COARSENING BIAS: HOW COARSE TREATMENT MEASUREMENT UPWARDLY BIASES INSTRUMENTAL VARIABLE ESTIMATES

JOHN MARSHALL*

FEBRUARY 2016

Political scientists increasingly use instrumental variable (IV) methods, and must often choose between operationalizing their endogenous treatment variable as discrete or continuous. For theoretical and data availability reasons, researchers frequently coarsen treatments with multiple intensities (e.g. treating a continuous treatment as binary). I show how such coarsening can substantially upwardly bias IV estimates by subtly violating the exclusion restriction assumption, and demonstrate that the extent of this bias depends upon the first stage and underlying causal response function. However, standard IV methods using a treatment where multiple intensities are affected by the instrument—even when fine-grained measurement at every intensity is not possible—recover a consistent causal estimate without requiring a stronger exclusion restriction assumption. These analytical insights are illustrated in the context of identifying the long-run effect of high school education on voting Conservative in Great Britain. I demonstrate that coarsening years of schooling into an indicator for completing high school upwardly biases the IV estimate by a factor of three.

*PhD candidate, Department of Government, Harvard University. jlmarsh@fas.harvard.edu. I thank Matt Blackwell, John Bullock, Anthony Fowler, Andy Hall, Torben Iversen, Horacio Larreguy, Rakeen Mabud, Daniel Moskowitz, Arthur Spirling, Brandon Stewart, Dustin Tingley, and Tess Wise for illuminating discussions or useful comments. Replication materials are available online as Marshall (2016*b*).

1 Introduction

Instrumental variables are becoming a standard component of the political scientist's methodological toolkit (Sovey and Green 2011). Instrumental variable (IV) analyses have illuminated the effects of political institutions on economic development (Acemoglu, Johnson and Robinson 2001), international trade agreements on foreign direct investment (Büthe and Milner 2008), and campaign spending on election outcomes (Gerber 1998). Given that an appropriate instrument can identify important causal relationships that cannot be easily disentangled, it is not surprising to find that the number of articles published in the *American Journal of Political Science* (AJPS) and the *American Political Science Review* (APSR) using IV techniques has almost doubled over the last decade (see Figure A1 in the Online Appendix).

Although best practice for using IV methods is now receiving greater scrutiny (e.g. Angrist and Pischke 2008; Dunning 2008; Sovey and Green 2011), this article illustrates a potentially severe but rarely-recognized source of bias: how coarsening a continuous or multi-valued (endogenous) treatment variable can substantially upwardly bias IV estimates.[1] This bias is particularly important given that such coarsening may often appear appealing to applied researchers in cases where: (1) the researcher believes that coarsening aids interpretation, or avoids imposing linearity on a statistical relationship; (2) theory suggests that the treatment effect may be non-linear; or (3) more granular measures of the treatment are unavailable. Since 2005, 36% of AJPS and APSR publications using IV methods instrument for binary or ordinal treatments with few categories, and may be susceptible to coarsening bias.[2] Furthermore, coarsening bias counters the perceived wisdom—based on the case of classical measurement error—that IV methods necessarily help to correct for measurement error in the treatment variable.[3]

[1]Following Angrist, Imbens and Rubin (1996), I refer to the endogenous variable as the "treatment."

[2]I count ordinal treatments with five or fewer categories.

[3]Kane, Rouse and Staiger (1999) make a similar point for categorical treatment variables measured with non-classical error. However, this article shows that IV estimates are upwardly biased even when a categorical treatment variable is correctly measured.

Building on Angrist and Imbens's (1995) analysis of IV methods when the treatment can take multiple ordered values (or intensities), I show analytically and demonstrate empirically how coarsening an endogenous treatment variable can substantially bias IV estimates. Intuitively, coarsening bias arises when an instrument affects the intensity of the underlying treatment, which in turn affects the outcome of interest, but does not register a change in the treatment when it is operationalized as a binary variable.[4] For example, a dummy for completing high school would fail to register the effects of school leaving laws that encouraged some students to stay in school for an additional year *without* completing high school. But by grouping together multiple years of schooling (or treatment intensities) that each affect the outcome, coarsening falsely creates the impression that the sum of the effects at each intensity can be attributed to completing high school. Coarsening bias therefore violates the exclusion restriction assumption underpinning IV identification strategies (e.g. Angrist, Imbens and Rubin 1996), because the instrument affects the outcome through an avenue not captured by the treatment when operationalized as a dichotomous variable.

I demonstrate that coarsening bias is especially large when: (1) any change in treatment intensity affects the outcome, e.g. if the treatment's true effect is linear; and (2) when the instrument has large effects, or a large "first stage," on receiving treatment intensities other than the coarsened treatment threshold.[5] In general, only when the effect of a treatment is concentrated at a single intensity such as completing high school (among the intensities affected by the instrument) *and* the researcher is able to both recognize and measure the value of the treatment where this occurs, will the IV estimate associated with a coarsened treatment variable be consistently estimated. These conditions constitute a demanding assumption that I refer to as the *strong exclusion restriction*.

In many observational studies this strong exclusion restriction may not hold. Consider the

[4]The argument equally applies when only a single value of an ordinal treatment is affected by the instrument because the treatment effectively serves as a dummy variable.

[5]Angrist and Imbens (1995) only flag the second condition, and thus do not discuss the key role of the causal response function, or its interaction with the first stage. This article also shows how coarsening bias arises from a subtle exclusion restriction and how identification requires a stronger exclusion restriction depending upon both the first stage and underlying causal response function. I also exploit a rare applied opportunity to disentangle and quantify coarsening bias.

case of Pierskalla and Hollenbach (2013), who use communication regulations as instruments to identify the effect of a binary measure of local cell phone coverage on collective action. However, if favorable regulations increase cell phone usage without affecting their coarse measure of cell phone coverage, coarsening bias may explain why their IV estimates are 20 times larger than the corresponding OLS estimates. Similarly, Milligan, Moretti and Oreopoulos (2004) use the compulsory school laws of U.S. states to instrument for completing high school, and report that completing high school increases turnout by 30-40 percentage points and the likelihood that an individual follows politics by 40-85 percentage points. These estimates would be upwardly biased if these laws increased education levels below the point of completing high school, and such lower levels of education also affect political outcomes.

In experimental studies, the extent of coarsening bias depends upon the type of endogenous variable that a randomized instrument affects. Experiments inducing respondents to uptake truly binary treatments are not affected by this bias. For example, get-out-the-vote canvassing is unlikely to impact respondents that did not answer the door (e.g. Gerber and Green 2000). However, for Gerber, Huber and Washington (2010), who in one specification instrument for an individual's partisan identification and find effects 15 times larger than their corresponding OLS estimates, upward bias may occur if their randomized mailing causes voters to move toward a political party without passing the threshold required to register a new partisan identification.[6]

However, a consistent causal estimate can still be recovered without requiring that the strong exclusion restriction holds. When the treatment is coded as a multi-valued variable, where the instrument affects multiple treatment intensities, standard IV methods consistently estimate the *local average per-unit treatment effect*, or the average causal effect of a unit increase in the treatment among compliers (Angrist and Imbens 1995). My analysis offers novel justifications for this linearization, which is typically used when the treatment is believed to linearly affect the outcome. First, although a slightly different quantity of interest is returned, this approach always provides a

[6] Gerber, Huber and Washington (2010) also consider a 7-point partisanship scale, which is unlikely to be biased.

consistent estimate when the instrument affects at least two levels of the treatment—even when the treatment is non-linearly related to the outcome. Second, without the strong exclusion restriction, the researcher need not correctly specify the functional form relating the treatment to the outcome. Finally, under certain conditions, this approach is robust even without observing all treatment intensities. Nevertheless, when the treatment's effects are highly non-linear across intensities, the intent to treat (ITT) effect estimated by the reduced form may be preferred to an IV estimate relying on assumptions about how the treatment effect varies by intensity for its interpretation.

While coarsening bias is clear in theory, its threat to inference in applied settings is difficult to gauge. I illustrate its importance in the context of Marshall's (2016*a*) study of late high school education's downstream effect on vote choice in Great Britain. While the study uses the consistent linearization approach, this article demonstrates that instrumenting for an indicator for completing high school would have overestimated the causal effect of late high school education on voting Conservative later in life by a factor of three.[7] Several properties ensure that this example offers a rare opportunity to empirically differentiate coarsening bias from other potential explanations for large IV estimates. The results suggest that coarsening bias can explain why IV estimates are often orders of magnitude larger than the analogous OLS or reduced form estimates. To avoid drawing biased inferences, which could cause policy-makers to mistakenly adopt relatively ineffective policies, this article thus demonstrates that it is imperative that political scientists become aware of how coarsening an endogenous treatment variable can dramatically inflate IV estimates.

2 Instrumental variables with coarsened treatments

Given that many interesting treatment variables cannot be easily randomized, IV approaches offer an appealing means for researchers to estimate causal relationships where they suspect that the

[7] Replication data and code can be found on the *Political Analysis* dataverse (Marshall 2016*b*).

treatment is subject to endogeneity (Sovey and Green 2011).[8] The key idea is that a researcher finds an "instrument" which (1) affects the level of the (endogenous) treatment variable, but (2) does not itself affect the outcome except by affecting the treatment. In my running example of schooling, raising the minimum school leaving age is likely to keep students in school longer, but is unlikely to affect political preferences except by keeping students in school longer.

After identifying a plausible instrument, researchers must decide how to operationalize their endogenous treatment variable. For example, should education be coded as years of schooling or an indicator for completing high school? This article shows that the decision to coarsen a treatment can undermine the goal of IV estimation—consistent estimation of a treatment's causal effect. Failing to consistently estimate an effect means that, even as the sample size becomes large, the IV estimate will not converge to the true value and confidence intervals will exclude the true value with probability approaching 1.[9] This section analyzes this inconsistency, focusing on the simplest case where there is a binary instrument and the treatment is coarsened into a binary indicator.[10]

2.1 Setup and notation

IV analyses start from observing an instrument, a treatment, and an outcome. The binary instrument, for each observation $i = 1,...,N$, is denoted by $Z_i \in \{0,1\}$. The endogenous treatment intensity of observation i, $T_i \in \{1,...,J\}$, takes one of J ordered intensities. As J becomes large, the analysis generalizes to essentially continuous treatments. Finally, Y_i is i's outcome of interest.

Following Angrist, Imbens and Rubin (1996), I examine the IV estimator in term of heterogeneous potential outcomes. Before defining potential outcomes, we first assert that the stable unit treatment value assumption (SUTVA) holds, which implies that i's potential outcomes are not related to the instrument and treatment values received by any other individual. Given SUTVA,

[8] Another motivation for using IV is to address classical measurement error, which attenuates OLS estimates. This article shows that this motivation does not apply to non-classical measurement error induced by coarsening.

[9] IV estimators are biased but consistent in finite samples (Staiger and Stock 1997).

[10] All results naturally extend to multi-valued instruments, multiple instruments, multi-valued coarsened treatments, and the inclusion of control variables.

we may write $T_i(Z_i) \equiv T_{iZ}$ as i's potential outcomes of T_i for any value of the instrument Z_i that i receives, and $Y_i(Z_i, T_i(Z_i)) \equiv Y_{iZT_Z}$ correspondingly denotes i's potential outcomes of Y_i.[11]

In addition to SUTVA (A1), IV estimation relies on four additional assumptions. First, that the instrument is independent of potential outcomes and potential treatment intensities (A2). This follows from the (conditional) random assignment of an instrument, such as a leaving age reform that alters incentives to remain in school across otherwise similar cohorts. Second, that there exists a "first stage" (A3), such that the instrument affects the intensity of the treatment i receives. Third, monotonicity (A4) requires that, for all individuals, the instrument either never decreases or never increases the treatment intensity. This rules out "defiers" that leave school earlier in response to raising the leaving age. Fourth, the *weak* exclusion restriction requires that Z_i only affects Y_i through the treatment T_i (A5). Consequently, $Y_{izt} = Y_{it}$ for any z and t. The diagram in Figure 1(a) represents this weak exclusion restriction graphically, marking by "X" the exclusion of the possibility that a schooling leaving age reform affects political preferences through a path other than the additional education that an individual receives.

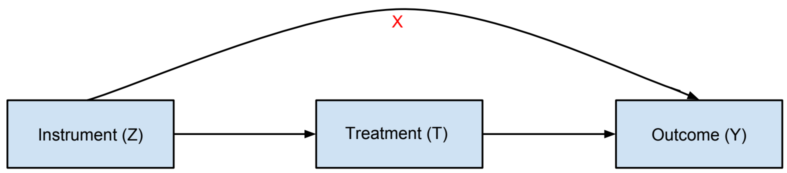
These standard assumptions, which are explained in greater detail elsewhere (e.g. Angrist, Imbens and Rubin 1996; Angrist and Pischke 2008; Sovey and Green 2011), are formalized below:

A1. **Stable unit treatment value assumption**: for all $i = 1, ..., N$, and where Z_{-i} and T_{-i} represent the vectors of instrument and treatment assignments for all observations except i, (1) for all w and w', $T_i(Z_i, Z_{-i} = w) = T_i(Z_i, Z_{-i} = w')$; and (2) for all w, w', v, and v', $Y_i(Z_i, T_i(Z_i), Z_{-i} = w, T_{-i}(w) = v) = Y_i(Z_i, T_i(Z_i), Z_{-i} = w', T_{-i}(w') = v')$.

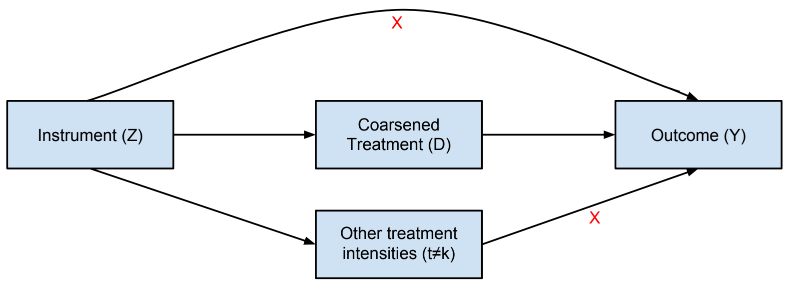
A2. **Instrument independence**: for all $i = 1, ..., N$, T_{i0}, T_{i1}, Y_{i0T_0}, and Y_{i1T_1} are jointly independent of Z_i.

A3. **First stage**: $\mathbb{E}[T_i|Z_i = 1] - \mathbb{E}[T_i|Z_i = 0] \neq 0$.

[11] Given SUTVA, observed outcomes relate to potential outcomes through $T_i = Z_i T_{i1} + (1 - Z_i) T_{i0}$ and $Y_i = Z_i Y_{i1T_1} + (1 - Z_i) Y_{i0T_0}$.



(a) Weak exclusion restriction



(b) Strong exclusion restriction

Figure 1: Graphical representation of weak and strong exclusion restrictions

A4. **Monotonicity**: for all $i = 1,...,N$, $T_{i1} - T_{i0} \geq 0$ or $T_{i1} - T_{i0} \leq 0$.

A5. **Weak exclusion restriction**: for all $i = 1,...,N$, $Y_{izt} = Y_{iz't}$ for any t and all z and z'.

Without loss of generality, I analyze the case where $T_{i1} - T_{i0} \geq 0$.

For any value t of the treatment, the local average treatment effect (LATE) of moving from treatment intensity $t-1$ to treatment intensity t because of the instrument is defined as $\beta_t \equiv \mathbb{E}[Y_{it} - Y_{it-1} | T_{i1} \geq t > T_{i0}]$. $Y_{it} - Y_{it-1}$ defines the difference in the outcome resulting from receiving treatment intensity t as opposed to intensity $t-1$, while $T_{i1} \geq t > T_{i0}$ defines this effect for the set of compliers that only reach intensity t when they receive the instrument. Collecting the LATE across all intensities defines the *causal response function* (CRF), and thus describes how each level of the treatment affects the outcome. The shape of the CRF in any given empirical application will prove crucial for understanding when coarsening is problematic.

2.2 Coarsening bias

Coarsening bias may occur when the researcher, whether by choice or constrained by data availability, coarsens their measure of the treatment. In particular, in the hope of identifying the LATE of obtaining intensity k beyond only obtaining the preceding level $k-1$, or β_k, the treatment intensity T_i is partitioned by defining the indicator $D_{ik} \equiv 1(T_i \geq k)$.[12] This coarsened binary variable indicates whether an individual receives at least treatment intensity k, and could represent the effect of completing high school beyond completing the penultimate grade of high school.

The coarsened IV estimator for the effect of D_{ik} on Y_i, denoted $\hat{\beta}_k^{IV}$, divides the reduced form effect of Z_i on Y_i by the first stage effect of Z_i on D_{ik} (see e.g. Angrist and Pischke 2008). Separate causal estimates of the reduced form and first stage are identified under assumptions A1 and A2. Further utilizing assumptions A4 and A5, the probability limit of the coarsened IV estimator can be expressed as the weighted sum of the causal effect for compliers moving from intensity $t-1$ to t, β_t, for each such interval (see Angrist and Imbens 1995 and the proof in the Online Appendix):

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}_k^{IV} \equiv \frac{\mathbb{E}[Y_i|Z_i=1]-\mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[D_{ik}|Z_i=1]-\mathbb{E}[D_{ik}|Z_i=0]} = \frac{\sum_{t=2}^{J} p_t\beta_t}{p_k}, \tag{1}$$

where $\operatorname*{plim}_{N\to\infty}$ denotes the probability limit as the sample sizes approaches ∞. Importantly, $p_t \equiv \Pr(T_{i1} \geq t > T_{i0}) = \mathbb{E}[D_{it}|Z_i=1]-\mathbb{E}[D_{it}|Z_i=0]$ denotes the probability that an individual only reaches treatment intensity t because they received the instrument $Z_i=1$, and thus represents the proportion of compliers at treatment intensity t in the population.[13] For example, p_{11} captures the probability that raising the school leaving age induces an individual that would have stayed in school for less than 11 years without the instrument to stay in school for at least 11 years. Each p_t, including p_k, can be estimated by regressing an indicator for $T_i \geq t$ on Z_i (Angrist and Imbens

[12] Angrist and Imbens (1995) describe this as an "incorrectly coded binary treatment."

[13] Where the instrument induces i to receive more than one additional intensity, i is counted in all first stages, so contributes once to each p_t (Angrist and Imbens 1995). When estimating the LAPTE (see below), the effect is thus effectively split between intensities

1995). The limit is well-defined under A3, provided $p_k > 0$. In the case of schooling, equation (1) shows that the coarsened IV estimator for the effect of completing high school converges to the sum of the effects of reaching *each* additional year of schooling—not just completing high school—weighted by the number of individuals completing each additional year of schooling only because they received the instrument.

However, under the standard assumptions, the coarsened IV estimator rarely consistently estimates our quantity of interest: the LATE of obtaining treatment intensity k. Inspection of equation (1) shows that $\hat{\beta}_k^{IV}$ converges to β_k when $\sum_{t \neq k} p_t \beta_t = 0$. Therefore, $\hat{\beta}_k^{IV}$ only consistently estimates β_k in four special cases when:

1. The instrument only affects reaching intensity k; or $p_t = 0, \forall t \neq k$.
2. The effect at all intensities other than k is zero; or $\beta_t = 0, \forall t \neq k$.
3. One of the preceding conditions holds for each intensity t, ensuring that $p_t \beta_t = 0, \forall t \neq k$.
4. The direction of the effects (weighted by p_t) differ across intensities, but ultimately cancel out; or $\sum_{t \neq k} p_t \beta_t = 0$.

These special cases are often implausible in practice. First, as the number of intensities increases, or as T_i becomes effectively continuous, it becomes increasingly implausible that only one intensity is affected. Second, as the discussion of the examples in the introduction suggest, it is often hard to believe that only one particular intensity of the treatment affects the outcome.

The following proposition summarizes these insights:

Proposition 1. *(Coarsening bias) Assume A1, A2, A3, A4 and A5 hold, and that $p_k > 0$. The probability limit of the coarsened IV estimator $\hat{\beta}_k^{IV}$ can then be expressed as:*

$$\operatorname*{plim}_{N \to \infty} \hat{\beta}_k^{IV} = \beta_k + \frac{\sum\limits_{t \neq k} p_t \beta_t}{p_k}. \tag{2}$$

Provided $sign(\beta_k) = sign(\beta_t)$ *for all* $t \neq k$ *where* $p_t \neq 0$*, the limit of the coarsened IV estimator accentuates the true causal effect:* $|\beta_k| \leq |\underset{N\to\infty}{\text{plim}}\ \hat{\beta}_k^{IV}|$.

Proof: All proofs are provided in the Online Appendix. ■

Provided that each treatment intensity causes Y_i to shift in the same direction (i.e. the CRF is monotonic in T_i), Proposition 1 further demonstrates that the coarsened IV estimator is *upwardly* biased in magnitude. This upward bias is increasing in both $p_t\,/\,p_k$ and $|\beta_t|$, for any $t \neq k$. In other words, coarsening bias is greater when the first stage is relatively large at intensities other than k and the causal effect at intensities other than k is large. If $\beta_t \in [a,b]$, for all $t \neq k$, and $p_t \geq 0$, then the bounds on the true value of β_k are $\beta_k \in \left[\hat{\beta}_k^{IV} - sb, \hat{\beta}_k^{IV} - sa\right]$, where $s \equiv \sum_{t\neq k} p_t\,/\,p_k$ is the relative strength of the first stage at intensities other than k.

The general inconsistency of the coarsened IV estimator reflects a subtle exclusion restriction violation that arises from coarsening T_i into D_{ik}. Assumption A5 requires that Z_i does not affect Y_i through avenues other than T_i. However, coarsening T_i into D_{ik} allows for T_i to affect Y_i without going through D_{ik}, but without violating A5. Consequently, while any effect of Z_i on Y_i through T_i is registered in the reduced form (the numerator of equation (1)), the first stage (the denominator) only registers cases where Z_i induces i to pass the threshold used to define the treatment indicator D_{ik} (i.e. moving from intensity $k-1$ to k). Coarsening bias thus arises if values of T_i other than k, such as years of high school before completing high school, also affect Y_i, because such changes in the reduced form are not captured in the first stage.

Consistent estimation of the effect of D_{ik} requires a stronger assumption. The *strong* exclusion restriction (A5*) that Z_i only affects Y_i through D_{ik} is sufficient:[14]

A5*. **Strong exclusion restriction**: for all $i = 1,..,N$, all t such that $p_t \neq 0$, and all z and z', (1) $Y_{izt} = Y_{iz't'}$ for all $t,t' \geq k$; and (2) $Y_{izt} = Y_{iz't'}$ for all $t,t' < k$.

[14]This captures the first, second and third of the four special cases described above. The second is probably the most empirically relevant, given that all β_t's typically go in the same direction and most instruments induce $p_t \neq 0$ for some $t \neq k$.

As illustrated in Figure 1(b), this assumption is more demanding than A5 because, in addition to requiring that the instrument only affect the outcome by altering the intensity of the treatment, it also requires that the researcher coarsen the treatment such that $\beta_t = 0$ for all intensities $t \neq k$ that are affected by the instrument.

The next proposition establishes the consistency of the coarsened IV estimator when A5* holds, and thus demonstrates the importance of the seemingly minor distinction between assumptions A5 and A5* when T_i is coarsened into D_{ik}.

Proposition 2. *(Consistency under the strong exclusion restriction) If $p_k > 0$ and assumptions A1, A2, A3, A4, and A5* hold, then $\hat{\beta}_k^{IV}$ consistently estimates β_k.*

If we cannot find an instrument that only affects completion of high school, Proposition 2 shows that consistent estimation of the effect of completing high school using the coarsened IV estimator is only possible when other levels of schooling impacted by the instrument do not affect the outcome.

2.3 Linearization as an alternative to coarsening

If coarsening an endogenous treatment can substantially upwardly bias IV estimates, what alternative strategies are available? In general, the CRF cannot be identified (see Abadie 2003). Although there exist non-linear and semi-parametric IV estimation strategies (e.g. Abadie 2003; Newey and Powell 2003), such approaches rely on stronger assumptions, require large amounts of data, and are technically challenging to estimate. However, a simple alternative to coarsening is to operationalize the treatment as a linear intensity, and thus produce a complier-weighted linear approximation to the CRF. Although this is the default approach for many researchers, its desirable theoretical properties, and the cases when it performs best, have not been clearly articulated.

The linearized IV estimator replaces D_{ik} with T_i in the first stage. The first stage could thus regress years of education, rather than an indicator for completing high school, on Z_i. However, by using T_i in the first stage to re-scale the reduced form estimate, Angrist and Imbens (1995) show that the estimand changes and instead identifies the local average per-unit treatment effect (LAPTE):

$$\underset{N\to\infty}{\text{plim}}\ \hat{\beta}^{IV}_{LAPTE} \equiv \frac{\mathbb{E}[Y_i|Z_i=1]-\mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[T_i|Z_i=1]-\mathbb{E}[T_i|Z_i=0]} = \frac{\sum_{t=2}^{J} p_t\beta_t}{\sum_{t=2}^{J} p_t}. \tag{3}$$

The LAPTE is thus a weighted linear approximation: the causal effect at each treatment intensity is weighted by the proportion of compliers at that intensity. Therefore, if the instrument primarily induces students to complete high school, the LAPTE disproportionately weights the effect of completing high school.

The LAPTE estimator is a more robust approach than coarsening in three important respects, which provide important justifications for researchers already using this approach. First, only the weaker exclusion restriction (A5) is required for consistent estimation of the LAPTE. The weaker assumptions required reflect the fact that the LAPTE "corrects" the first stage associated with the coarsened estimator by capturing all effects of the instrument on the endogenous treatment. We can see this by noting that only the denominator differs between equations (1) and (3). Provided that the instrument increases or decreases all affected treatment intensities, the signs of the LAPTE and coarsened estimators are always the same. Second, and consequently, even when the true CRF is not exactly linear, the LAPTE estimator represents a consistent linear approximation.

Third, the linearized IV approach can be robust even without observing all intensities. If the J observed intensities represent a coarsening of the true intervals (e.g. because T_i is continuous), a linear causal effect can still be recovered provided that the intervals between intensity categories are equally spaced.

Proposition 3. *(Robustness of LAPTE to missing intensity categories) Let only J equally-spaced categories of T_i be observed when there are in fact αJ equally-spaced categories, where $\alpha > 1$*

is finite and αJ is an integer. Denote $\hat{\beta}^{IV,J}_{LAPTE}$ and $\hat{\beta}^{IV,\alpha J}_{LAPTE}$ respectively as the IV estimators in the observed sample (denoted by superscript J) and unobserved sample (denoted by superscript αJ). Let assumptions A1-A5 hold, and assume $p_t > 0$ for at least two intensities. If the effect of T_i is linear such that $\beta^J_j = \tau$ for all intervals j, then $\underset{N\to\infty}{\text{plim}}\ \hat{\beta}^{IV,J}_{LAPTE} = \alpha \underset{N\to\infty}{\text{plim}}\ \hat{\beta}^{IV,\alpha J}_{LAPTE}$.

Consequently, obtaining the coefficient on the quantity of interest only requires an adjustment by factor α to identify the average linear causal effect for any unit interval.

In Proposition 3, a first stage is required for at least two treatment intensities to ensure that the IV estimate averages across coarsened categories. Without estimates at two different intensities to "draw a line through," IV estimates would be equally susceptible to coarsening bias. Consequently, if treatment measurement is sufficiently coarse that the instrument only affects a single intensity, researchers may require a new dataset, or a second dataset drawn from the same population that can separately estimate the first stage (see Angrist and Pischke 2008) . If such alternatives are not feasible, researchers should focus on estimating the reduced form ITT effect.

2.4 Coarsening in practice: CRFs, first stage weights, and bias

Proposition 1 demonstrated that the extent of coarsening bias depends upon the first stage and the shape of the causal response function. Specifically, bias depends upon the LATE at different treatment intensities (β_t) weighted by the relative effects of the instrument on the treatment intensities that do not define the coarsened treatment ($\sum_{t\neq k} p_t \,/\, p_k$). Since the CRF is not known in advance, it is essential to understand the types of causal relationships for which the strong exclusion restriction required for Proposition 2 is tenable, and what causal quantities the linear approximation can recover when only the weak exclusion restriction is tenable. This subsection compares the coarsened and linearized IV approaches to estimating "single jump," linear, and non-linear CRFs.

2.4.1 Single jumps in the CRF

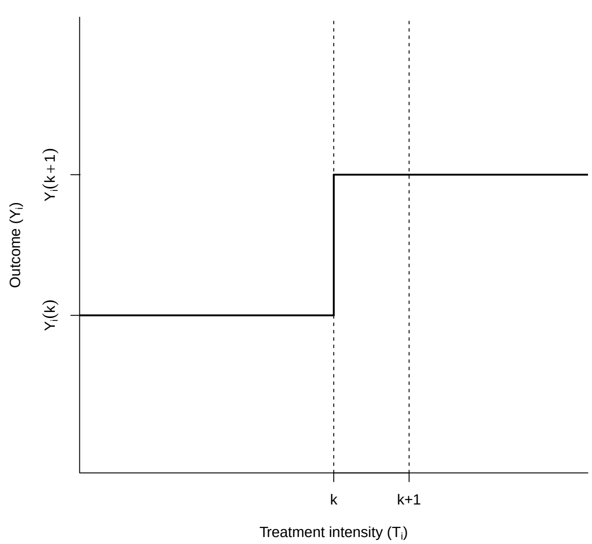
When the CRF exhibits a single jump, coarsening can be appropriate. Figure 2(a) depicts the case where no level of schooling other than completing high school (i.e. reaching treatment intensity k) affects the outcome. Provided that the researcher is able to correctly identify intensity k—the only point at which there is a (positive) causal effect in the figure—as the key jump, then β_k can be consistently estimated using the coarsened IV estimator $\hat{\beta}_k^{IV}$ when a suitable instrument ensures $p_k > 0$. This works because $\beta_t = 0$ for all $t \neq k$ for this CRF, and thus the coarsened IV estimator is consistent regardless of whether $p_t > 0$ for some other $t \neq k$.

In practice, however, it is hard to know whether k correctly captures the true jump in the CRF. In general, tipping points are not straight-forward to predict. If the researcher incorrectly surmises that $k+1$ is the correct threshold, at best they fail to detect the existence of the effect of intensity k but consistently identify no effect at $k+1$. In the simple example of Figure 2(a), where $\beta_t = 0$ for all $t \neq k$, the researcher correctly concludes that $\beta_{k+1} = 0$ only if their instrument does not induce subjects to reach intensity k. In other words, $p_k = 0$ ensures a consistent estimate of a quantity that was probably not of primary interest. However, when $p_k > 0$ and $p_{k+1} > 0$, the coarsened IV estimator will produce the following inconsistent estimate of the LATE at intensity $k+1$:

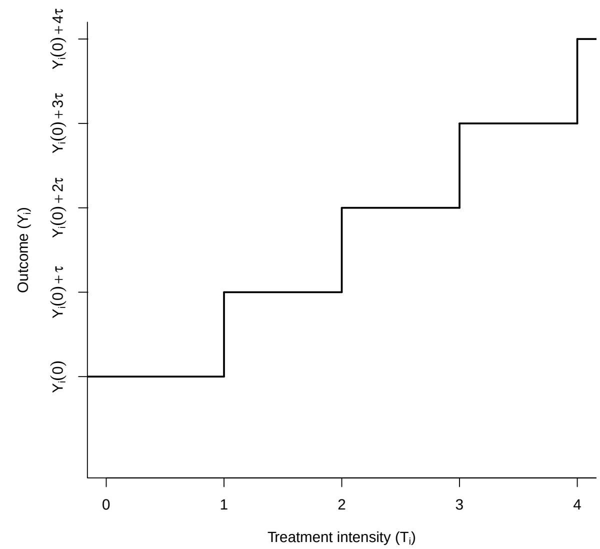
$$\operatorname*{plim}_{N\to\infty} \hat{\beta}_{k+1}^{IV} = \frac{p_k \beta_k}{p_{k+1}} > 0. \tag{4}$$

While approximately correct in the sense that there is a causal effect *nearby*, this estimator both wrongly attributes the effect to intensity $k+1$ and does not even consistently estimate β_k unless $p_{k+1} = p_k$.

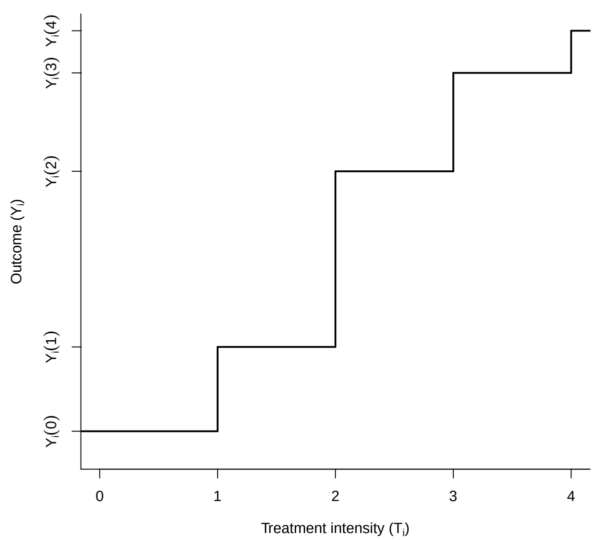
In experiments where subjects cannot be partially treated it is easier to determine clear cutoffs. As noted above, in the case of Gerber and Green's (2000) get-out-the-vote canvassing, knocking on a door is only likely to affect respondents that opened the door to receive the treatment. But even experiments can be hard to evaluate if there is partial compliance, such that individuals can



(a) Single jump in the CRF



(b) Linear CRF



(c) Non-linear CRF

Figure 2: Examples of causal response functions

experience some of the treatment without being designated as treated. This could occur, for example, if subjects partially learn about a treatment by reading an email heading, but are not registered as treated because they did not open the email. If such information affects political behavior, then the strong exclusion restriction is violated.

The linearized IV estimator with an uncoarsened treatment faces different issues when the CRF is singularly discontinuous. The linearization introduces a trade-off between consistent estimation and the quantity of interest. On the one hand, $\hat{\beta}^{IV}_{LAPTE}$ provides a consistent causal estimate regardless of the CRF, and is thus robust to misspecifying the threshold of the discontinuity. On the other hand, weighting the effect at each intensity by its relative first stage will underestimate the effect of the sharp jump at intensity k when intensities other than k do not affect the outcome. If additional years of schooling do not affect an outcome except upon completion of high school, the LAPTE will positively weight the effect of completing high school *and* the zero effects at years other than completing high school. While this remains a well-defined causal quantity, it may differ from the researcher's primary quantity of interest. However, CRFs which approximate sharp jumps may be empirically rare, and hard to compellingly identify in practice.

2.4.2 Linear CRFs

The bias associated with the coarsened IV estimator for category k can be particularly large when the true CRF is linear. As depicted in Figure 2(b), a linear CRF entails a casual effect of size $\beta_t = \tau$ at each treatment intensity interval. The probability limit of the coarsened IV estimator is:

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}^{IV}_k = \beta_k + s\tau \tag{5}$$

where the coarsening bias in the second term is increasing in s, the first stage at intensities other than k relative to the first stage at k. In particular, more than one half of all compliers must reach intensity k for the coarsened IV estimator to converge upon an estimate less than double the size of the true effect.[15] This concern also increases with how close the treatment intensity categories are to one another (i.e. increases in the number of categories J), because it becomes increasingly implausible that any instrument could only increase the likelihood of receiving a particular intensity, so s becomes increasingly large.

However, when the CRF is linear and T_i is observable, the LAPTE estimator is highly desirable. From inspection of equation (3), it is easy to see that when the true effect is τ at each interval, τ is recovered exactly by $\hat{\beta}^{IV}_{LAPTE}$ as the sample becomes large. Moreover, Proposition 3 shows that even when some categories of T_i are unobserved, the LAPTE still consistently estimates τ. Except in the rare instance where $p_t = 0$, for all $t \neq k$, the LAPTE strictly improves upon the inconsistent coarsened IV estimator. Although this is not surprising, this fact is important because many multi-valued treatments such as education are believed to exert relatively linear effects. Furthermore, similar advantages apply when the CRF is non-linear.

2.4.3 Non-linear CRFs

In practice, many CRFs likely reflect more complex non-linear relationship than a single jump at a given intensity. For example, each additional unit of schooling increases subsequent income, but certain thresholds such as completing high school or university may have particularly large effects. Figure 2(c) illustrates such a case, although innumerable other non-linear CRFs can be imagined. When are the coarsened and LAPTE estimators appropriate approximations to non-linear CRFs?

As demonstrated above, the coarsened IV estimate will generally be upwardly biased whenever the CRF is not discontinuous at a single coarsened treatment intensity. Conversely, the LAPTE remains a consistent estimate of the complier-weighted average across intervals regardless of the CRF's functional form. Therefore, even when the causal effect varies substantially across intensities, the complier-weighted LAPTE is informative about the average effect among compliers in the population that the sample represents, even though it may not provide the non-linear quantity

[15] Note that $s = \sum_{t \neq k} p_t / p_k = (\overline{p} - p_k)/p_k < 1$, only when $p_k > \overline{p}/2$, where $\overline{p} \equiv \sum_{t=2}^{J} p_t$.

the researcher would ideally estimate. Furthermore, in the limit, the LAPTE always yields a coefficient at least as small in magnitude as the coarsened estimator, and may thus be regarded as a more conservative approach (Angrist and Imbens 1995). Given the extent of upward bias that coarsening can introduce, and the difficulty of efficiently implementing non-linear IV estimates, this trade-off may be preferable in many applications.

However, in some cases, the ITT effect of the instrument may be the only informative estimate. When the instrument affects a wide range of intensities and the CRF significantly departs from linearity, the coarsened IV estimator may be substantially biased, and plausible bounds may be hard to establish. Moreover, an estimate of the LAPTE provides misleading estimates of the treatment's effect when the instrument affects many intensities of a non-monotonic CRF. In an extreme example, a null LAPTE estimate could mask the impact of a treatment where the LATE changes sign across intensities. In contrast, the ITT—estimated by the reduced form—captures the overall impact of the instrument in the population. The ITT does not make any claims about the endogenous treatment's effect, but its interpretation does not rely on assumptions about the CRF. Moreover, as in the case of a school leaving age reform, the ITT is often the policy-relevant quantity of interest. In such a context, and given that policy-makers often weight different types of compliers differently, researchers should report estimates of all the p_t's alongside the ITT. Although this cannot identify the LATE in different parts of the CRF, it may help readers to appropriately scale the ITT.

2.5 Implications for applied research

The preceding analysis clearly shows that coarsening an endogenous treatment can substantially upwardly bias IV estimates, except in the special cases where the instrument only induces subjects to reach the intensity where the treatment is coarsened or the CRF only registers an effect at a single and correctly-identified intensity. Therefore, as the simulation analysis in the Online Appendix illustrates, estimating the LAPTE is more appropriate for most CRFs. Furthermore, even when the CRF is not linear, treating the endogenous treatment as linear has three important advantages: the

LAPTE (1) does not suffer from coarsening bias, and always provide a consistent causal estimate; (2) does not rely on researchers correctly specifying the intensity where the effect is large; and (3) is often robust when the researcher cannot observe all treatment intensities. Nevertheless, the easily-interpretable and often policy-relevant ITT estimate of the instrument's effect may become increasingly appealing when researchers are faced with a non-linear CRF for which average effects across intensities are difficult to interpret.

In general, researchers seeking to implement an IV analysis must rely on their theoretical intuitions and descriptive data—including the reduced form relationship, separate first stage regressions, and the (endogenous) OLS relationship—to determine the appropriate specification in the standard case where only a single instrument is available. Under special circumstances where multiple instruments are available, I show below that a sharper empirical assessment is possible. This will allow me to quantify the coarsening bias associated with instrumenting for a dichotomous measure of schooling.

3 Observational application: high school education's effect on voting Conservative in Great Britain

Building on a recent study by Marshall (2016*a*), this section illustrates the risk of coarsening bias in an applied setting. Marshall (2016*a*) uses Britain's 1947 school leaving reform to instrument for measures of schooling, and finds that each additional year of late high school significantly increases the probability that an individual votes for the right-wing Conservative party later in life. Figure 3 shows that the reform, which increased the school leaving age in Britain from 14 to 15 for students who reached the age of 14 after April 1st 1947, dramatically increased the proportion of students staying in school until age 15. Importantly, the reform also increased the proportion leaving at age 16, and thus completing high school, but did not impact higher levels of education.[16]

[16]Both the reform and research design are discussed in detail in Marshall (2016*a*).

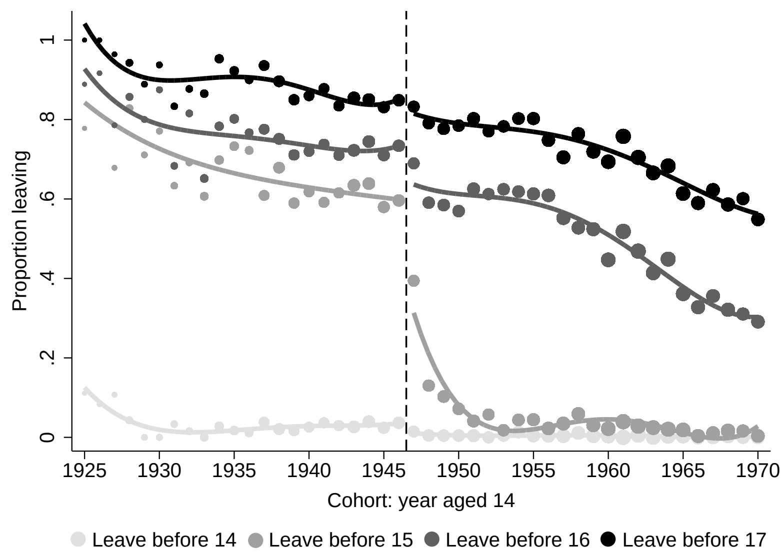


Figure 3: 1947 compulsory schooling reform and student leaving age by cohort

Notes: Data are from the British Election Survey. Curves represent fourth-order polynomial fits. Grey dots are birth-year cohort averages, and their size reflects their weight in the sample.

The study's findings are not vulnerable to coarsening bias, because it focuses on the LAPTE. However, several particular features of the study's setting enable me to quantify the coarsening bias that would arise if the author instead instrumented for an indicator for completing high school. By differentiating the effects of coarsening from other factors that could explain large IV estimates, the objective of this section is to demonstrate that coarsening can substantially bias IV estimates.

3.1 Estimation: coarsening and the LAPTE

I use the British Election Survey survey data and follow the research design described in Marshall (2016*a*).[17] The outcome is an indicator for the 34% of respondents that reported voting for the Conservative party at the most recent election. The instrument is an indicator for the 49% respondents that turned 14 in 1947 or later, and were thus affected by the higher school leaving age. Depart-

[17]The sample differs slightly from Marshall (2016*a*) because I focus on the eight elections from 1979 to 2010 for which data on the outcome, instrument and treatment measures are available, and restrict attention to working age respondents (aged below 70).

ing from the original study, schooling—the endogenous treatment variable in this application—is measured *both* as a linear intensity and a coarsened indicator. The linear intensity is the number of years of schooling. The coarsened treatment is an indicator for the 64% of respondents that completed high school. Since the purpose of this application is to analyze the decision to coarsen, not assess the results substantively, operationalization is described in the Online Appendix.

To estimate the effect of late high school education on voting Conservative, Marshall (2016*a*) uses Britain's 1947 leaving age reform as an instrument for schooling in the context of a "fuzzy" regression discontinuity (RD) design. In particular, whether an individual was affected by the reform is a discontinuous function of their birth year cohort. However, since the reform could not force every student to remain in school for (at least) an additional year, the 1947 reform is used as an instrument that discontinuously increases the *probability* that a student received greater schooling. Figure 3 indicates a strong and monotonic first stage, while Marshall (2016*a*) provides evidence that Conservative voting is continuous across cohorts in all covariates other than the school leaving age at the reform discontinuity (implying assumption A2 above) and that the reform only affected Conservative voting by increasing years of schooling.[18]

The design entails estimating the following structural equation for individuals i from cohort c:

$$Vote\ Conservative_{ic} = \beta Schooling_{ic} + f(Birth\ year_c) + \varepsilon_{ic}, \tag{6}$$

where f is a flexible function of birth year cohort. Specifically, following Marshall (2016*a*), I estimate local linear regressions (LLRs) using observations within the Imbens and Kalyanaraman (2012) optimal bandwidth (of 11.514 cohorts) and weight using a triangular kernel. The corresponding first stage is given by:

$$Schooling_{ic} = \alpha Post\ 1947\ reform_c + f(Birth\ year_c) + \varepsilon_{ic}. \tag{7}$$

[18]The weak exclusion restriction need not hold for the purposes of identifying the effects of coarsening, since the exclusion restriction similarly impacts both approaches to defining the treatment.

Conditioning on $f(Birth\ year_c)$, the combination of equations (6) and (11) approximately estimate equations (3) and (1), depending upon the measure of $Schooling_{ic}$. Instrumenting for years of schooling corresponds to the LAPTE in equation (3), while instrumenting for the indicator for completing high school aims corresponds with the coarsened IV estimator in equation (1).

3.2 Illustrating coarsening bias

The first stage estimates in Table 1 verify that the 1947 reform instrument substantially increased schooling. Column (1) shows that, on average, the reform increased years of schooling by 0.61 years. The large F statistic confirms the strong first stage. Column (2) shows that—in addition to keeping students in school until 15—the reform also increased the probability of completing high school by 15 percentage points. This increase is also highly statistically significant (F statistic of 24.8), and thus confirms that any bias in the IV estimates does not simply reflect weak instruments (Staiger and Stock 1997).

I now examine the IV estimates of schooling's effect on voting behavior. I first instrument for years of schooling to estimate the LAPTE. Reinforcing Marshall's (2016*a*) findings, the estimate of the LAPTE in column (3) shows that each additional year of late high school significantly increases a complier's probability of voting Conservative in later life by 11 percentage points.

To evaluate the impact of coarsening the treatment variable on the IV estimates, I turn to the case where schooling is dichotomized as an indicator for completing high school. Column (4) reports that voters induced to complete high school by the reform are fully 46 percentage points more likely to vote Conservative in later life. This statistically significant estimate seems implausibly large, particularly when considering that compliers represent a relatively large share of the student population (so are unlikely to be an anomalous group experiencing unusually large effects), and given that the average effect of an additional year's schooling is four times smaller.

Furthermore, the conditions under which coarsening bias is large—highlighted in equation (2)—appear to be satisfied. First, the reform predominantly kept students in school until age 15,

Table 1: Estimates of schooling's effect on voting Conservative

	Years of schooling	Completed high school	Vote Con.	Vote Con.
	LLR	LLR	LLR IV	LLR IV
	(1)	(2)	(3)	(4)
Post 1947 reform	0.614***	0.150***		
	(0.135)	(0.030)		
Years of schooling			0.113**	
			(0.054)	
Completed high school				0.464**
				(0.209)
Observations	4,820	4,820	4,820	4,820
First stage F statistic	20.7	24.8	20.7	24.8

Notes: Specifications (1) and (2) present the first stage estimates where years of schooling and completing high school are respectively the endogenous treatment variable. Specifications (3) and (4) are respectively the IV estimates for years of schooling and completing high school. All specifications are estimated using local linear regression (LLR) with a triangular kernel and the Imbens and Kalyanaraman (2012) optimal bandwidth of 11.514. Robust standard errors are given in parentheses. * denotes $p < 0.1$, * denotes $p < 0.05$, *** denotes $p < 0.01$.

but did not compel most students to complete high school (i.e. large p_t/p_k). Second, there are reasons to believe that an additional year of late high school, without completing high school, imparts skills that ultimately affect political preferences (i.e. $\beta_t \neq 0, t \neq k$). Third, since both IV estimates rely on the weak exclusion restriction, differences cannot be explained by the 1947 reform affecting vote choice through channels other than schooling. There are thus good reasons to believe that the estimate for completing high school suffers from substantial coarsening bias.

Nevertheless, although such large effects of completing high school are surprising, they are not completely implausible. Alternatively, the huge effect of completing high school could perhaps be squared with the smaller average effect of an additional year of schooling if the effect of staying until 15 is actually zero. To confirm that the large estimate actually reflects coarsening bias, I also exploit Britain's 1972 reform raising the school leaving age from 15 to 16 to separate the effects of the penultimate and final year of high school.[19] As the Online Appendix shows, the 1972 reform also significantly increased the proportion of students leaving school at age 16 without affecting further education. Consequently, I can use both reforms as instruments to identify the effects of both the penultimate year of high school and completing high school.[20] Given that the availability of two instruments that only affect two levels of the treatment is uncommon, this application represents a rare opportunity to demonstrate and quantify coarsening bias.

The results, described in detail in the Online Appendix to save space, show that completing the penultimate year of high school increases the probability of voting Conservative later in life by 10 percentage points, while completing high school similarly increases Conservative voting by a further 17 percentage points. This implies that the effect of additional late high school on Conservative voting is relatively linear. These estimates are consistent with the LAPTE estimate of 11 percentage points for each additional year in column (3) of Table 1. However, by failing to return the coarsened IV estimate for completing high school of 0.46, the results demonstrate that the strong exclusion restriction fails to hold. In fact, coarsening is responsible for upwardly biasing the IV estimate by a factor of nearly three.

4 Conclusion

This article shows that coarsening bias should be a major concern for applied researchers using IV methods. Coarsening bias arises where a treatment with multiple intensities is transformed into a binary indicator or short scale. Except in special cases that are likely to be rare in empirical applications, such coarsening subtly violates the exclusion restriction requiring that the instrument only affects the outcome through the measured treatment. Although IV methods can address classical measurement error in a treatment variables, coarsening bias shows that non-classical measurement

[19]Using years of schooling also represents a coarsening of days. However, the estimates below do not suggest a bias when estimating the penultimate and ultimate years of high school separately.

[20]This is because $p_t \approx 0$ for all t other than leaving school at 15 and leaving school at 16.

error can dramatically inflate IV estimates.

As demonstrated in the case of high school's long-run effects on vote choice in Britain, the upward bias can be considerable. I show that using a binary indicator for completing high school overestimated the causal effect of the final year of high school on voting Conservative by nearly three times. The extent of coarsening bias clearly has important implications for the interpretation of academic findings, but could also substantially mislead policy-makers seeking to choose the most efficacious policy. As IV methods become a standard part of a political scientist's methodological toolkit, it is thus essential that researchers become aware of the risks of coarsening bias.

Coarsening bias may be common in applied research. Coarsening an endogenous treatment variable may appear to be theoretically appealing or easier to interpret, to avoid strong linearity assumptions, or be necessitated by data unavailability. However, in general, only when the treatment's effects are concentrated at a certain level of the treatment *and* the researcher is able to pinpoint the specific threshold where this causal effect occurs can coarsening consistently estimate the desired causal quantity. In most cases, it is better to avoid invoking the strong exclusion restriction required to coarsen an endogenous treatment, and instead consistently estimate the LAPTE by measuring the treatment intensity using a linear treatment variable. Unlike coarsening bias, which arises when the true causal response function is not discontinuous, this approach provides consistent causal estimates even when the underlying CRF is not linear. Moreover, because this approach is also robust when some categories cannot be observed, it provides an effective solution in the frequent case that a good measure of the treatment is not available. Nevertheless, where the true CRF is highly non-linear, this estimate may be misleading, and researchers are instead advised to report the ITT effect and the first stages at treatment intensity.

References

Abadie, Alberto. 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics* 113(2):231–263.

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–1401.

Angrist, Joshua D. and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity." *Journal of the American Statistical Association* 90(430):431–442.

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

Angrist, Joshua D. and Jörn-Steffan Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Büthe, Tim and Helen V. Milner. 2008. "The politics of foreign direct investment into developing countries: increasing FDI through international trade agreements?" *American Journal of Political Science* 52(4):741–762.

Dunning, Thad. 2008. "Model specification in instrumental-variables regression." *Political Analysis* 16(3):290–302.

Gerber, Alan. 1998. "Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables." *American Political Science Review* 92(2):401–411.

Gerber, Alan S. and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3):653–663.

Gerber, Alan S., Gregory A. Huber and Ebonya Washington. 2010. "Party affiliation, partisanship, and political beliefs: A field experiment." *American Political Science Review* 104(4):720–744.

Imbens, Guido W. and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *Review of Economic Studies* 79(3):933–959.

Kane, Thomas J., Cecilia Elena Rouse and Douglas Staiger. 1999. "Estimating Returns to Schooling when Schooling is Misreported." NBER Working Paper 7235.

Marshall, John. 2016*a*. "Education and voting Conservative: Evidence from a major schooling reform in Great Britain." *Journal of Politics* 78(2).

Marshall, John. 2016*b*. "Replication Data for: Coarsening bias: How coarse treatment measurement upwardly biases instrumental variable estimates." http://dx.doi.org/10.7910/DVN/J7HUX3.

Milligan, Kevin, Enrico Moretti and Philip Oreopoulos. 2004. "Does education improve citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics* 88:1667–1695.

Newey, Whitney K. and James L. Powell. 2003. "Instrumental variable estimation of nonparametric models." *Econometrica* 71(5):1565–1578.

Pierskalla, Jan H. and Florian M. Hollenbach. 2013. "Technology and collective action: The effect of cell phone coverage on political violence in Africa." *American Political Science Review* 107(2):207–224.

Sovey, Allison J. and Donald P. Green. 2011. "Instrumental variables estimation in political science: A readers' guide." *American Journal of Political Science* 55(1):188–200.

Staiger, Douglas and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3):557–586.

Online Appendix

Trends in the use of instrumental variables

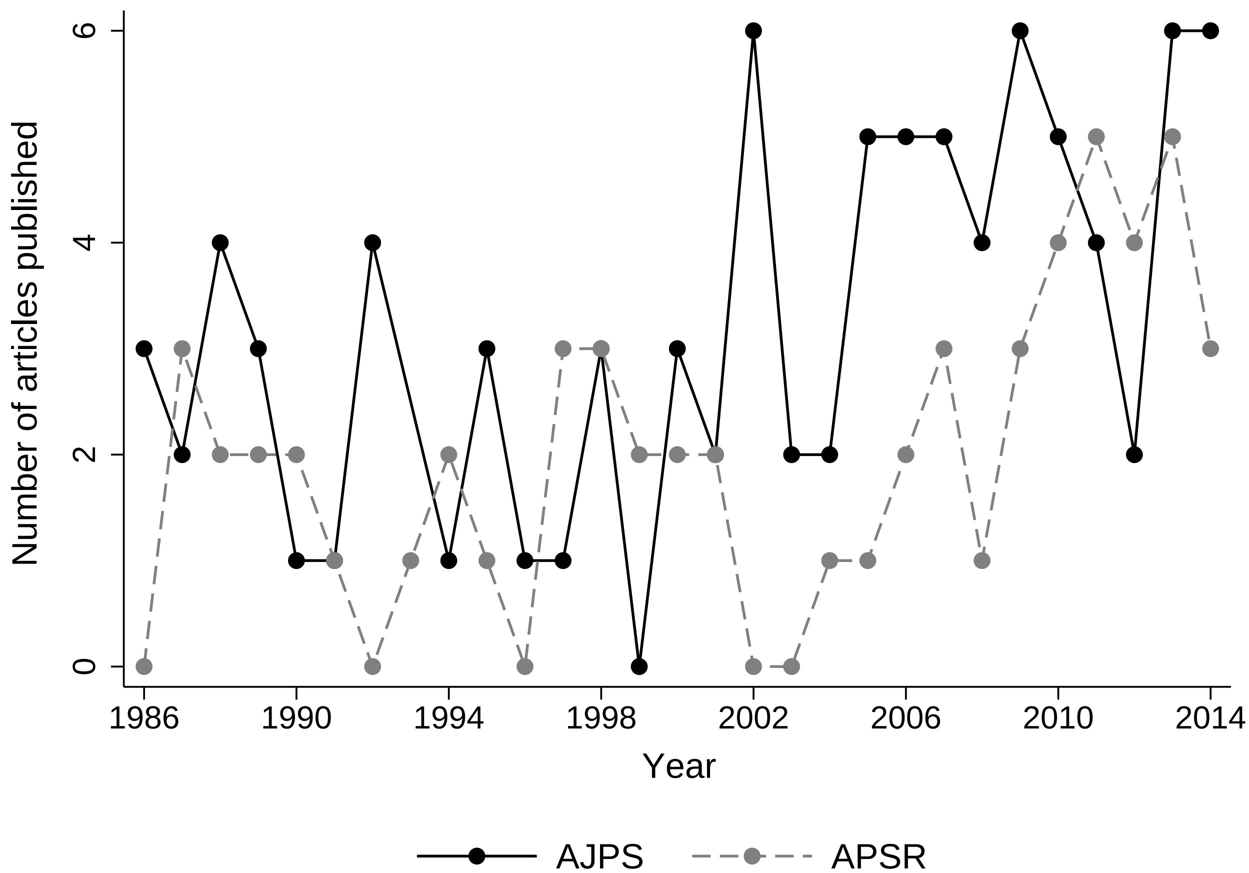


Figure A1: Annual trends in the use of instrument variable techniques in political science

Notes: Counts are based on data provided by Allison Carnegie (from Sovey and Green 2011) and the author's own reading of AJPS and APSR articles. Any reference to implementing an IV technique is included.

Proofs

Proof of Proposition 1. First note that by the law of large numbers and Slutsky's theorem:

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}_k^{IV} = \operatorname*{plim}_{N\to\infty} \left(\frac{\sum_{\{i:Z_i=1\}} Y_i - \sum_{\{i:Z_i=0\}} Y_i}{\sum_{\{i:Z_i=1\}} D_{ik} - \sum_{\{i:Z_i=0\}} D_{ik}} \right) = \frac{\mathbb{E}[Y_i|Z_i=1] - \mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[D_{ik}|Z_i=1] - \mathbb{E}[D_{ik}|Z_i=0]}.$$

By A1, the potential outcomes of Y_i and T_i can respectively be written as Y_{izt} and T_{iz}. Without loss of generality, let A4 hold such $T_{i1} - T_{i0} \geq 0$, and thus $p_t \geq 0$.[21] Following Angrist and Imbens (1995), the probability limit of the reduced form can be written as:

$$\begin{aligned}
\mathbb{E}[Y_i|Z_i=1] - \mathbb{E}[Y_i|Z_i=0] &= \mathbb{E}[Y_{i1T_{i1}} - Y_{i0T_{i0}}] \\
&= \mathbb{E}\left[\sum_{t=2}^{J} I(T_{i1} \geq t)[Y_{i1t} - Y_{i1t-1}] - \sum_{t=2}^{J} I(T_{i0} \geq t)[Y_{i0t} - Y_{i0t-1}] \right] \\
&= \mathbb{E}\left[\sum_{t=2}^{J} [I(T_{i1} \geq t) - I(T_{i0} \geq t)][Y_{it} - Y_{it-1}] \right] \\
&= \sum_{t=2}^{J} \Pr[T_{i1} \geq t > T_{i0}]\mathbb{E}[Y_{it} - Y_{it-1}|T_{i1} \geq t > T_{i0}] \\
&= \sum_{t=2}^{J} p_t \beta_t,
\end{aligned}$$

where the first line uses A2, the second decomposes the first over all treatment intensity levels, the third uses A5, and the fourth uses A4 (which implies that $T_{i1} \geq t > T_{i0}$ is either 0 or 1).

By definition, D_{ik} is purely a function of t. Using A2, the probability limit of the first stage for

[21] All results hold for $T_{i0} - T_{i1} \geq 0$, where β_T and p_t are respectively redefined as $\beta_k \equiv \mathbb{E}[Y_{ik} - Y_{ik-1}|T_{i0} \geq k > T_{i1}]$ and $p_t \equiv \Pr(T_{i0} \geq t > T_{i1})$.

the coarsened treatment indicator D_{ik} is given by:

$$\begin{aligned}
\mathbb{E}[D_{ik}|Z_i=1]-\mathbb{E}[D_{ik}|Z_i=0] &= \mathbb{E}\left[\sum_{t=2}^{J}[I(T_{i1}\geq t)-I(T_{i0}\geq t)][D_{ik}(t)-D_{ik}(t-1)]\right] \\
&= \mathbb{E}\left[I(T_{i1}\geq k)-I(T_{i0}\geq k)\right] \\
&= \Pr(T_{i1}\geq k>T_{i0}) \\
&= p_k,
\end{aligned}$$

where the second line follows from the fact that $D_{ik}(k)-D_{ik}(k-1)=1$ and $D_{ik}(t)-D_{ik}(t-1)=0,\forall t\neq k$, while the third line follows from A4.

Combining the first stage and reduced form, the probability limit of the Wald IV estimator is:

$$\operatorname*{plim}_{N\to\infty}\hat{\beta}_k^{IV}=\frac{\sum\limits_{t=2}^{J}p_t\beta_t}{p_k}=\beta_k+\frac{\sum\limits_{t=2,t\neq k}^{J}p_t\beta_t}{p_k},$$

where the final term is the bias of the estimator (beyond finite sample bias). The condition $p_k\neq 0$ ensures that β_k^{IV} is well-defined.

It is immediate that the bias is positive whenever $\sum_{t=2,t\neq k}^{J}p_t\beta_t>0$, given that A4 ensures $p_t\geq 0,\forall t$. $\beta_t>(<)0$, for all t such that $p_t\neq 0$ is a sufficient condition for positive (negative) bias. Consequently, $sign(\beta_k)=sign(\beta_t)$, for all t such that $p_t\neq 0$ implies $|\beta_k|\leq|\beta_k^{IV}|$. ■

Proof of Proposition 2. Consistency requires that $\sum_{t=2,t\neq k}^{J}p_t\beta_t=0$. I now show that A5* is a sufficient condition. Both A5 and A5* entail that $[Y_{izt}-Y_{izt-1}]=[Y_{it}-Y_{it-1}]$. Furthermore, A5* entails that $[Y_{it} - Y_{it-1}] = 0$ for all $t \neq k$ where $p_t \neq 0$. Consequently,

$$\begin{aligned}
\mathbb{E}[Y_i|Z_i=1] - \mathbb{E}[Y_i|Z_i=0] &= \mathbb{E}\left[\sum_{t=2}^{J}[I(T_{i1} \geq t) - I(T_{i0} \geq t)][Y_{it} - Y_{it-1}]\right] \\
&= \mathbb{E}\left[\sum_{t=2, t\neq k}^{J}[I(T_{i1} \geq t) - I(T_{i0} \geq t)]\right][Y_{it} - Y_{it-1}] + \\
&\quad \mathbb{E}\left[[I(T_{i1} \geq k) - I(T_{i0} \geq k)][Y_{ik} - Y_{ik-1}]\right] \\
&= p_k \beta_k,
\end{aligned}$$

where the first line follows from A5 and the third line requires A5*. Under A5*, it is thus clear that the Wald estimator then yields:

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}_k^{IV} = \frac{\mathbb{E}[Y_i|Z_i=1] - \mathbb{E}[Y_i|Z_i=0]}{\mathbb{E}[D_{ik}|Z_i=1] - \mathbb{E}[D_{ik}|Z_i=0]} = \frac{p_k \beta_t}{p_k} = \beta_k.$$

Therefore, $\hat{\beta}_k^{IV}$ is a consistent estimator under A1, A2, A3, A4 and A5* (which implies A5). ■

Proof of Proposition 3. Using the proof of Proposition 1, $\operatorname*{plim}_{N\to\infty} \hat{\beta}_{LAPTE}^{W,J} = \frac{\sum_{t=2}^{J} p_t \beta_t^J}{\sum_{t=2}^{J} p_t} = \tau$ and $\operatorname*{plim}_{N\to\infty} \hat{\beta}_{LAPTE}^{W,\alpha J} = \frac{\sum_{t=2}^{\alpha J} p_t \beta_t^{\alpha J}}{\sum_{t=2}^{\alpha J} p_t} = \tau/\alpha$, where the linearity of the causal effect at each intensity interval implies $\alpha \beta_t^{\alpha J} = \beta_t^J$. The result follows. ■

Illustrating the analytical results: simulated data

In addition to examining coarsening bias in a political application with observational data, as I do in the main paper, it is useful to also illustrate and validate the analytical results using purely simulated data that can control the true empirical relationships. The simulation results reinforce the analytical results showing that the size of the effect at each treatment intensity on the outcome (i.e. the shape of the CRF) and the strength of the first stage at intensity k relative to the first stage at all other intensities ($\sum_{t\neq k} p_t / p_k$) are critical in determining the extent of coarsening bias.

Monte Carlo simulations

The analytical insights are captured in a simple simulation framework. For each simulated dataset, I draw a random sample of 1,000 independent observations. Using complete randomization, half the observations are randomly assigned to treatment and control. The observation's endogenous treatment intensity T_i is given by the nearest round number to $Z_i + \xi_i$, where $\xi_i \sim Normal(10, \sigma^2)$. The treatment thus takes multiple integer intensities, where the mean for observations that did not receive the instrument is 10 and the mean for observations that did is 11. Absracting from noise, in terms of potential outcomes, $T_i = (1 - Z_i)T_{i0} + Z_i T_{i1} = (1 - Z_i) * 10 + Z_i * (10 + 1)$. Importantly, by increasing the variance parameter σ^2, the instrument can be allowed to exert a relatively larger effect on reaching treatment intensities other than 11. Finally, the treatment is coarsened as an indicator for receiving an intensity of at least 11: $D_i \equiv 1(T_i \geq 11)$.

I consider two types of causal relationship. I first examine a *linear CRF* where each additional intensity level increases the outcome by 0.05 units:

$$Y_i = 0.05 * \sum_{t=1}^{T} 1(T_i \geq t) + \eta_i \tag{8}$$

where $\eta_i \sim Normal(0.4, 0.2)$ determines the location of the outcome and its variation. This is held constant across CRFs. Abstracting from the noise, potential outcomes are thus $Y_{it} = Y_{i0} + 0.05t = 0.4 + 0.05t$. I also consider a *single jump CRF* where only the eleventh intensity level increases the outcome by 0.05 units:

$$Y_i = 0.05 * 1(T_i \geq 11) + \eta_i. \tag{9}$$

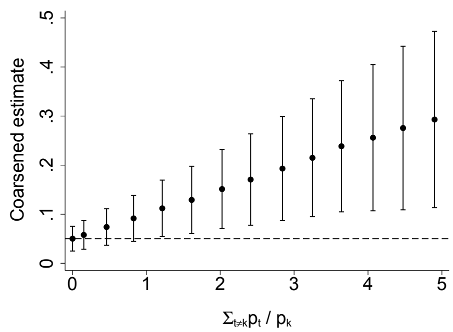
Abstracting from the noise, potential outcomes are thus $Y_{it} = Y_{i0} + 0.05t = 0.4 + 0.05I(t \geq 11)$. In both cases, the standard IV assumptions (A1-A5) hold. However, the strong exclusion restriction (A5*) only holds for the single jump CRF.

To illustrate the bias associated with coarsening, I vary σ^2 to examine how two stage least squares (2SLS) estimates for the two different CRFs depend upon the relative first stage at intensities other than k, or $\sum_{t \neq k} p_t / p_k$, which increases in σ^2. For each σ^2, I examine 1,000 simulated datasets and present the coarsened IV estimates (using the indicator D_i for reaching the eleventh intensity) and the LAPTE (treating T_i as linear).[22]

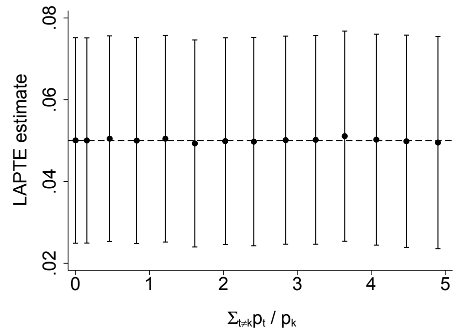
I first analyze the linear CRF. As demonstrated in Proposition 1, Figure 2(a) shows that the bias associated with coarsening the treatment can be substantial and is increasing linearly in $\sum_{t \neq k} p_t / p_k$. When only the coarsened intensity is affected by the instrument (i.e. $p_k \neq 0$ and $p_t = 0, \forall t \neq k$), or as $\sum_{t \neq k} p_t / p_k \downarrow 0$, the IV estimate is essentially unbiased. However, bias increases in the relative magnitude of the effect of the instrument on other intensities. Once only 50% (25%) of the instrument's effect occurs at intensity k, the IV estimate is fully two (four) times larger than the true causal effect. This bias reflects that fact that the instrument is inducing many changes in the outcome through intensities that are not captured by the coarsened first stage. Conversely, Figure 2(b) shows that the LAPTE estimate correctly identifies the true causal effect regardless of which intensities are affected by the instrument, and does not lose precision because each intensity is equally relevant. This is not surprising given that the LAPTE is designed precisely for this case.

Second, consider the single jump CRF, where the researcher correctly identifies the sole effect at the eleventh intensity. Reinforcing Proposition 2, Figure 2(c) confirms that coarsening the treatment unsurprisingly produces an unbiased estimate of the effect at the eleventh intensity in this rare instance where the strong exclusion restriction holds. The precision of this estimate is decreasing in the relative first stage at other (uninformative) intensities. Turning to the LAPTE estimate, Figure 2(d) demonstrates that—as shown analytically in the main paper—the LAPTE only returns the causal effect at the eleventh intensity when the instrument only affects this intensity. The size of the LAPTE declines with $\sum_{t \neq k} p_t / p_k$, as the relative weight attached to compliers with zero effects increases. However, it is important to reiterate that the LAPTE remains a consistent causal

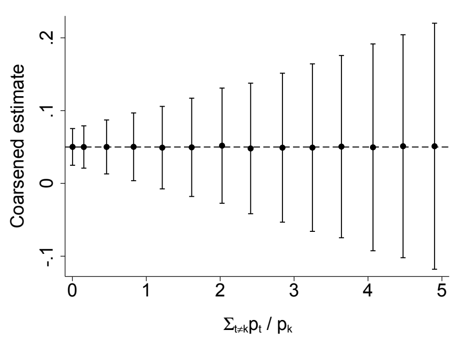
[22]Specifically, I consider 14 values of σ^2, increasing from 0.17 to 2.33 in intervals of 0.17.



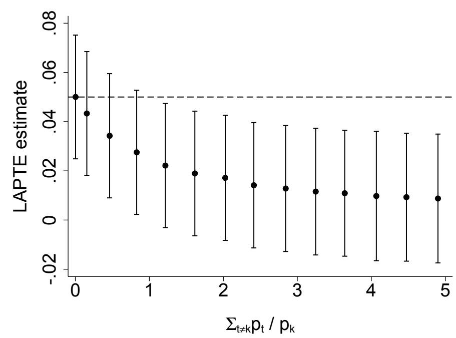
(a) Linear CRF, coarsened IV estimate



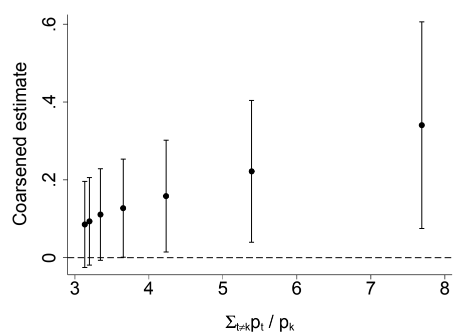
(b) Linear CRF, LAPTE estimate



(c) Single jump CRF, correctly coarsened IV estimate



(d) Single jump CRF, LAPTE estimate



(e) Single jump CRF, incorrectly coarsened IV estimate

Figure A2: Monte Carlo simulations illustrating how IV estimates of a 0.05 causal effect depend upon the CRF, coarsening, and the relative strength of the instrument on different treatment intensities (95% confidence intervals)

Notes: Estimates at each level of $\sum_{t \neq 11} p_t / p_k$ are based on 1,000 simulated datasets containing 1,000 observations. For the linear CRF, the effect of each additional intensity is 0.05. For the Single jump CRF, the effect is 0.05 only at intensity $k = 11$. When correctly coarsened, the coarsened treatment is defined as $D_i = 1(T_i \geq k)$. When incorrectly coarsened, the coarsened treatment is defined as $D_i' = 1(T_i \geq k+1)$. See the main text for the underlying equations and distributions. Point estimates are the average estimate across simulations; the horizontal dashed line denotes the true causal effect the researcher seeks to estimate. 95% confidence intervals are based on variation within and between simulated estimates.

estimate—it just weights the effects at all the intensities affected by the instrument. Furthermore, the comparison of Figures 2(c) and 2(d) shows the general results that the LAPTE will always be smaller than the coarsened IV estimate. When the CRF is neither linear nor discontinuous at a single threshold, the LAPTE thus provides both consistent and conservative estimates.

Finally, consider the same single jump CRF, but where the researcher incorrectly coarsens the treatment at intensity 12. Figure 2(e) shows that coarsening often incorrectly ascribes a significant positive effect to the twelfth intensity where no such effect exists. This occurs where, in addition to inducing subjects to receive intensity 12 (where there is no effect), the instrument also induces subjects to receive the eleventh treatment intensity where there exists a positive effect. As the figure shows, this is not simply an issue of interpretation because the points systematically fail to even recover the true effect of the eleventh intensity.

Monte Carlo simulation code

The `R` code generating the simulation results above is provided below:

```
# Load SEM package for IV estimation
library(sem)

# Set number of simulated datasets
n <- 1000

# Set total number of variance parameters to examine
largest <- 14

######## Linear CRF

# Generating matrices to hold the results
FS <- FSK <- RF <- IV1 <- IV2 <- matrix(nrow=largest,ncol=4)
poverp <- matrix(nrow=largest,ncol=2)

# Set seed
set.seed(12345)

# Loop over the variance parameter sigma^2 (=j/6)
```

```
for (j in 1:largest){
# Set variance
par <- j/6

# Create temporary holding matrices
fs <- fsk <- ols <- iv1 <- iv2 <- matrix(nrow=n,ncol=2)

# First stage (linear), first stage (coarsened), reduced form, and IV regressions  for each of the
# 1,000 datasets
for (i in 1:n){
z <- sample(rep(c(0,1),n/2))
x <- round(z + rnorm(n,10,par))
y <- 0.05*as.numeric(x>=1) + 0.05*as.numeric(x>=2) + 0.05*as.numeric(x>=3) +
  0.05*as.numeric(x>=4) + 0.05*as.numeric(x>=5) + 0.05*as.numeric(x>=6) + 0.05*as.numeric(x>=7) +
  0.05*as.numeric(x>=8) + 0.05*as.numeric(x>=9) + 0.05*as.numeric(x>=10) + 0.05*as.numeric(x>=11) +
  0.05*as.numeric(x>=12) + 0.05*as.numeric(x>=13) + 0.05*as.numeric(x>=14) + 0.05*as.numeric(x>=15) +
  0.05*as.numeric(x>=16) + 0.05*as.numeric(x>=17) + 0.05*as.numeric(x>=18) + rnorm(n,0.4,.2)
d <- as.numeric(x>=11)

fs[i,] <- summary(lm(x ~ z))$coefficients[2,1:2]
fsk[i,] <- summary(lm(d ~ z))$coefficients[2,1:2]
ols[i,] <- summary(lm(y ~ z))$coefficients[2,1:2]
iv1[i,] <- summary(tsls(y ~ d, ~ z))$coefficients[2,1:2]
iv2[i,] <- summary(tsls(y ~ x, ~ z))$coefficients[2,1:2]
}

# First stage ratio shown on the x axis
poverp <- ( mean(fs[,1]) - mean(fsk[,1]) ) / mean(fsk[,1])

FS[j,] <- cbind( par, poverp, mean(fs[,1]), mean(fs[,2])+(1+1/n)*var(fs[,1]) )
FSK[j,] <- cbind( par, poverp, mean(fsk[,1]), mean(fsk[,2])+(1+1/n)*var(fsk[,1]) )
RF[j,] <- cbind( par, poverp, mean(ols[,1]), mean(ols[,2])+(1+1/n)*var(ols[,1]) )
IV1[j,] <- cbind( par, poverp, mean(iv1[,1]), mean(iv1[,2])+(1+1/n)*var(iv1[,1]) )
IV2[j,] <- cbind( par, poverp, mean(iv2[,1]), mean(iv2[,2])+(1+1/n)*var(iv2[,1]) )
}

FS; FSK; RF; IV1; IV2

######## DISCONTINUOUS CRF
```

```
FS <- FSK <- RF <- IV1 <- IV2 <- matrix(nrow=largest,ncol=4)

set.seed(12345)

for (j in 1:largest){
par <- j/6

fs <- fsk <- ols <- iv1 <- iv2 <- matrix(nrow=n,ncol=2)

for (i in 1:n){
z <- sample(rep(c(0,1),n/2))
x <- round(z + rnorm(n,10,par))
y <- 0.05*as.numeric(x>=11) + rnorm(n,0.4,.2)
d <- as.numeric(x>=11)

fs[i,] <- summary(lm(x ~ z))$coefficients[2,1:2]
fsk[i,] <- summary(lm(d ~ z))$coefficients[2,1:2]
ols[i,] <- summary(lm(y ~ z))$coefficients[2,1:2]
iv1[i,] <- summary(tsls(y ~ d, ~ z))$coefficients[2,1:2]
iv2[i,] <- summary(tsls(y ~ x, ~ z))$coefficients[2,1:2]
}

poverp <- ( mean(fs[,1]) - mean(fsk[,1]) )/mean(fsk[,1])

FS[j,] <- cbind( par, poverp, mean(fs[,1]), mean(fs[,2])+(1+1/n)*var(fs[,1]) )
FSK[j,] <- cbind( par, poverp, mean(fsk[,1]), mean(fsk[,2])+(1+1/n)*var(fsk[,1]) )
RF[j,] <- cbind( par, poverp, mean(ols[,1]), mean(ols[,2])+(1+1/n)*var(ols[,1]) )
IV1[j,] <- cbind( par, poverp, mean(iv1[,1]), mean(iv1[,2])+(1+1/n)*var(iv1[,1]) )
IV2[j,] <- cbind( par, poverp, mean(iv2[,1]), mean(iv2[,2])+(1+1/n)*var(iv2[,1]) )
}

FS; FSK; RF; IV1; IV2

######## INCORRECTLY IDENTIFIED DISCONTINUOUS CRF

FS <- FSK <- RF <- IV1 <- IV2 <- matrix(nrow=largest,ncol=4)

set.seed(12345)
```

```
for (j in 3:20){
par <- j/12

fs <- fsk <- ols <- iv1 <- iv2 <- matrix(nrow=n,ncol=2)

for (i in 1:n){
z <- sample(rep(c(0,1),n/2))
x <- round(z + rnorm(n,10,par))
y <- 0.05*as.numeric(x>=11) + rnorm(n,0.4,.2)
d <- as.numeric(x>=12)

fs[i,] <- summary(lm(x ~ z))$coefficients[2,1:2]
fsk[i,] <- summary(lm(d ~ z))$coefficients[2,1:2]
ols[i,] <- summary(lm(y ~ z))$coefficients[2,1:2]
iv1[i,] <- summary(tsls(y ~ d, ~ z))$coefficients[2,1:2]
iv2[i,] <- summary(tsls(y ~ x, ~ z))$coefficients[2,1:2]
}

poverp <- ( mean(fs[,1]) - mean(fsk[,1]) )/mean(fsk[,1])

FS[j,] <- cbind( par, poverp, mean(fs[,1]), mean(fs[,2])+(1+1/n)*var(fs[,1]) )
FSK[j,] <- cbind( par, poverp, mean(fsk[,1]), mean(fsk[,2])+(1+1/n)*var(fsk[,1]) )
RF[j,] <- cbind( par, poverp, mean(ols[,1]), mean(ols[,2])+(1+1/n)*var(ols[,1]) )
IV1[j,] <- cbind( par, poverp, mean(iv1[,1]), mean(iv1[,2])+(1+1/n)*var(iv1[,1]) )
IV2[j,] <- cbind( par, poverp, mean(iv2[,1]), mean(iv2[,2])+(1+1/n)*var(iv2[,1]) )
}

FS; FSK; RF; IV1; IV2
```

BES survey data

All variables are from the British Election Survey (BES). The BES uses a multi-stage design, randomly selecting postal addresses from several wards from randomly sampled constituencies (stratifying by region). The BES has been conducted following every general election since 1964, although I only use the surveys since 1979 where appropriate variables are available. As noted in the main paper, the sample is restricted to working age respondents (i.e. aged 70 or below), and those aged 18 at the time of the survey. The sample is restricted to those aged below 70 given the likelihood that education affects political behavior through earned income (see Marshall 2016*a*). Summary statistics for the RD and full samples are provided in Table A1.

- *Vote Conservative*. Indicator coded one for respondents who reported voting for the Conservative party at the last general election. Only respondents which refused to respond, did not answer or did not vote were excluded.

- *Years of schooling*. Years of schooling is calculated as the age that the respondent left full time education minus five (the age at which students start formal schooling). Years of schooling is top-coded at 13 years to ensure comparability and focus on state-provided education. Indicators for 10 and 11 years of schooling are defined according to this measure.

- *Completed high school*. Indicator coded one for respondents that answered that either: (1) possess a grade 1 Certificate of Secondary Education (CSE), 5 O-levels at A-C, 5 General Certificates of Secondary Education (GCSEs) at A*-C or a lower grade on the Scottish Certification of Education (SCE); or (2) left school at age 16 or later.

- *Birth year*. Birth-year is estimated by subtracting age at the date of the survey from the year in which the survey was conducted. I then add 14 for year aged 14. Non-responses were deleted.

- *Post 1947 reform*. Indicator coded one for students aged 14 or below in 1947, and aged 15 or above in 1972.

- *Post 1972 reform*. Indicator coded one for students aged 14 or below in 1972.

- *Male*. Indicator coded one for respondents identifying as male. Non-responses were deleted.

- *Age*. Standardized age at the date of the survey.

Table A1: Summary statistics: RD and full BES samples

	RD sample (1947 reform)					**Full BES sample**				
	Obs.	Mean	Std. dev.	Min.	Max.	Obs.	Mean	Std. dev.	Min.	Max.
Dependent variable										
Vote Conservative	4,820	0.409	0.492	0	1	13,853	0.343	0.475	0	1
Endogenous treatment variables										
Years of schooling	4,820	10.317	1.417	0	13	13,853	10.982	1.436	0	13
Completed high school	4,820	0.400	0.490	0	1	13,853	0.642	0.480	0	1
Excluded instruments										
Post 1947 reform	4,820	0.589	0.492	0	1	13,853	0.493	0.500	0	1
Post 1972 reform	4,820	0.000	0.000	0	0	13,853	0.314	0.464	0	1
Pre-treatment covariates										
Male	4,820	0.474	0.499	0	1	13,853	0.466	0.499	0	1
White	4,820	0.984	0.126	0	1	13,853	0.965	0.183	0	1
Black	4,820	0.006	0.079	0	1	13,853	0.009	0.094	0	1
Asian	4,820	0.007	0.084	0	1	13,853	0.016	0.127	0	1
Age	4,820	54.732	8.323	35	69	13,853	43.365	14.073	18	69
Father manual/unskilled job	3,820	0.724	0.447	0	1	9,922	0.672	0.470	0	1
Survey	4,820	1988.912	7.342	1979	2010	13,853	1991.697	8.926	1979	2010
Birth year	4,820	1934.180	6.465	1922	1944	13,853	1948.332	16.153	1910	1992

- *Race*. Indicators coded one for respondents who respectively identify their ethnicity as white, black, or Asian (including South Asian ethnicities and Chinese).

- *Father manual/unskilled job*. Indicator coded one for respondent's who answered that their father had a manual or unskilled job.

- *Survey year*. Year in which the survey was conducted.

Using the 1972 school leaving reform to identify coarsening bias

As noted in the main text, the availability of two instruments offers the opportunity to precisely estimate the extent of coarsening bias. In particular, the availability of two instruments means that it is possible to instrument for two treatment variables—in this case completing the penultimate year of high school (i.e. $Penultimate_{ic} = 1(Schooling_{ic} = 10))$ and completing at least the final year of high school (as defined in the main paper). Since Figure 5 in the main paper and Figure A3 respectively show that the 1947 and 1972 reforms did not affect students leaving at older or younger ages, the reforms only affected whether students remained in school for the penultimate or final year of high school. Consequently, the first stage at all other levels of the treatment is zero (i.e. $p_t = 0, \forall t \neq 10, 11$). Therefore, instrumenting for two indicator variables—completing the penultimate and final years of high school—does not suffer from coarsening bias because the special case that $p_t = 0$ at all other t is satisfied. This enables me to estimate β_t for both additional years of high school, and thus to exactly identify the effect of completing high school beyond completing the penultimate year of high school. This latter quantity can be then be compared to the estimate for completing high school in the main paper to adduce the extent of coarsening bias.

To simultaneously estimate the effects of the penultimate and final year of high school, I use 2SLS to estimate the following structural equation:

$$Vote\ Conservative_{ic} = \beta_1 Penultimate_{ic} + \beta_2 Completed\ high\ school_{ic} + f(Birth\ year_c) + \varepsilon_{ic}, \quad (10)$$

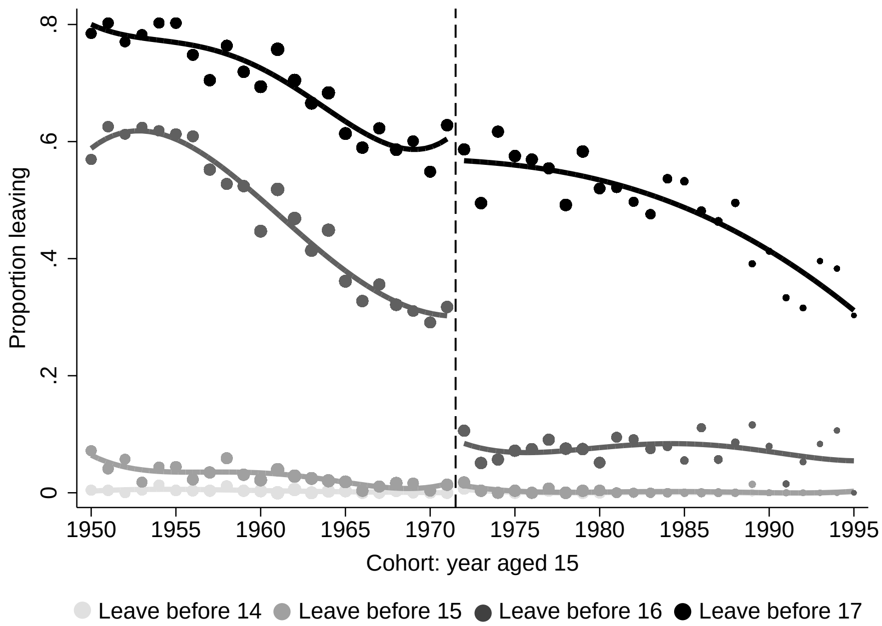


Figure A3: 1972 compulsory schooling reform and student leaving age by cohort

Notes: Data from the BES. Curves represent fourth-order polynomial fits. Grey dots are birth-year cohort averages, and their size reflects their weight in the sample.

where the first stage regressions generating exogenous variation in educational attainment are given by:

$$\mathit{Penultimate}_{ic} = \alpha_1 \mathit{Post\ 1947\ reform}_c + \alpha_2 \mathit{Post\ 1972\ reform}_c + f(\mathit{Birth\ year}_c) + \varepsilon_{ic} \tag{11}$$

$$\mathit{Completed\ high\ school}_{ic} = \gamma_1 \mathit{Post\ 1947\ reform}_c + \gamma_2 \mathit{Post\ 1972\ reform}_c + f(\mathit{Birth\ year}_c) + \varepsilon_{ic} \tag{12}$$

Because I now use two discontinuities as instruments, this system cannot be estimated using local linear regression, as in the main paper. However, I adopt a similar approach by using the full BES sample and letting f include cubic global polyomials in the running variable (birth year). These flexible polynomial terms are designed to capture general trends in Conservative support across cohorts around both reforms, and thus mimic the local linear approach in the main paper.[23]

[23]Encouragingly, column (1) of Table A2 shows that the cubic polynomials report a similar estimate for the 1947 reform as the local linear regression approach in the main paper. I also found similar results when using fourth, fifth, sixth and seventh-order polynomials.

Table A2: Using the 1972 reform to identify the extent of coarsening bias

	Vote Con. OLS (1)	Vote Con. 2SLS (2)	Vote Con. 2SLS (3)
Post 1947 reform	0.059***		
	(0.021)		
Post 1972 reform	0.079***		
	(0.030)		
Years of schooling		0.130***	
		(0.046)	
Penultimate year of high school			0.098**
			(0.039)
Completed high school			0.269**
			(0.124)
Observations	13,853	13,853	13,853
First stage F statistic		28.5	658.4/52.1

Notes: All specifications include standardized cubic polynomials in birth year. Robust standard errors in parentheses (which are always larger than errors clustered by cohort). * denotes $p < 0.1$, ** denotes $p < 0.05$, *** denotes $p < 0.01$.

Table A2 reports the reduced form and IV estimates when exploiting both reforms simultaneously. Column (1) presents the reduced form estimates, and demonstrates that both reforms significantly increase the probability that an individual votes Conservative. Consistent with the 1947 reform having a larger effect on education, the 1947 reform also has a larger effect on Conservative voting: the 1947 increased Conservative voting per cohort by 6 percentage points, while the 1972 reform added a further two percentage points. Column (2) estimates the LAPTE for an additional year of schooling by using the 1947 and 1972 reforms as instruments for years of schooling. Similarly to the main paper, which reports a 11.3 percentage point increase, the estimates suggest that an additional year increases the probability of voting Conservative later in life by 14.8 percentage points. Furthermore, a Sargan overidentification χ^2_1 test fails to reject the null hypothesis that the instruments produce different IV estimates ($p = 0.70$). This provides evidence that both reforms had similar effects on voters, and again suggests that an additional year of compulsory education substantially increases the probability of voting Conservative. The similarity of the effects means that despite using two different instruments suggests that I am able to use both instruments to separate out the effects of the penultimate and final years of schooling.

Finally, I turn to the key element of this exercise: determining the extent of coarsening bias. Column (3) estimates equation (10), and reports the estimates for completing the penultimate and final year of high school. The results indicate that the penultimate year of high school increases Conservative voting by 10 percentage points, while completing high school adds a further 17 percentage points. This implies that the effect of completing high school, over completing a lower level of schooling, is to increase the probability of voting Conservative by 17 percentage points. Therefore, this demonstrates that coarsening bias almost triples the size of this estimate (46 percentage points in the main paper). Although the coefficients come from different samples, plugging the estimates for β_t and p_t into equation (4) in the main paper yields an IV estimate of around 0.5. This is similar to the 0.46 estimate in the main paper.