PA

Profiling Compliers and Noncompliers for Instrumental-Variable Analysis

Moritz Marbach[1] and Dominik Hangartner[1,2]

[1] Center for Comparative and International Studies, ETH Zurich, 8092 Zurich, Switzerland
[2] Department of Government, London School of Economics, London WC2A 2AE, UK. Email: d.hangartner@lse.ac.uk

Abstract

Instrumental-variable (IV) estimation is an essential method for applied researchers across the social and behavioral sciences who analyze randomized control trials marred by noncompliance or leverage partially exogenous treatment variation in observational studies. The potential outcome framework is a popular model to motivate the assumptions underlying the identification of the local average treatment effect (LATE) and to stratify the sample into compliers, always-takers, and never-takers. However, applied research has thus far paid little attention to the characteristics of compliers and noncompliers. Yet, profiling compliers and noncompliers is necessary to understand what subpopulation the researcher is making inferences about and an important first step in evaluating the external validity (or lack thereof) of the LATE estimated for compliers. In this letter, we discuss the assumptions necessary for profiling, which are weaker than the assumptions necessary for identifying the LATE if the instrument is randomly assigned. We introduce a simple and general method to characterize compliers, always-takers, and never-takers in terms of their covariates and provide easy-to-use software in R and STATA that implements our estimator. We hope that our method and software facilitate the profiling of compliers and noncompliers as a standard practice accompanying any IV analysis.

Keywords: instrumental variables, local average treatment effects, non-compliance

1 Introduction

Instrumental-variable (IV) estimation is an important tool for applied researchers across the social and behavioral sciences to address noncompliance with the assigned treatment in randomized experiments and unobserved confounding in observational studies (see, e.g., Sovey and Green 2011; Dunning 2012). To date, the predominant framework for IV, introduced by Angrist, Imbens, and Rubin (1996), has been based on potential outcomes. In contrast to classic structural equation models (Haavelmo 1943), which assume constant treatment effects, IV analysis based on potential outcomes allows for heterogeneous treatment effects (under the assumption of monotonicity, see below). This framework uses principal stratification (Frangakis and Rubin 2004) to classify the population into distinct, nonoverlapping but unobserved groups: compliers, always-takers, never-takers, and defiers.

The principal strata are defined by their reaction to the instrument—the assignment to take the treatment (or not). *Compliers* always comply with the assigned treatment: they take the treatment only when assigned to take it and do not take the treatment when they are not assigned. *Always-takers* always take the treatment, independent of whether they are assigned or not. *Never-takers* never take the treatment, independent of whether they are assigned or not. *Defiers* always do the opposite of the assignment: when assigned, they do not take the treatment and when not assigned, they do take it. One implication of observations' agency over whether or not to comply with the assigned treatment is that comparing those who take the treatment to those who refuse it does not provide a causal estimate of the treatment effect since the reasons for (non-)compliance might well be correlated with other characteristics that affect the outcome. In this context, IV analysis based on the potential outcome framework offers a clearly defined set of assumptions that point-identify the treatment effect for the complier subpopulation—the

local average treatment effect (LATE). The corresponding treatment effects for noncompliers (i.e., always-takers, never-takers, and defiers) are, by definition, not identified (Angrist, Imbens, and Rubin 1996).

The focus of IV analysis on compliers raises questions about how different this group is from the noncompliers and by how much the LATE differs from the average treatment effect (ATE). The latter is often the main quantity of interest for applied researchers but different from the LATE if treatment effects are heterogeneous. These questions typically become more important the weaker the instrument is, and the larger the proportion of noncompliers. The "localness" of the LATE led some researchers to doubt its usefulness for the study of economic and political phenomena (Deaton 2009; Heckman and Urzua 2010). Like other scholars (e.g., Imbens 2010), we believe that rather than abandoning IV altogether, researchers should pay more attention to the potential limitations of the LATE. A transparent discussion of the external validity (or lack thereof) of the LATE beyond compliers should build on an explicit comparison of compliers, always-takers, and never-takers. In this letter, we introduce such a method for profiling compliers and noncompliers, corresponding statistical software, and provide an illustration of the approach.

Profiling compliers and noncompliers as discussed in this article can play three important roles in assessing the generalizability of the LATE. First, heterogeneity in treatment responses across compliance strata (and, therefore, how much the LATE and ATE differ) is typically driven by both observable and unobservable variables. Thus, finding that compliers and noncompliers are highly similar in terms of their observable covariates does not imply that they are also similar in terms of unobserved variables, nor does it suggest that we can generalize the LATE to the ATE without invoking additional assumptions (see, e.g., Aronow and Carnegie (2013) and their re-weighting method). But, if profiling reveals that compliers and noncompliers are different with regard to observable covariates that are likely to be predictive of treatment effect size, we know that any attempt to directly learn about the ATE from the LATE is prone to lead to biased inferences. Second, different instruments for the same treatment often estimate different LATEs. One example for this is provided by Angrist and Pischke (2009), who compare two instruments that increase the propensity to have a third child; namely, if the first two children are twins or have the same gender. Other examples where different instruments for the same treatment exist can be found in the literature on the economic and political returns to education and on the effects of get-out-the-vote campaigns. Comparing the compliers defined by the different instruments can help explain the differences in the corresponding LATEs. Third, one can often conjecture how a slightly stronger (or weaker) instrument than the one actually used might be able to convert some marginal never-takers to compliers (and vice versa). For example, in the context of the *Washington Post* study discussed below, a slightly stronger instrument might take the form of a financial reward for actually reading the newspaper (in addition to the free subscription). Comparing the actual never-takers and compliers allows us to form expectations about the treatment effects of the additional compliers encouraged by a stronger instrument and their similarity to the LATE in the original study. And if the researcher cannot (further) increase the strength of the instrument, this comparison allows for a better understanding of the never-takers who cannot be pushed to take the treatment.

In the following, we provide a simple method to characterize compliers and noncompliers in terms of their covariates. We discuss the assumptions required for profiling and show that they are weaker than the assumptions needed for identification of the LATE if the instrument is randomly assigned. Our method can be applied to IV analysis in both experimental and observational studies and whether or not the exclusion restriction holds.

While the idea to profile compliers and noncompliers is not new, it is rarely done in practice. We reviewed all seventy-one papers using IV (including fuzzy regression discontinuity) designs that

were published in the *American Journal of Political Science*, the *American Political Science Review*, the *Journal of Politics*, and *Political Analysis* between January 2013 and December 2018. None of these papers provides information as to if and how compliers differ from the rest of the sample.[1] In economics, the share is only slightly higher: of the 280 papers using IV or fuzzy regression discontinuity designs that were published during the same time period in the *American Economic Review*, *Econometrica*, the *Journal of Political Economy*, and the *Quarterly Journal of Economics*, 10 articles compared compliers to the entire sample in terms of their covariates.[2] A cursory reading of recent empirical papers in demography, epidemiology, and sociology reveals that profiling of compliers and noncompliers is similarly rare. Profiling is likely unpopular due to both a lack of awareness among researchers and the limitations or complexity of existing methods.

Prior research has used a few approaches for profiling. Pinotti ([2017](#)) profiles compliers by regressing an interaction of the covariate and the treatment on the instrumented treatment using two stage least squares (2SLS) and, based on this, compares the covariate means of compliers to the total sample. This approach is inefficient as it only leverages compliers assigned to treatment but disregards information about compliers assigned to control. Angrist and Pischke ([2009](#)) exploit variation in the first stage across covariate groups to estimate the relative likelihood that compliers (compared to the entire sample) will have a particular characteristic. Since it is focused on ratios, this method is limited to binary (or dichotomized) covariates. As an alternative method that is not limited to binary covariates, Baiocchi, Cheng, and Small ([2014](#)) reweight the sample to estimate the covariate mean of compliers. A similar approach is Abadie's ([2003](#)) κ-weighting scheme. In addition to weighting the local average response function (LARF) regression of the outcome on the treatment and the covariates, the κ weights can also be used to weight binary and continuous covariates to the subsample of compliers. In the standard LARF regression that conditions on covariates, iterative least squares algorithms might fail to find the weights that correspond to the global residual-sum-of-squares minimum since the estimated weights will be negative for always-takers and never-takers by construction.[3] Note, however, that when the instrument is randomized, the κ weights can be simplified such that no optimization is needed to estimate covariate means for compliers. In that case, the κ weights offer an alternative but equivalent estimator for the complier mean to ours.

Given the limited popularity of these approaches, we developed a general, simple—and, arguably, more intuitive—method. In the following, we detail the assumptions needed to identify the sample means of covariates for compliers, always-takers, and never-takers. In conjunction with the software that implements our estimator for R and STATA, we hope that this will facilitate the profiling of compliers and noncompliers as a standard practice accompanying every IV analysis.

The rest of this letter is structured as follows: the next section provides an informal summary that conveys the core idea of our approach in a nonmathematical fashion. The following section then discusses the assumptions needed to identify the covariate means of complier, always-taker, and never-taker strata in more formal terms. We then apply this estimator to provide descriptive statistics of compliers and noncompliers in a randomized encouragement experiment on the effect of reading the *Washington Post* on voting behavior and public opinion. A second application, presented in the Supplementary Materials (SM), focuses on compliers and noncompliers in a study on the effects of watching *Fox News* on voting in a referendum on affirmative action. Both examples show how profiling compliers and noncompliers is an important first step in gauging the external validity of the effect estimates. A brief conclusion discusses possible ways to generalize the

---

1 The only political science paper we could find, Pianzola *et al.* ([2019](#)), is published in 2019.
2 These papers are: Abdulkadiroğlu, Angrist, and Pathak ([2014](#)), Dahl, Køstol, and Mogstad ([2014](#)), Jacob, Kapustin, and Ludwig ([2014](#)), Gelber, Isen, and Kessler ([2015](#)), Abdulkadiroğlu *et al.* ([2016](#)), Deshpande ([2016](#)), Huang *et al.* ([2017](#)), Pinotti ([2017](#)), Solis ([2017](#)), and Dobbie, Goldin, and Yang ([2018](#)).
3 It is, however, possible to use alternative optimizers or to take iterated expectations over the κ weights such that they are strictly nonnegative; see, e.g., page 181 in Angrist and Pischke ([2009](#)).

Table 1. Potential treatments and the instrument define four subpopulations: compliers, never-takers, always-takers, and defiers.

Principal strata	Instrument and potential treatments
Compliers	$D(0) = 0$ and $D(1) = 1$, i.e., $D(1) > D(0)$
Always-takers	$D(0) = D(1) = 1$
Never-takers	$D(0) = D(1) = 0$
Defiers	$D(0) = 1$ and $D(1) = 0$, i.e., $D(1) < D(0)$

proposed method beyond the binary instrument and binary treatment case and points the reader to the software that implements our estimator.

2 Intuition

Before we more formally discuss our method of profiling compliers and noncompliers, this section provides a nonmathematical summary of the core idea. Consider an IV scenario with a binary treatment, a binary, scalar instrument, and two-sided noncompliance. Assume that the instrument is independently assigned and that there are no defiers (such that the study group consists of compliers, always-takers, and never-takers). Under these two assumptions, we can directly identify the compliance status of some units by comparing their instrument and treatment values: subjects assigned to the control group who take the treatment are "observable" always-takers, and subjects assigned to the treatment group who do not take the treatment are "observable" never-takers. Because observable and nonobservable always-takers and never-takers, respectively, have the same covariate distribution if the instrument is independently assigned, we can directly estimate the covariate means for these two subpopulations. In contrast, compliers cannot be identified at the individual level since compliers assigned to the control group are, with respect to their realized instrument and treatment values, observably identical to never-takers assigned to control; and compliers assigned to the treatment group are observably identical to always-takers assigned to the treatment. However, by subtracting the weighted covariate mean of observable always-takers and never-takers from the covariate mean of the entire sample, we can back out the covariate mean for compliers.

3 Method

Following the notation of Angrist, Imbens, and Rubin (1996) who discuss IV analysis within the potential outcome framework, we assume that every observation has two binary, fixed, and unobserved potential treatment indicators ($D(0)$ and $D(1)$) that realize a binary, observed treatment indicator (D). The treatment D depends on the binary, scalar instrument, Z, i.e., $D = D(Z)$. If the unit is assigned to treatment, then $Z = 1$, and $Z = 0$ otherwise. Together, the instrument and treatment indicators define four subpopulations, the principal strata, as shown in Table 1.

Since only D, but not $D(0)$ and $D(1)$, is observed, the principal strata that a unit belongs to is unknown without further assumptions. In order to identify, and estimate, the mean of pretreatment (and preinstrument) covariate X, including uncertainty estimates for each of the four subpopulations, we impose two identifying assumptions. Since these two assumptions are similar to the assumptions needed to identify the LATE and weaker if the instrument is randomly assigned, profiling compliers and noncompliers comes at little additional cost.

ASSUMPTION 1 (Monotonicity). $D(1) \geq D(0)$.

ASSUMPTION 2 (Independence of the instrument). $D(0), D(1), X \perp\!\!\!\perp Z$.

The first assumption, monotonicity, rules out defiers, for whom $D(1) < D(0)$. The second assumption, independence of the instrument, implies that the instrument is assigned independently of a unit's compliance stratum and covariate value. Note that Assumption 2 is both weaker and stronger than what is needed to identify the LATE. On the one hand, for profiling, researchers do not have to invoke the exclusion restriction, which assumes that the instrument is also independent of the two potential outcomes. This implies that we can characterize compliers and noncompliers even when the exclusion restriction is violated and the instrument affects the outcome through channels other than the treatment.[4] On the other hand, for profiling, we have to assume that the covariate X is independent of the instrument. The latter assumption will hold trivially if the instrument is randomly assigned. In addition to these two assumptions, the probability of assignment has to be bounded between 0 and 1, $0 < \mathbb{P}(Z = 1) < 1$, and must influence the treatment, $\mathbb{E}[D|Z = 1] \neq \mathbb{E}[D|Z = 0]$ (first stage), as otherwise the fraction of compliers is 0.

These assumptions are sufficient to identify the covariate means for always-takers

$$\mathbb{E}[X|D(1) = D(0) = 1] = \mathbb{E}[X|D = 1, Z = 0] \tag{1}$$

by focusing on the observable subset of nonencouraged always-takers. Equivalently, we can identify the covariate means of never-takers by focusing on the observable never-takers:

$$\mathbb{E}[X|D(1) = D(0) = 0] = \mathbb{E}[X|D = 0, Z = 1]. \tag{2}$$

Next, we turn to compliers. First, we employ the Law of Iterated Expectations to decompose the population mean of X into a linear combination of the strata means, weighted by the size of each strata:

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X|D(1) > D(0)]\mathbb{P}[D(1) > D(0)] \\
&+ \mathbb{E}[X|D(1) = D(0) = 1]\mathbb{P}[D(1) = D(0) = 1] \\
&+ \mathbb{E}[X|D(1) = D(0) = 0]\mathbb{P}[D(1) = D(0) = 0].
\end{aligned} \tag{3}$$

Invoking the monotonicity and independence assumption, we can substitute all potential treatment indicators with their observed counterparts, but for $\mathbb{E}[X|D(1) > D(0)]$,

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X|D(1) > D(0)](\mathbb{E}[D|Z = 1] - \mathbb{E}[D|Z = 0]) \\
&+ \mathbb{E}[X|D = 1, Z = 0]\mathbb{P}[D = 1|Z = 0] \\
&+ \mathbb{E}[X|D = 0, Z = 1]\mathbb{P}[D = 0|Z = 1].
\end{aligned} \tag{4}$$

Solving equation (4) for $\mathbb{E}[X|D(1) > D(0)]$, we get

$$\begin{aligned}
\mathbb{E}[X|D(1) > D(0)] &= (\mathbb{E}[X] - \mathbb{E}[X|D = 1, Z = 0]\mathbb{P}[D = 1|Z = 0] \\
&- \mathbb{E}[X|D = 0, Z = 1]\mathbb{P}[D = 0|Z = 1]) \\
&(\mathbb{E}[D|Z = 1] - \mathbb{E}[D|Z = 0])^{-1}
\end{aligned} \tag{5}$$

such that the entire right-hand side is written in terms of observed treatments, such that the covariate means for compliers are identified.

---

4 While in most cases there is little point in profiling compliers and noncompliers when the exclusion restriction does not hold, we can imagine some scenarios, for example, when the direction of the bias is known and the goal of the LATE is to bound (rather than point-identify) the true treatment effect for compliers, where profiling might still be useful.

Table 2. Stratification of the population by realized treatment (D) and assignment (Z) into compliers, never-takers, and always-takers.

		Z	
		0	1
D	0	Compliers and never-takers	Never-takers
	1	Always-takers	Compliers and always-takers

Building on the intuition developed in the preceding section, Table 2 stratifies the population by realized treatment (D) and assignment (Z) to help examine these identification results. After ruling out defiers using the monotonicity assumption, we can identify the always-taker and never-taker strata from their directly observable subsets in the off-diagonal cells. The independence of the instrument assumption ensures that the observable and nonobservable always-takers and never-takers are exchangeable. The identifiability of the compliers that are not directly observable (on the main diagonal) follows from the observation that we can subtract the contribution of the always-takers and never-takers from the overall population mean to back out the mean of the complier strata.

Having established the assumptions needed to identify the different strata, we now turn to estimation. We construct our estimator by replacing the population means and shares in equations (1), (2), and (5) with their sample analogs (Manski 1988). While the derivation of the estimators is relegated to the SM Section 1, we briefly sketch here the sample analog of the main result from equation (5).

Let N be the sample size, K_{nt} the number of observable never-takers, K_{at} the number of observable always-takers, and $N_{Z=z}$ the number of units with realized instrument value z. Let x_i be the covariate value for the ith unit. We write the estimators for the covariate mean of the entire sample, $\hat{\mu}$, for always-takers, $\hat{\mu}_{at}$, and for never-takers, $\hat{\mu}_{nt}$, as

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} x_i, \quad \hat{\mu}_{at} = \frac{1}{K_{at}} \sum_{i=1}^{K_{at}} x_i, \quad \hat{\mu}_{nt} = \frac{1}{K_{nt}} \sum_{i=1}^{K_{nt}} x_i. \quad (6)$$

By weighting these covariate means by the estimated sample shares of compliers, $\hat{\pi}_{co}$, always-takers, $\hat{\pi}_{at}$, and never-takers, $\hat{\pi}_{nt}$, as in equation (5), we can estimate the covariate mean for compliers, $\hat{\mu}_{co}$, as follows:

$$\hat{\mu}_{co} = \frac{1}{\hat{\pi}_{co}} \hat{\mu} - \frac{\hat{\pi}_{nt}}{\hat{\pi}_{co}} \hat{\mu}_{nt} - \frac{\hat{\pi}_{at}}{\hat{\pi}_{co}} \hat{\mu}_{at}. \quad (7)$$

Because the shares of the complier, always-takers, and never-takers are unknown and have to be estimated, the derivation of the standard error for the complier mean is somewhat tedious. Given the extremely low computational costs in this context, we use the bootstrap method to obtain standard errors that reflect the estimation uncertainty in both the covariate means and the sample proportions. The SM detail the results from a series of Monte Carlo experiments that verify that the empirical coverage rate closely tracks the nominal coverage rate of the 95% bootstrap confidence interval across different sample sizes. These simulations also confirm that the bias in means decreases at the expected rate as the sample size grows.

4 Application

Gerber, Karlan, and Bergan (2009) report the results of a randomized field experiment in which subjects were sent a free ten-week subscription for either the *Washington Post* or the *Washington Times*. For this analysis, we focus on those $N = 503$ subjects that were randomized to the *Post* or the control group and completed the baseline and the follow-up survey. The baseline survey

was conducted in September 2005 and asked the respondents to, inter alia, indicate their gender, age, past turnout, and their preference for a Democratic or Republican governor. The follow-up survey was conducted during the week after the November 2005 Virginia gubernatorial election and asked the respondents about the newspapers they receive and how frequently they read them and their voting behavior during the gubernatorial election as well as their attitudes on a range of topics. Based on these data, and using ordinary least squares regression to estimate intention-to-treat (ITT) effects, the authors estimate the effect of receiving a free *Post* on a range of outcomes including political knowledge, policy positions, turnout, and voting for the Democratic or Republican candidate. Among all the outcomes considered, only the ITT effect of voting for the Democratic candidate is statistically significant: getting a free subscription to the *Post* increases the probability of selecting the Democrat by 7.9 percentage points ($p < 0.082$) among voters.[5] SM Table S.1 shows the corresponding IV analysis. For this, we code the $D = 1$ if the subjects report that they received the *Post* and 0 otherwise.[6] Using 2SLS, we estimate a LATE of 22.3 percentage points ($p < 0.083$) for respondents that comply with the *Post* assignment.

Next, we use the estimators described above to profile compliers and noncompliers. Figure 1 shows the covariate means for the entire sample and the sample shares and covariate means for compliers, always-takers, and never-takers across eight socio-political characteristics, all measured pretreatment in the baseline survey. Numerical estimates are provided in SM Table S.2. About 34.1% of the subjects are compliers, 20.3% are always-takers, and 45.6% are never-takers.

The upper panel of Figure 1 shows that all strata have relatively similar turnout levels in 2001, 2002, and 2004. However, we find some meaningful differences in terms of socio-demographic characteristics. Compliers tend to be younger than never-takers, who, in turn, are younger than always-takers; and compliers are less likely to be female compared to always-takers and never-takers (the latter group features the highest share of women). In terms of political preferences, compliers and always-takers are both less likely to support the Republican candidate compared to never-takers, while always-takers are slightly more likely to prefer none of the candidates compared to compliers and never-takers.

These differences have two main implications for our ability to generalize the estimates from the subsample of compliers. First, since the causal effect of treatment receipt is only defined for compliers, any attempt to directly generalize those estimates to always-takers and never-takers is purely speculative. Second, who is (and is not) a complier is a function of the instrument and is fixed only in the context of the particular study analyzed. Therefore, a slightly stronger instrument might be expected to encourage some never-takers to become compliers. In this study, participants in the encouragement group were simply sent a postcard announcing that they had won a free subscription to the *Post*. Thus, combining the free subscription with, e.g., a financial incentive to read the newspaper might strengthen the instrument by incentivizing some marginal never-takers to become compliers. Would we expect the treatment effect for these additional compliers to be the same as the LATE in the original study? Our profiling method is crucial for answering this question: given the differences in age and gender between compliers and never-takers and the preexisting differences in support for the Republican candidate—all factors that are likely related to changes in party attachment and party vote switching (see, e.g., Campbell *et al.* 1980)—we have little reason to assume that the LATE estimated for compliers can be generalized to other study subjects.

A second application, discussed in the SM, profiles compliers and noncompliers in a randomized encouragement experiment on the effect of watching *Fox News* on support for

---

5 Note that this estimate is slightly smaller with a correspondingly higher *p*-value compared to the ITT reported in Gerber, Karlan, and Bergan (2009) that leverages all treatment groups.

6 There are several reasons as to why subjects that were randomized to the *Post* subscription might not report to receiving it; e.g., because the newspaper was not delivered, because they are receiving it but are unaware of it, or because they misinterpret the survey question (see Gerber, Karlan, and Bergan (2009) for a more detailed discussion).
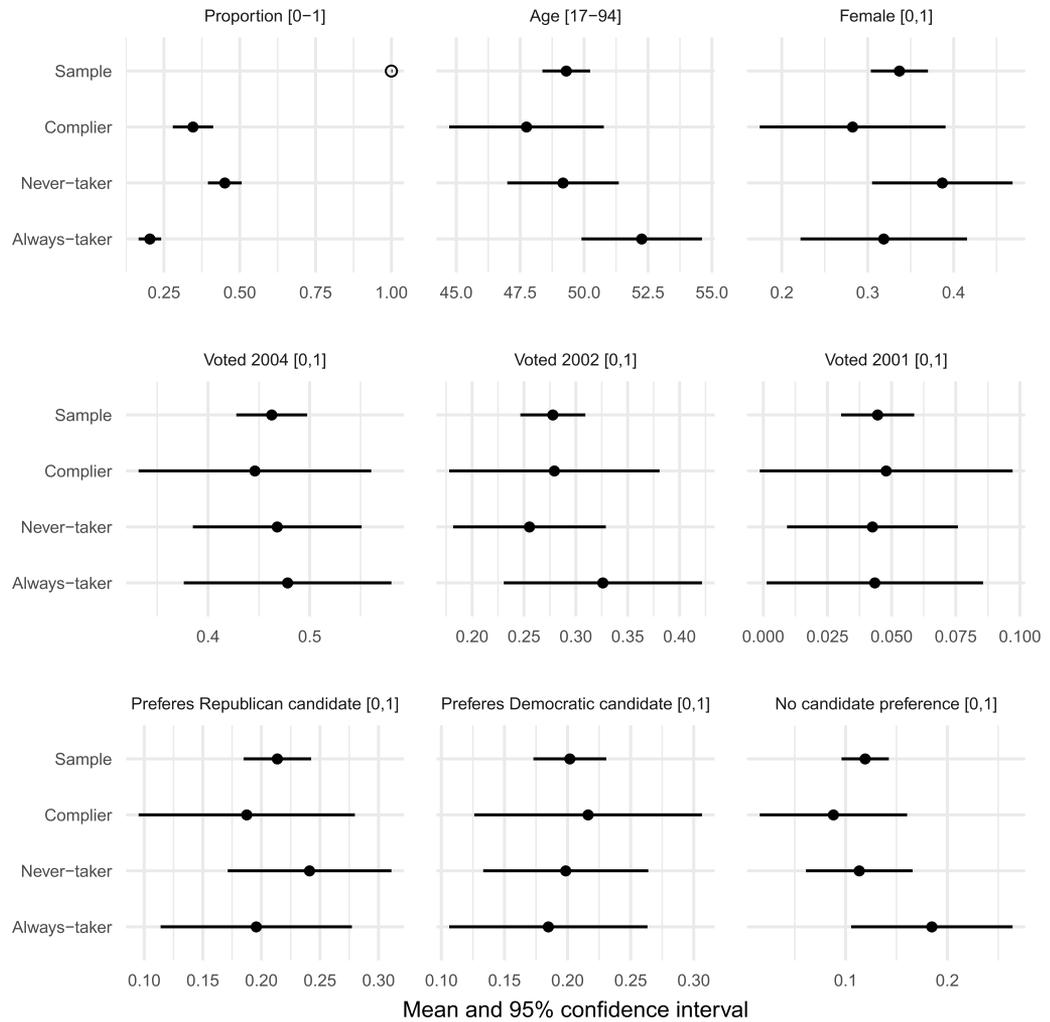
Figure 1. Descriptive statistics (mean and 95% bootstrap confidence intervals) for the complier and noncomplier subpopulations in a field experiment assessing the effects of receiving the *Washington Post* on voting behavior and public opinion. While compliers and noncompliers tend to have similar pretreatment turnout levels, compliers are younger, less likely to be female, and less likely to prefer the Republican candidate in the 2005 Virgina gubernatorial election.

affirmative action (Albertson and Lawrence 2009). In this context, we find politically meaningful and statistically significant differences in the level of political interest and frequency of media consumption between compliers and never-takers.

5 Conclusion

In this letter, we introduced a simple method of profiling compliers, always-takers, and never-takers based on their (pretreatment) covariate characteristics. Like many prior studies on IV, we focused on a case with a binary treatment and a scalar, binary instrument. In principle, our proposed method could be generalized to categorical or continuous treatments, and nonbinary (and even multivariate) instruments, by considering all combinations of instrument and treatment levels. However, the compliers will likely change across instrument and treatment levels, which creates a heretofore unsolved aggregation problem (see Abadie 2003). For the case of a binary instrument and a categorical or continuous treatment, one could dichotomize the treatment variable, and profile compliers and noncompliers at the treatment level at which the first stage is strongest. A similar method could be used to dichotomize a categorical or continuous instrument.

When resorting to this approximation, researchers should keep in mind that compliers and noncompliers who comply at different instrument and treatment levels might look different.

At least for the binary treatment, binary instrument case considered here, we recommend that researchers make it a standard practice to augment any IV analysis with reporting descriptive statistics of pretreatment covariates for the complier and noncomplier samples along with their shares. These statistics should form the basis of an informed discussion of the differences between compliers, always-takers, and never-takers. Such a discussion will increase our understanding of how "local" the LATE is, and forms the first step toward gauging the extent to which the findings derived for compliers can or cannot be generalized to other strata. We hope that this paper, in conjunction with the accompanying, easy-to-use software package `ivdesc` (for `R` available at CRAN and for `STATA` at `http://github.com/sumtxt/ivdesc`), facilitates the adoption of this practice.

Acknowledgements

Data Availability Statement

The replication materials for this paper can be found at Marbach and Hangartner (2019).

Supplementary material

For supplementary material accompanying this paper, please visit
https://doi.org/10.1017/pan.2019.48.

References

Abadie, A. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics* 113(2):231–263.

Abdulkadiroğlu, A., J. Angrist, and P. Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." *Econometrica* 82(1):137–196.

Abdulkadiroğlu, A., J. D. Angrist, P. D. Hull, and P. A. Pathak. 2016. "Charters without Lotteries: Testing Takeovers in New Orleans and Boston." *American Economic Review* 106(7):1878–1920.

Albertson, B., and A. Lawrence. 2009. "After the Credits Roll: The Long-term Effects of Educational Television on Public Knowledge and Attitudes." *American Politics Research* 37(2):275–300.

Angrist, J. D., G. W. Imbens, and D. B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–455.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Aronow, P. M., and A. Carnegie. 2013. "Beyond LATE: Estimation of the Average Treatment Effect with an Instrumental Variable." *Political Analysis* 21(4):492–506.

Baiocchi, M., J. Cheng, and D. S. Small. 2014. "Instrumental Variable Methods For Causal Inference." *Statistics in Medicine* 33(13):2297–2340.

Campbell, A., P. E. Converse, W. E. Miller, and D. E. Stokes. 1980. *The American Voter*. Chicago: University of Chicago Press.

Dahl, G. B., A. R. Kostøl, and M. Mogstad. 2014. "Family Welfare Cultures." *The Quarterly Journal of Economics* 129(4):1711–1752.

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

Deshpande, M. 2016. "Does Welfare Inhibit Success? The Long-term Effects of Removing Low-income Youth from the Disability Rolls." *American Economic Review* 106(11):3300–3330.

Dobbie, W., J. Goldin, and C. S. Yang. 2018. "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *American Economic Review* 108(2):201–240.

Dunning, T. 2012. *Natural Experiments in the Social Sciences. A Design-based Approach*. Cambridge: Cambridge University Press.

Frangakis, C. E., and D. B. Rubin. 2004. "Principal Stratification in Causal Inference." *Biometrics* 58(1):21–29.

Gelber, A., A. Isen, and J. B. Kessler. 2015. "The Effects of Youth Employment: Evidence From New York City Lotteries." *The Quarterly Journal of Economics* 131(1):423–460.

Gerber, A. S., D. Karlan, and D. Bergan. 2009. "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions." *American Economic Journal: Applied Economics* 1(2):35–52.

Haavelmo, T. 1943. "The Statistical Implications of a System of Simultaneous Equations." *Econometrica* 11(1):1–12.

Heckman, J. J., and S. Urzua. 2010. "Comparing IV With Structural Models: What Simple IV Can and Cannot Identify." *Journal of Econometrics* 156(1):27–37.

Huang, Z., L. Li, G. Ma, and L. C. Xu. 2017. "Hayek, Local Information, And Commanding Heights: Decentralizing State-owned Enterprises in China." *American Economic Review* 107(8):2455–2478.

Imbens, G. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2010)." *Journal of Economic Literature* 48(2):399–423.

Jacob, B. A., M. Kapustin, and J. Ludwig. 2014. "The Impact of Housing Assistance on Child Outcomes: Evidence From a Randomized Housing Lottery." *The Quarterly Journal of Economics* 130(1):465–506.

Manski, C. F. 1988. *Analog Estimation Methods in Econometrics*. New York: Chapman and Hall.

Marbach, M., and D. Hangartner. 2019. "Replication Data For: Profiling Compliers and Non-compliers For Instrumental-Variable Analysis." https://doi.org/10.7910/DVN/TRLTPY, Harvard Dataverse, V1.

Pianzola, J., A. H. Trechsel, K. Vassil, G. Schwerdt R., and M. Alvarez. 2019. "The Impact of Personalized Information on Vote Intention: Evidence From a Randomized Field Experiment." *The Journal of Politics* (forthcoming).

Pinotti, P. 2017. "Clicking on Heaven's Door: The Effect of Immigrant Legalization on Crime." *American Economic Review* 107(1):138–168.

Solis, A. 2017. "Credit Access and College Enrollment." *Journal of Political Economy* 125(2):562–622.

Sovey, A. J., and D. P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55(1):188–200.