey from the STR. The SALTY survey was conducted in 2009–2010 in a total sample of 11,482 individuals born between 1943 and 1958, and contains measures of, among other things, psychological constructs, economic behavior, moral and political attitudes, and behavior and health measures. Importantly for our purposes, the survey contains a comprehensive battery of 34 political preference measures. We use these 34 items as our outcome space. The items present specific policy proposals spanning issue dimensions from economic and social policy (e.g., *Taxes should be cut* or *Decrease income inequality in society*) to environmental and foreign policy (e.g., *Ban private cars in the inner cities* or *Sweden should leave the EU*), and ask the respondent to indicate to what degree they agree with these proposals on a 1–5 scale. The choice of policy preferences used in the survey overlaps with previous waves of the Swedish Election Study (Holmberg and Oscarsson 2017), which in turn partially overlaps with election studies in several other countries. A full list of the items can be found in Supplementary Appendix A.

Data for the predictors outlined in the background section are gathered from a number of register and survey sources. Precise definitions and sources can also be found in Supplementary Appendix A.

3.1 Additional Datasets

The external validity of the main results is limited by three factors. First, the twin population might differ from the nontwin population simply by virtue of being twins. Second, the subsample of the STR used in this study consists of individuals who have agreed to participate in genotyping (Magnusson *et al.* 2013), which may signal civic-mindedness that makes them different from the rest of the population. Third, Swedes might differ from other nationals.

To check the external validity of the empty versus naive model changes, data from election surveys in Sweden, Denmark, Norway, and the UK will be leveraged. These contain attitude data that can be matched to some of the outcomes used in the main data, as well as variables for a few predictors and controls. Details on these models can be found in Supplementary Appendix C.

4 Method

The method employed follows three steps for each predictor separately. First, three regression models (empty, naive, and within, as outlined below) are run for each political preference outcome in the sample of complete twin pairs. Second, a meta-analytical average for all outcomes, per model, is calculated. Third, this average effect size is compared across models to see how it changes with specification.[2]

This procedure is intended to solve two fundamental problems. The first problem is the one broadly outlined above: it allows us to infer how much confounding each specification successfully captures. The second problem that it solves is that statistical precision is often severely reduced when moving from between- to within-pair estimates, for two reasons. First, since there are half as many pairs as there are individuals, adding pair fixed effects decreases the degrees of freedom by $n/2$. Other things being equal, the standard errors should then be inflated by almost the square root of 2, that is, roughly 1.4. Second, since we are removing all factors shared by the twins, within-pair differences are going to be much smaller than the differences between any two randomly selected individuals in the population. This results in less variation in the exposure of interest and therefore less precision (Vitaro *et al.* 2009). As a consequence, a change in the effect size when going from a naive to a within-pair model is more likely to come about by pure chance than when simply comparing two different naive specifications.

The precision problem is at least partially solved by the aggregation of many outcomes: while we should expect standard errors to be higher in the discordant models, the coefficients should not change in any systematic direction if the naive effect sizes are unbiased. Systematic changes in the *average* effect size across the different preference items is therefore a consequence of model choice (and, we argue, a reduction in bias) rather than variance artefacts.

4.1 Models

4.1.1 Empty. Three models of increasing robustness will be tested in two stages of comparisons. The first (the "empty," e) model will be used as a reference point and controls only for sex, age fixed effects, and their interaction:

$$y_{ij}^e = a + b_j^e x_{ki} + b_2 \text{sex}_i + \sum_{a=1} \left(c_a \text{age}_{ia} + d_a \text{sex}_i \times \text{age}_{ia} \right) + e_i, \tag{1}$$

where i denotes an individual twin and j denotes the outcome.

4.1.2 Naive. The second model (the "naive" model, n), and hence the first model comparison, adds a comprehensive set of controls available in the register data. The ambition is to produce as robust a model as possible with conventional statistical controls. The controls include possible contextual (municipal fixed effects), familial (parental birth years, income, and education) and individual (occupational codes, income, and education) confounders. In total, this should produce a model that is fairly conservative:

$$y_{ij}^n = a + b_j^n x_{ki} + b_2 \text{sex}_i + \sum_{a=1} \left(c_a \text{age}_{ia} + d_a \text{sex}_i \times \text{age}_{ia} \right) + \mathbf{b}\chi_{\mathbf{i}} + e_i, \tag{2}$$

where $\chi_{\mathbf{i}}$ is the vector of naive controls. Complete definitions of all naive controls can be found in Supplementary Appendix A.[3]

---

2 The code for reproducing our results, as well as intermediate level data, are available from this article's Dataverse at https://doi.org/10.7910/DVN/MGEN32.
3 Models without education and income as individual controls are also reported in Supplementary Appendix B.

Rafael Ahlskog and Sven Oskarsson | Political Analysis

4.1.3 *Within.* Finally, the third model (the "within" model, *w*) adds twin-pair fixed effects, producing a discordant twin design.[4] This controls for all unobserved variables *shared* within an identical twin pair, that is, genetic factors, ubpringing and home environment, as well as possible neighborhood and network effects:

$$y_{ij}^w = a + b_j^w x_{ki} + b_2 \text{sex}_i + \sum_{a=1} \left(c_a \text{age}_{ia} + d_a \text{sex}_i \times \text{age}_{ia} \right) + \mathbf{b}\chi_{\mathbf{i}} + \sum_{p=1} \phi_p P_{ip} + e_i, \quad (3)$$

where $\sum \phi_p P_{ip}$ are the twin-pair fixed effects for twin pair p. Note that when adding these pair fixed effects, the age and sex variables as well as many of the controls will automatically be dropped since they are shared within pairs.

4.2 Changes Between Models

In all models, standardized regression coefficients are used to facilitate aggregation and comparison. Furthermore, to be able to calculate meaningful averages of coefficients across the full range of outcomes, all outcomes y_j are transformed to correspond to positive coefficients in the baseline model (i.e., empty when comparing empty vs. naive, and naive when comparing naive vs. within), such that

$$y_j^{e*} = |y_j^e|,$$
$$y_j^{n*} = 6 - y_j \text{ if } b_j^e < 0 \quad (4)$$

is used when introducing the naive controls, and

$$y_j^{n*} = |y_j^n|,$$
$$y_j^{w*} = 6 - y_j \text{ if } b_j^n < 0 \quad (5)$$

is used when moving from the naive to the within model.[5] To see why this transformation is necessary, consider its absence—the average effect size would reflect both negative and positive effects and thus go toward zero, and the average effect size would be the result of the arbitrary coding of the items. Norming the sign by the previous model makes sure that changes to the average when new controls are introduced have an interpretation. Since there are two model comparisons (naive following empty, and within following naive), the naive model will also exist in two versions: one normed with Equation (4) to compare with the empty model (and hence with coefficients on both sides of the null), and one renormed with Equation (5) used as the departure point for the within model (and hence with non-negative coefficients).

When moving from the naive model to the discordant twin model, we should expect the precision in the estimates to decrease more than when moving from the empty to the naive model. This, as elaborated above, follows from the introduction of twin-pair fixed effects and implies that point estimates for any given preference measure can change substantially due to random chance. *In expectation*, however, the change between models is zero under the null hypothesis that the naive model captures all confounding. Leveraging the fact that there are 34 *different* outcomes in the same sample therefore allows us to make inferences about the average degree of confounding remaining: systematic differences between the naive and within models should *not* arise as a consequence of statistical imprecision.

---

4 This is implemented in Stata using the *reghdfe* command and absorbing twin-pair identifiers.
5 The correct mirror transformation is $\max(y) + \min(y) - y_j$, which reduces to $6 - y_j$ since all outcomes range from 1 to 5.

The first step in assessing how the effects change as a consequence of model choice is to calculate the overall average effect size B for all outcomes, for each specification and predictor:

$$B_v^m = \frac{1}{k} \sum_{j=1}^{k} b_{vj}^{m*}. \tag{6}$$

Here, k is the number of dependent variables (i.e., 34 for the main models), m is the model (out of e, n, or w) and v is the predictor.

To know whether changes between models are meaningful, we also need standard errors for this mean. Since political preferences tend to vary along given dimensions (such as left–right) is also necessary to account for the correlation structure of the outcome space. Unless completely independent, the effective number of outcomes is lower than 34. The standard error for B, when adjusted for the correlation matrix of the outcome space, is therefore the meta-analytical standard error for overlapping samples (Borenstein *et al.* 2009, Ch. 24):

$$SE_{B_v^m} = \sqrt{V\left(\frac{1}{k} \sum_{j=1}^{k} b_{vj}^{m*}\right)} = \left(\frac{1}{k}\right) \sqrt{\left(\sum_{j=1}^{k} V_{b_{vj}^{m*}} + \sum_{j \neq l}(r_{jl} \sqrt{V_{b_{vj}^{m*}}} \sqrt{V_{b_{vl}^{m*}})}\right)}, \tag{7}$$

where r_{jl} is the pairwise correlation between preference measures j and l.[6] To see how this correction affects the standard error, we can consider how it changes as the correlations between the preference measures change. In the special case where the outcomes are completely uncorrelated, such that $r_{jl} = 0 \forall j, l$, the formula collapses to $\frac{1}{k} \sqrt{\sum_{j=1}^{k} V_{b_{vj}^{m*}}}$ and is decreasing in the number of outcomes: each preference measure adds independent information. As the correlation between outcomes increases, the standard error also increases by the corresponding fraction of the product of the individual item standard errors, and approaches, in the case where $r_{jl} = 1 \forall j, l$ the simple average of all included standard errors, such that any additional outcome provides *no* additional information.

5 Results

The first set of results, shown in Figures 1 and 2, includes all 34 political preference outcomes.[7] The lightest bars represent the meta-analyzed empty models, that is, the average standardized coefficients across all 34 outcomes, with only sex, age, and their interaction as controls.[8] The largest average coefficient is evident for education years at 11.9% of a standard deviation, trailing all the way down to 3.4% for utilitarian judgment.

We can also see changes to the average effect sizes when moving from the empty to the naive models. It is evident that the extensive controls introduced in the naive models in many cases draw the average effect size substantially toward the null. For example, the naive effect estimate for education years is roughly 67% of the empty effect estimate, for IQ roughly 65% and for work income roughly 47%. In most cases, the reduction in the effect estimate is also itself significant, the exceptions being altruism, antisocial attitudes, and utilitarian judgment, where the average empty estimates were close to zero to begin with.

---

6 A minor detail is that this formula is for completely overlapping samples. While the samples for the different preference outcomes are going to vary slightly due to nonresponse on certain preference issues, this nonoverlap is negligible and will lead to the estimated standard errors being marginally too conservative.

7 An alternative way of viewing specific political preferences is to see them as instantiations of latent ideological constructs. In Supplementary Appendix B, we also present detailed results using indices derived from the first five principal components of the outcome space. The results with these reduced attitude dimensions are qualitatively identical to those presented here.

8 Tables with all details can be found in Supplementary Appendix B, and histograms of the complete effect size distributions for all models and predictors are found in Supplementary Appendix D.
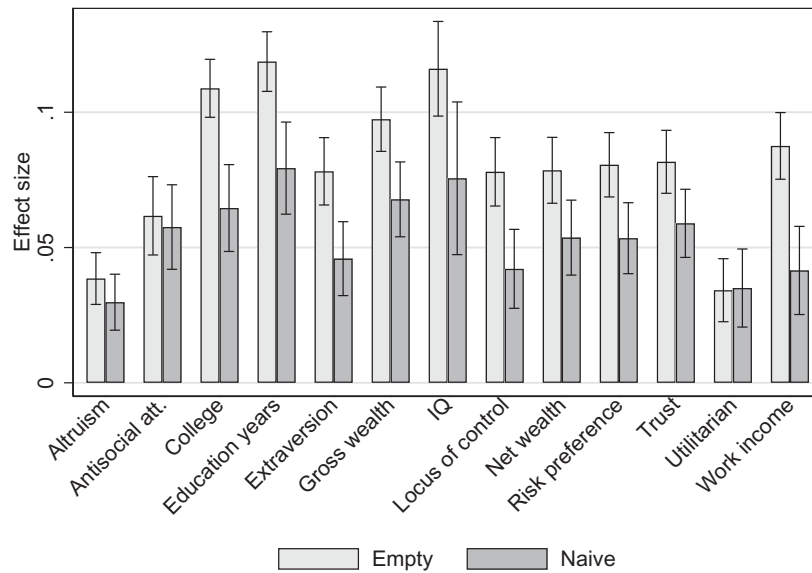

Figure 1. Main results, all outcomes, empty versus naive. Average beta coefficients across all outcomes, per model and predictor. 90% confidence intervals shown.
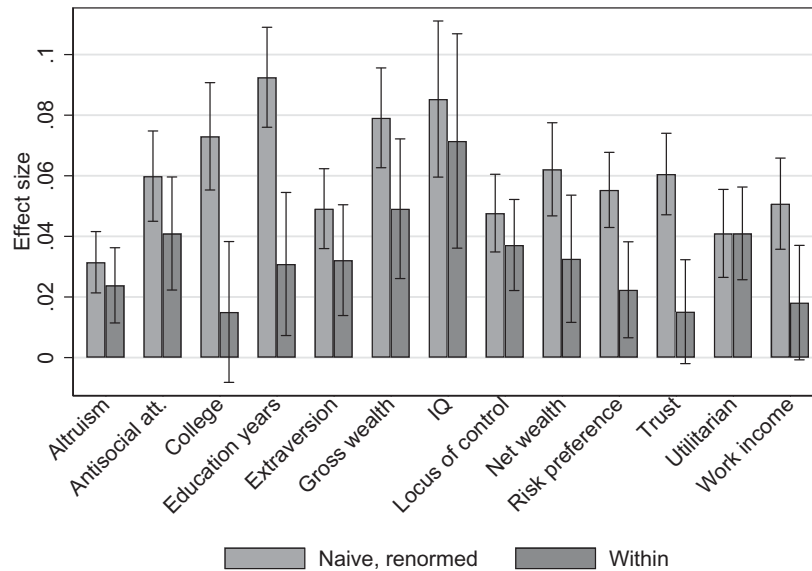

Figure 2. Main results, all outcomes, naive versus within. Average beta coefficients across all outcomes, per model and predictor. 90% confidence intervals shown.

Furthermore, we can see what happens when we move from the naive models to the within models in Figure 2. Comparing the renormed naive to the within models, substantial chunks of the effect sizes are again removed. For example, for education years only 33% of the naive effect size remains, for work income 36% and for trust roughly 25%. For these, as well as college education and risk preferences, the majority of the naive effect size appears to be attributable to unmeasured confounding shared within twin pairs, whereas for net wealth and extraversion it is roughly half. This reduction is itself significant for education years, college, gross and net wealth, work income, trust, and risk preferences. In other cases, most notably utilitarian judgment, the within models are not at all or only slightly closer to zero than the naive models indicating that unmeasured familial confounding is not biasing these results appreciably.

Not all political preferences are theoretically plausibly connected to each predictor, however. Including all political preference outcomes is therefore going to push estimates for all models
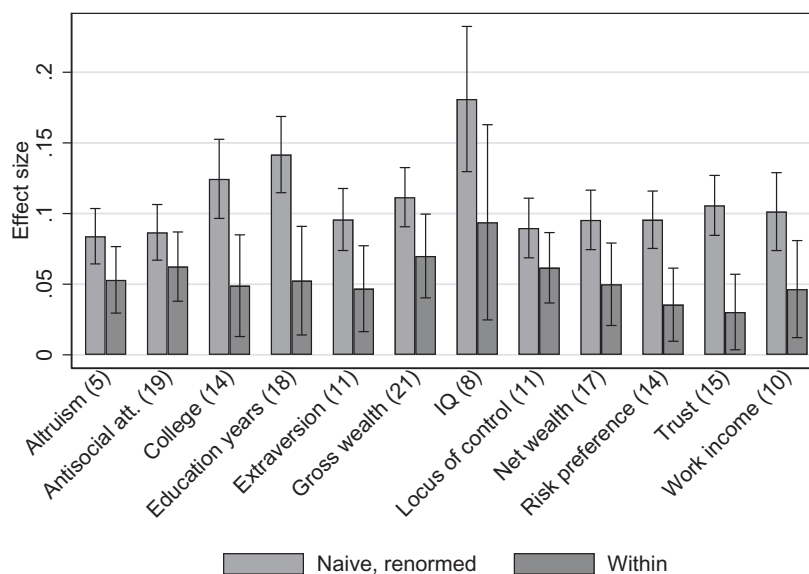
Figure 3. Winner's curse: naive significance selection. Average beta coefficients for outcomes with $p < 0.05$ in naive model, per model and predictor. 90% confidence intervals shown. Only predictors with at least five included outcomes shown (number of included outcomes in parentheses).

toward zero. For this reason, it is also interesting to restrict the analysis to some set of key outcomes per predictor. Barring a carefully constructed set of hypotheses, this can be done in an atheoretical fashion in two different ways. First of all, we report results restricted to the outcomes that were initially statistically significant ($p < 0.05$) in the naive models. Selecting on significance is known to produce a phenomenon known as the "winner's curse"—meaning that the expected effect size is inflated. Apart from testing a more "lenient" set of outcomes, this procedure therefore also partially mimics the effects of publication bias (Young, Ioannidis, and Al-Ubaydli 2008).

In Figure 3, the Winner's curse results selected on naive significance are presented (restricted to predictors with at least five outcomes passing the threshold). As expected, there is a general inflation in effect sizes (the extent to which this is due to the Winner's curse versus a better selection of outcomes is not possible to evaluate with the current methodology). However, there is still a substantial reduction in the average effect sizes for all predictors when moving to within-pair variation. For example, whereas previously the within pair effect of IQ was 84% of the naive effect and the reduction not significant, it is now 52% and significant. The reductions are also significant for all shown predictors except locus of control and antisocial attitudes.

Another way of atheoretically picking the "right" outcomes to include for each predictor is to look at the naive effect size instead of statistical significance. We have set the (somewhat arbitrary) threshold of having a standardized coefficient of at least $\beta > 0.1$. This filters on substantive rather than statistical significance and is going to lead to higher average effect sizes in both models for purely mechanical reasons, in a fashion similar to the Winner's curse.

Figure 4 presents the results selected on naive effect sizes (again restricted to predictors with at least five outcomes passing the threshold). The picture is consistent with the previous results—reductions when moving to within-pair variation are substantial (the remaining effect is generally in the range of 40–60%) and are significant for the shown predictors except antisocial attitudes and IQ.

5.1 Robustness

To evaluate the external validity of the results, we have matched political preference items from Swedish, Danish, Norwegian, and British election surveys to the corresponding items in SALTY
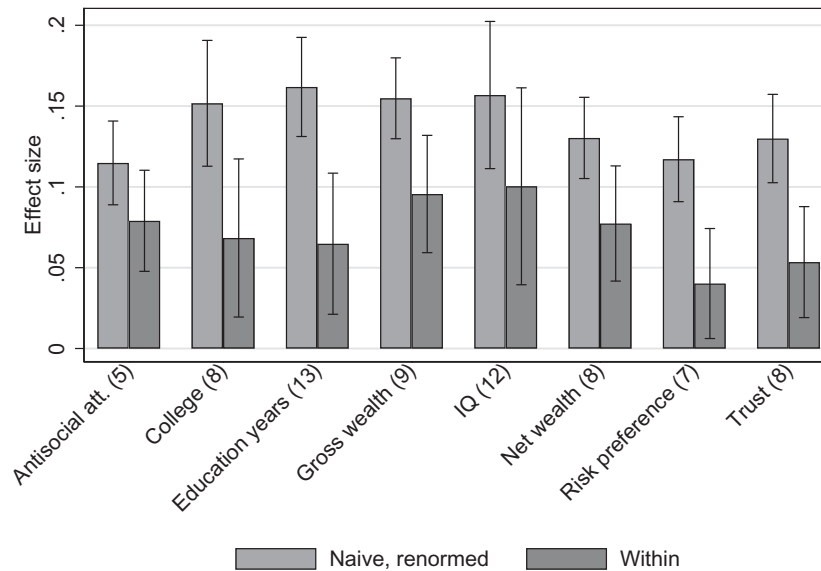

Figure 4. Naive effect size selection. Average beta coefficients for outcomes with $\beta > 0.1$ in naive model, per model and predictor. 90% confidence intervals shown. Only predictors with at least five included outcomes shown (number of included outcomes in parentheses).

and run empty and naive models for a small subset of predictors. These results can be found in Supplementary Appendix C. Both average effect sizes and reductions in effect sizes between models are comparable when using data from the Swedish election study. This indicates that the overall results are not driven by the particular characteristics of the twin sample. Furthermore, effect size reductions are roughly comparable when using data from other Nordic countries. The least consistent results are obtained for college and income in the British sample, where the included controls make sizable dents in the effect sizes in the STR data, but almost none in the British data. This could stem from differences in data quality, but could also be taken to imply that the remaining familial bias is even larger in the British data. It is also possible that it reflects other institutional differences between the UK and the other included countries. In almost all cases, the matched within-pair models in the STR data further cuts the effect sizes dramatically, which indicates that the overall pattern of results should be externally valid.

Furthermore, to check the robustness of the results to violations of the independence assumption, we tested a set of contact rate interaction models. These are outlined in detail in Supplementary Appendix B. In summary, effect size reductions may be slightly inflated, but not significantly so for any predictor except IQ.

6 Discussion

The feasibility of observational methods for capturing causal effects on political preferences depends on to what extent they are able to remove confounding variation. In this study, we have shown that for a fairly large set of predictors, even conservative observational models will often still suffer from substantial bias. This bias can be found across predictors in all of the three domains we were able to investigate, but was particularly pronounced in the domain most often assumed to be crucial for political preference formation—socioeconomic factors like education, income, and wealth.

Using discordant twin analyses to estimate the remaining level of confounding is not without problems. This study was partly designed to overcome one of the main issues—decreased precision—but others can be addressed. In particular, there is always a risk that some remaining factor in the *unique* environment confounds the relationship. We have gone to great lengths to

---

include as powerful statistical controls as possible, but this risk can never be fully ameliorated. However—unique environmental confounders will be missing for both naive and within-pair models, and will therefore not detract from the main takehome message: observational models with a seemingly robust and conservative set of controls are likely substantially biased by familial factors that are difficult to capture.

Another issue that could be important in the present case is random measurement error in the predictor. This becomes important since the attenuating effect of measurement error is potentially magnified when adding twin-pair fixed effects, at least in bivariate models (Griliches 1979).[9] In some of the more important cases, this would be less of an issue since data often come from registers with high reliability (e.g., education, wealth and income etc.), but for some of the psychological constructs it might cause an artifically large "bias reduction" that is actually attributable to magnified attenuation bias. This problem does not apply to problems with systematic measurement error—while a problem in its own right, it will attenuate estimates from all models equally. Unfortunately, without a good idea of the degree of measurement error (i.e., test–retest reliability ratios) for the specific items used to operationalize the predictors in this study, the possible magnitude of this problem is difficult to assess, and such an assessment would only be valid for bivariate comparisons.

Finally, an issue with discordant twin designs that has been discussed in the econometric literature (dating back to Griliches (1979)) is that while twin-fixed effects does filter out much of the endogenous variation, it also filters out exogenous variation. If the proportion between the two is unaltered, we would be in no better situation than without the discordant design altogether (see e.g., Bound and Solon (1999) for a more thorough discussion). However, departing from the plausible assumption that the net effect of unobserved confounders is to inflate, rather than suppress, effect estimates, our within-pair models are still going to provide an upper bound. As such, we would not be finding bias reductions that are artifically large, but too small. This assumption can be bolstered by the fact that the net effect of the *observable* confounders (comparing empty to naive models in Figure 1) is inflationary.

Taking our aggregate effect size reductions at face value, a reasonable heuristic appears to be that we should expect roughly half of the effect size from naive observational methods to be composed of confounding. This result is largely in line with what tends to show up in discordant twin designs with other political outcomes. In Weinschenk and Dawes (2019), the estimated effect of education on political knowledge is reduced by about 72% when going from a naive to a discordant twin model. Similarly, in a recent paper on the relationship between political attitudes and participation by Weinschenk *et al.* (2021), the effect size decreased by 60%, 38% and 35% in Germany, the United States, and Sweden, respectively (in Denmark it was found to disappear completely). Looking at education and political participation, Dinesen *et al.* (2016) found that the effect decreased by 53% in the United states, and disappeared completely in Denmark and Sweden. Thus, it appears likely that our proposed general rule of thumb would also apply to outcomes other than political preferences, like political participation, civicness, knowledge and similar types of behaviors.

The pattern of results shown in this paper should be a strong reminder that observational estimates are likely to be substantially biased—even when a conservative set of controls are utilized. In short, causal conclusions in these situations are rarely warranted. This should not discourage researchers from the well-established approach of using the multiple regression toolkit on observational data—on the contrary, as we point out in the introduction, it is in many cases

---

9 In bivariate models with twin-pair fixed effects, the attenuation becomes a simple factor of the within pair correlation ρ in the predictor: $\beta = \dfrac{\hat{\beta}}{1 - r_e/(1-\rho)}$ where r_e is the ratio of measurement error. In this case, attenuation bias is the same or larger in the within-pair models, but in models with multiple independent variables, the change could go in either direction.

PA

the only tool available to us. However, it underscores the necessity of refraining from using causal language, and making policy recommendations that will, in many cases, fall short in the real world.

Conflicts of Interest

The authors would like to declare no conflicts of interest.

Data Availability Statement

This paper is based on proprietary register data held by Statistics Sweden and the Swedish Twin Register. The full code for obtaining the results reported in the paper, as well as intermediate (aggregate) data is available from the Political Analysis Dataverse at https://doi.org/10.7910/DVN/MGEN32 (Ahlskog and Oskarsson (2022)). The Dataverse also contains a detailed description on how to apply for access to the register data.

Supplementary Material

For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2022.2.

References

Ahlskog, R., and A. Brännlund. 2021. "Uncovering the Source of Patrimonial Voting: Evidence from Swedish Twin Pairs." *Political Behavior*. https://doi.org/10.1007/s11109-020-09669-4

Ahlskog, R., and S. Oskarsson. 2022. "Replication Code for: Quantifying Bias from Measurable and Unmeasurable Confounders across Three Domains of Individual Determinants of Political Preferences." *Harvard Dataverse*. https://doi.org/10.7910/DVN/MGEN32 UNF:6:phnYW+g6qZF/LeaiAs2EzA==[fileUNF]

Alford, J. R., C. L. Funk, and J. R. Hibbing. 2005. "Are Political Orientations Genetically Transmitted?" *American Political Science Review* 99 (2): 153–167.

Angrist, J. D., and J. S. Pischke. 2010. "The Credibility Revolution in Empirical Economics." *Journal of Economic Perspectives* 24 (2): 3–30.

Berning, C., and C. Ziller. 2016. "Social Trust and Radical Rightwing Populist Party Preferences." *Acta Politica* 52: 1–20.

Bjornskov, C., and G. T. Svendsen. 2013. "Does Social Trust Determine the Size of the Welfare State?" *Public Choice* 157 (1–2): 269–286.

Borenstein, M., L. Hedges, J. Higgens, and H. Rothstein. 2009. *Introduction to Meta-Analysis*. West Sussex: Wiley.

Bostyn, D. H., A. Roets, and A. Van Hiel. 2016. "Right-Wing Attitudes and Moral Cognition." *Personality and Individual Differences* 96: 164–171.

Bound, J., and G. Solon. 1999. "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling." *Economics of Education Review* 18: 169–182.

Cai, M., P. Liu, and H. Wang. 2020. "Political Trust, Risk Preferences, and Policy Support." *World Development* 125: 104687.

Carney, D. R., J. T. Jost, S. D. Gosling, and J. Potter. 2008. "The Secret Lives of Liberals and Conservatives." *Political Psychology* 29 (6): 807–840.

Chabris, C. F., J. J. Lee, D. Cesarini, D. J. Benjamin, and D. I. Laibson. 2015. "The Fourth Law of Behavior Genetics." *Current Directions in Psychological Science* 24 (4): 304–312.

Clark, W. R., and M. Golder. 2015. "Big Data, Causal Inference, and Formal Theory: Contradictory Trends in Political Science?: Introduction." *PS: Political Science and Politics* 48 (1): 65–70.

Deary, I., D. Barry, and C. Gale. 2008. "Childhood Intelligence Predicts Voter Turnout, Voting Preferences, and Political Involvement in Adulthood." *Intelligence* 36: 548–555.

Deaton, A. 2009. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." National Bureau of Economic Research Working Paper 14690.

Dimick, M., and D. Stegmueller. 2015. "The Political Economy of Risk and Ideology." SOEP Papers 809-2015.

Dinesen, P. T., et al. 2016. "Estimating the Impact of Education on Political Participation." *Political Behavior*, 38: 579–601.

Druckman, J. N., D. P. Green, J. H. Kuklinski, and A. Lupia. 2006. "The Growth and Development of Experimental Research in Political Science." *The American Political Science Review* 100 (4): 627–635.

Dunn, K. 2011. "Left-Right Identification and Education in Europe: A Contingent Relationship." *Comparative European Politics* 9: 292–316.

Epper, T., E. Fehr, and J. Senn. 2020. "Other-Regarding Preferences and Redistributive Politics." ECON, Working Paper 339, Department of Economics, University of Zurich.

Gerber, A. S., G. A. Huber, D. Doherty, C. M. Dowling, and S. E. Ha. 2010. "Personality and Political Attitudes." *American Political Science Review* 104 (1): 111–133.

Gootnick, A. T. 1974. "Locus of Control and Political Participation of College Students." *Journal of Consulting and Clinical Psychology* 42 (1): 54–58.

Graham, J., B. A. Nosek, J. Haidt, R. Iyer, S. Koleva, and P. H. Ditto. 2011. "Mapping the Moral Domain." *Journal of Personality and Social Psychology* 101 (2): 366–385.

Graham, J., et al. 2013. "Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism." *Advances in Experimental Social Psycholog* 47: 55–130.

Griliches, Z. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy* 87: S37–S64.

Hatemi, P. K., et al. 2014. "Genetic Influences on Political Ideologies: Twin Analyses of 19 Measures of Political Ideologies from Five Democracies and Genome-Wide Findings from Three Populations." *Behavioral Genetics* 44: 282–294.

Herreros, F., and H. Criado. 2009. "Social Trust, Social Capital and Perceptions of Immigration." *Political Studies* 57 (2): 337–355.

Holmberg, S., and H. E. Oscarsson. 2017. "Svensk valundersökning 2010." Svensk Nationell Datatjänst. Version 1.0.

Huber, J. 2013. "Is Theory Getting Lost in the 'Identification Revolution'?" The Political Economist, Summer: 1–3.

Jennings, M. K., and R. G. Niemi. 1968. "The Transmission of Political Values from Parent to Child." *American Political Science Review*, 62 (1): 169–184.

Jennings, M. K., and R. G. Niemi. 1981. *Generations and Politics: A Panel Study of Young Adults and their Parents*. Princeton: Princeton University Press.

Koivula, A., A. Saarinen, and P. Räsänen. 2017. "Political Party Preferences and Social Trust in Four Nordic Countries." *Comparative European Politics* 15: 1030–1051.

Lewis-Beck, M., R. Nadeau, and M. Foucault. 2013. "The Compleat Economic Voter." *British Journal of Political Science* 43 (2): 241–261.

Lipset, S. M. 1960. *Political Man*. New York: Anchor Books.

Magnusson, P. K. E., et al. 2013. "The Swedish Twin Registry: Establishment of a Biobank and Other Recent Developments." *Twin Research and Human Genetics* 16 (1): 317–329.

Marshall, J. 2016. "Education and Voting Conservative: Evidence from a Major Schooling Reform in Great Britain." *Journal of Politics* 78 (2): 382–395.

Marx, K. 1977. *A Contribution to the Critique of Political Economy*. Moscow: Progress Publishers.

McDermott, R. 2012. "Internal and External Validity." In *Cambdrige Handbook of Experimental Political Science*, edited by J. I. Druckman, D. P. Greene, J. H. Kuklinski, and A. Lupia, 27–40. Cambridge: Cambridge University Press.

Meltzer, A. H., and S. F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89 (5): 914–927.

Morton, R., J. R. Tyran, and E. Wengström. 2011. "Income and Ideology: How Personality Traits, Cognitive Abilities, and Education Shape Political Attitudes." Department of Economics, University of Copenhagen, Discussion Paper 11-08.

Oskarsson, S., et al. 2015. "Linking Genes and Political Orientations: Cognitive Ability as Mediator Hypothesis." *Political Psychology* 36 (6): 649–665.

Polderman, T. J. C., et al. 2015. "Meta-Analysis of the Heritability of Human Traits Based on Fifty Years of Twin Studies." *Nature Genetics* 47: 702–709.

Quinlan, S. and M. Okolikj. 2019. "*Patrimonial Economic Voting: A Cross-National Analysis of Asset Ownership and the Vote*." *Journal of Elections, Public Opinion and Parties*. https://doi.org/10.1080/17457289.2019.1655758.

Schoon, I., H. Cheng, C. Gale, D. Batty, and I. Deary. 2010. "Social Status, Cognitive Ability, and Educational Attainment as Predictors of Liberal Social Attitudes and Political Trust." *Intelligence* 38: 144–150.

Smith, K. B., D. R. Oxley, M. V. Hibbing, J. R. Alford, and J. R. Hibbing. 2011. "Linking Genetics and Political Attitudes: Reconceptualizing Political Ideology." *Political Psychology* 32 (3): 369–397.

Sweetser, K. D. 2014. "Partisan Personality: The Psychological Differences Between Democrats and Republicans, and Independents Somewhere in Between." *American Behavioral Scientist* 58 (9): 1183–1194.

Van Lange, P. A. M. 2015. "Generalized Trust: Four Lessons from Genetics and Culture." *Current Directions in Psychological Science* 24 (1): 71–76.

Vitaro, F., M. Brendgen, and L. Arseneault. 2009. "The Discordant MZ-Twin Method: One Step Closer to the Holy Grail of Causality." *International Journal of Behavioral Development* 33 (4): 376–382.

Weakliem, D. L. 2002. "The Effects of Education on Political Opinions." *International Journal of Public Opinion Research* 14 (2): 141–157.

Weinschenk, A. C., and C. T. Dawes. 2019. "The Effect of Education on Political Knowledge: Evidence from Monozygotic Twins." *American Politics Research* 47 (3): 530–548.

Weinschenk, A. C., C. T. Dawes, S. Oskarsson, R. Klemmensen, and A. S. Norgaard. 2021. "The Relationship between Political Attitudes and Political Participation." *Electoral Studies* 69: 102269.

Young, A. I. 2019. "Solving the Problem of Missing Heritability." *PLoS Genetics* 15 (6): e1008222.

Young, N. S., J. Ioannidis, and O. Al-Ubaydli. 2008. "Why Current Publication Practices May Distort Science." *PLoS Medicine* 5 (10): e201.

Zagai, U., P. Lichtenstein, N. L. Pedersen, and P. K. E. Magnusson. 2019. "The Swedish Twin Registry: Content Management as a Research Infrastructure." *Twin Research and Human Genetics* 22 (6): 672–680.

Zettler, I., and B. E. Hilbig. 2010. "Attitudes of the Selfless: Explaining Political Orientation with Altruism." *Personality and Individual Differences* 48 (3): 338–342.