The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables[†]

By Magne Mogstad, Alexander Torgovitsky, and Christopher R. Walters*

Empirical researchers often combine multiple instrumental variables (IVs) for a single treatment using two-stage least squares (2SLS). When treatment effects are heterogeneous, a common justification for including multiple IVs is that the 2SLS estimand can be given a causal interpretation as a positively weighted average of local average treatment effects (LATEs). This justification requires the well-known monotonicity condition. However, we show that with more than one instrument, this condition can only be satisfied if choice behavior is effectively homogeneous. Based on this finding, we consider the use of multiple IVs under a weaker, partial monotonicity condition. We characterize empirically verifiable sufficient and necessary conditions for the 2SLS estimand to be a positively weighted average of LATEs under partial monotonicity. We apply these results to an empirical analysis of the returns to college with multiple instruments. We show that the standard monotonicity condition is at odds with the data. Nevertheless, our empirical checks reveal that the 2SLS estimate retains a causal interpretation as a positively weighted average of the effects of college attendance among complier groups. (JEL C26, I23, I26, J24, J31, R23)

Instrumental variables (IVs) are widely used to estimate causal relationships. In practice, researchers often combine multiple IVs using two-stage least squares (2SLS). In Section I, we report a survey of empirical papers using IV that were published in leading journals since 2000. More than half of these papers report results from a specification with multiple IVs for a single treatment, typically combined using 2SLS.

The textbook motivation for combining multiple IVs is statistical efficiency. However, this requires an assumption of constant treatment effects. In contrast, allowing for heterogeneous treatment effects is a key motivation in the modern

program evaluation literature, and one which is supported by a large body of empirical work.[1] In an influential paper, Imbens and Angrist (1994; "IA" hereafter) provided an alternative justification for using 2SLS with multiple IVs, one which allows for heterogeneous treatment effects. They showed that the 2SLS estimand can be interpreted as a positively weighted average of local average treatment effects (LATEs) for subpopulations whose treatment status is affected by the instruments. This result holds for any number of instruments, as long as IA's "monotonicity" condition is satisfied.[2]

The fact that widespread empirical practice rests on the monotonicity condition raises a number of questions. What requirements does this condition place on choice behavior when there are multiple IVs? Can we expect these requirements to be satisfied? If not, is it still possible to retain a causal interpretation of the 2SLS estimand while allowing for unobserved heterogeneity in both treatment effects and choice behavior? The contribution of our paper is to answer these questions and, by doing so, offer theoretical and empirical guidance for researchers who wish to use 2SLS with multiple IVs.

In Section II, we begin our theoretical analysis by showing that the monotonicity condition cannot be satisfied with more than one instrument without restricting choice behavior to be effectively homogeneous.[3] For example, if the treatment is college attendance and the instruments are tuition and proximity, the monotonicity condition requires all individuals to respond more to tuition than to proximity, or vice versa. This is a concerning implication; it shows that appealing to IA monotonicity to justify combining multiple IVs comes at the cost of assuming homogeneity in choice behavior.

Motivated by this observation, we then consider a weaker, *partial* version of the monotonicity condition. The partial monotonicity condition is that the IA monotonicity condition is satisfied for each instrument separately, holding all of the other instruments fixed. We show that partial monotonicity is satisfied if each instrument makes every individual weakly more likely to choose treatment. For example, a sufficient condition for partial monotonicity is that all individuals are at least as likely to attend college if they live closer to a college or face lower tuitions. However, unlike the IA monotonicity condition, partial monotonicity does not restrict heterogeneity in the relative impacts of different instruments; it allows for some individuals to respond more to tuition than to proximity, and for others to respond more to proximity than to tuition.

In Section III, we show that even though partial monotonicity permits heterogeneous choice behavior, it can still be sufficient to give the 2SLS estimand a causal

[1] Heckman (2001) compiled a list of empirical evidence on heterogeneous treatment effects prior to 2001. More recent papers that find evidence of unobserved heterogeneity in treatment effects include Bitler, Gelbach, and Hoynes (2006); Doyle (2007); Moffitt (2008); Carneiro and Lee (2009); Firpo, Fortin, and Lemieux (2009); Carneiro, Heckman, and Vytlacil (2011); Maestas, Mullen, and Strand (2013); Bitler, Hoynes, and Domina (2014); Walters (2018); Felfe and Lalive (2018); French and Song (2014); Havnes and Mogstad (2015); Kirkeboen, Leuven, and Mogstad (2016); Kline and Walters (2016); Carneiro, Lokshin, and Umapathi (2016); Cornelissen et al. (2018); Nybom (2017); and Brinch, Mogstad, and Wiswall (2017).

[2] In discussing the use of multiple IVs instead of a single binary instrument, Angrist and Pischke (2009, p. 173) write that "The econometric tool remains 2SLS and the interpretation remains fundamentally similar to the basic LATE result, with a few bells and whistles."

[3] Our analysis here builds upon points made by Heckman and Vytlacil (2005, Section 6) and Heckman, Urzua, and Vytlacil (2006, Section III.D).

interpretation as a positively weighted average of LATEs. Moreover, we characterize sufficient and necessary conditions for this interpretation, and show that the conditions can be checked empirically. In a simple case with two binary instruments this amounts to verifying that the *unconditional* correlations between the treatment and each instrument have the expected sign. More generally, our results provide a set of tests that researchers can report alongside 2SLS estimates formed from multiple IVs. These tests are implemented in our companion Stata module, mivcausal.[4]

In Section IV, we apply our results to estimating the returns to college attendance, using the same data as Carneiro, Heckman, and Vytlacil (2011; CHV hereafter). We focus on two instruments used by those authors: local labor market conditions and distance to college. We show that these instruments generate different estimated marginal treatment effect (MTE) schedules, a finding which is inconsistent with IA monotonicity. Nevertheless, our empirical checks show that under partial monotonicity the 2SLS estimand retains an interpretation as a positively weighted average of the causal effects of college attendance among complier groups. This finding illustrates how assumptions weaker than IA monotonicity can be sufficient for the modest empirical goal of using 2SLS to recover a weighted average of LATEs.

I. Survey on the Use of Multiple Instruments

Empirical researchers often combine multiple IVs using 2SLS. To document this practice, we searched the Web of Science database for articles published between January 2000 and October 2018 containing the words "instrument" or "instrumental variable" in the abstract, title, or topic words. We restricted the search to the following five journals: *Journal of Political Economy*, *American Economic Review*, *Quarterly Journal of Economics*, *Review of Economic Studies*, and *Econometrica*. In total, 266 articles matched our search criteria (Mogstad, Torgovitsky, and Walters 2021).

We restrict our attention to the 122 of these papers that use at least one IV in an empirical application. The other 144 papers not included were either methodological papers without an empirical application, or were papers that used the word "instrument" in a different context, such as to describe a policy or financial instrument. Table 1 categorizes the remaining 122 papers based on the number of instruments, the number of IVs compared to the number of endogenous variables, the choice of IV estimator, and the relationship between the instruments.[5]

Column 1 counts papers containing at least one specification in the main body of the paper that included more than one IV. The number of IVs is defined as the number of variables excluded from the outcome equation. If a paper has multiple IVs, but each reported specification used no more than one instrument, then we do not count it in column 1. The bottom row of column 1 reveals that more than half of the papers in our sample used more than one instrument.

In column 2 of Table 1, we count the set of papers that have at least one specification in the main body of the paper with more IVs than endogenous variables. This is typically referred to as an "overidentified" specification in a constant-effects model. Comparing column 2 to column 1, we see that most papers that used more than one

---

[4] The module is available for download from the Boston College Statistical Software Components archive.
[5] Online Appendix A reports the survey results separately by journal.

Table 1—Multiple IVs in Top Economics Journals

	Papers that use multiple IVs (1)	Papers with more IVs than endogenous variables (2)	Papers with more IVs than endogenous variables using 2SLS/GMM (3)	Relationship between IVs		
				Case A (4)	Case B (5)	Case C (6)
Number	65	57	52	35	7	10
Percentage	53	47	43	67	13	19

Notes: This table displays the results of a survey on the use of IVs in articles published in the *American Economic Review*, *Quarterly Journal of Economics*, *Journal of Political Economy*, *Econometrica*, and the *Review of Economic Studies*. The sample includes 122 articles published between January 2000 and October 2018 containing the words "instrument" or "instrumental variable" in the abstract, title, or topic words and using at least one IV in an empirical application. Columns 1–3 report numbers and percentages of these papers that use multiple IVs. Columns 4–6 report the number and percentages of the papers from column 3 by category of relationship between IVs. Case A includes studies that use multiple economically distinct instruments. Case B includes studies that use covariate interactions with a single instrument. Case C includes studies that use multiple functions of a single instrument.

IV had a specification with fewer endogenous variables than IVs. Nearly all of these specifications have just a single endogenous variable, consistent with the framework that we will use in our analysis.

A few of the papers in column 2 used more IVs than endogenous variables in a way that was either nonstandard or unclear to us. In column 3, we remove these papers and focus on only those that combined multiple IVs using 2SLS or optimally weighted generalized method of moments (GMM). This leaves 43 percent of papers across the five journals. This confirms our claim that combining more instruments than endogenous variables through 2SLS (or occasionally GMM) is widespread empirical practice. This practice is the focus of this paper.

Columns 4–6 of Table 1 classify the papers from column 3 into three types by the relationship between their multiple IVs. Case A is when at least two of the IVs are economically distinct quantities. For example, CHV instrument for college attendance with measures of local labor market conditions, distance to college, and local tuition in public four-year colleges. As shown in column 4, the vast majority of papers fall into case A. Given its empirical importance, our analysis will be centered around this case.

The other two cases are less common. Case B are papers in which all IVs in the multiple IV specification are formed via interactions between predetermined covariates and a single exogenous instrument. For example, Angrist and Krueger (1991) use a full set of indicators interacted between state and quarter of birth as instruments for years of schooling. As we see in column 5, case B accounts for only 13 percent of the papers, making it the least common of the three types of papers with multiple IVs. Case C is when the multiple IV specification uses functions of a single exogenous variable. For example, using distance and its square as instruments for college attendance would fall into this case. As shown in column 6, case C is also relatively rare, accounting for only 19 percent of the papers.

II. Monotonicity Conditions with Multiple IVs

This section contains our results on the interpretation of the IA and partial monotonicity conditions when there are multiple IVs. We first develop an equivalent

characterization of the IA condition, which facilitates a graphical interpretation of its content. Among other things, the graphical interpretation reveals that the IA condition has nothing to do with "monotonicity" in the usual sense of the phrase. Then, we argue that IA monotonicity severely restricts choice heterogeneity. This motivates interest in the weaker, partial monotonicity assumption. Throughout the paper, we consider the binary treatment case, since allowing for multiple treatments already complicates the IA monotonicity condition even with a single instrument (Angrist and Imbens 1995; Heckman, Urzua, and Vytlacil 2006). For our theoretical analysis, we condition on covariates nonparametrically and keep this conditioning implicit in the notation.

A. *Definition of the IA Monotonicity Condition*

Consider a population of individuals $i \in \mathcal{I}$. Denote individual i's potential treatment status if some instrument vector Z_i were set to z by $D_i(z) \in \{0, 1\}$, where z takes values in some subset $\mathcal{Z}$ of $\mathbb{R}^L$. We assume that the support of Z_i is contained in $\mathcal{Z}$, possibly as a proper subset. When $L > 1$, we view each component of the vector Z_i as comprising an economically distinct quantity. That is, if $L = 2$ then $Z_{i,1}$ and $Z_{i,2}$ will denote the two distinct instruments, each of which can take two or more values.

IA introduced the following assumption on the potential treatment states, which they described as "monotonicity."

ASSUMPTION IAM (IA Monotonicity): *For all $z, z' \in \mathcal{Z}$ either $D_i(z) \geq D_i(z')$ for all $i \in \mathcal{I}$, or $D_i(z) \leq D_i(z')$ for all $i \in \mathcal{I}$.*

Heckman and Vytlacil (2005, pp. 715–16) observed that Assumption IAM requires uniformity across individuals, not monotonicity in the instrument. The results ahead provide further justification of their observation. Nevertheless, to conform with the existing literature, we still refer to Assumption IAM as "IA *monotonicity*." For clarity, we refer to the usual definition of monotonicity as "*actual* monotonicity."

ASSUMPTION AM (Actual Monotonicity): *If $z' \geq z$ in the vector sense (component-wise), then $D_i(z') \geq D_i(z)$ for all $i \in \mathcal{I}$.*

We show ahead that IA monotonicity (Assumption IAM) neither implies nor is implied by actual monotonicity (Assumption AM).

B. *A Graphical Characterization of IA Monotonicity*

Assumption IAM is a comparison across all individuals for any two values of Z_i. To interpret this condition when Z_i is a vector, it is useful to rephrase it as a comparison across all values of Z_i for any two individuals. The equivalent condition is that for any two individuals j and k, either j must take treatment under all instrument values that k does, or the opposite. This is the content of the following proposition.[6]

---

[6] Proofs for all propositions are contained in the Appendix.

PROPOSITION 1: *For any $i \in \mathcal{I}$, define $\mathcal{Z}_i \equiv \{z \in \mathcal{Z} : D_i(z) = 1\}$ as the set of all instrument values for which individual i would take treatment. Assumption IAM holds if and only if for all $j, k \in \mathcal{I}$, either $\mathcal{Z}_j \subseteq \mathcal{Z}_k$, or $\mathcal{Z}_k \subseteq \mathcal{Z}_j$.*

Proposition 1 shows that Assumption IAM can be interpreted as a "nesting condition" among the sets of instrument values that induce different agents to take treatment.[7] This means that with two instruments one can gain intuition about the content of Assumption IAM by drawing sets in $\mathbb{R}^2$.

For example, in Figure 1 panel A, we have drawn two sets $\mathcal{Z}_j$ and $\mathcal{Z}_k$ that are not nested. Proposition 1 says that Assumption IAM fails, which can be verified by comparing the choices individuals j and k would make at the points marked z_j and z_k. Yet, for both individuals j and k, the instrument has a monotonic effect in the sense that $D_i(z)$ is increasing in z. That is, if $z' \geq z$ as a vector (component-wise), then $D_i(z') \geq D_i(z)$. This shows that Assumption AM does not imply Assumption IAM.

Figure 1 panel B depicts the opposite case, in which $\mathcal{Z}_k \subseteq \mathcal{Z}_j$. If $\mathcal{I}$ only consists of individuals like j and k, then Proposition 1 implies that Assumption IAM is satisfied. However, the instrument does not have a monotonic effect on treatment choice. For example, moving from z to $z' \geq z$ moves both individuals' choices from zero to one, but moving from z' to $z'' \geq z'$ moves their choices back to zero. This shows that Assumption IAM does not imply monotonicity in the usual sense of Assumption AM.

C. *Implications for Heterogeneity in Choice Behavior*

In this section, we examine the restrictions that Assumption IAM places on choice behavior. To do this, we use a random utility model. Assume that individual i's indirect utility from choosing d when the instrument is z is given by $V_i(d, z)$. The individual chooses $D_i(z) = 1$ if and only if $V_i(1, z) \geq V_i(0, z)$:

$$(1) \qquad D_i(z) = \underset{d \in \{0,1\}}{\mathrm{argmax}} \, V_i(d, z) \equiv \begin{cases} 1, & \text{if } V_i(z) \geq 0; \\ 0, & \text{if } V_i(z) < 0, \end{cases}$$

where $V_i(z) \equiv V_i(1, z) - V_i(0, z)$ and ties are resolved in favor of treatment.

For concreteness, consider the familiar setting of the returns to schooling in which $D_i(z)$ represents a binary decision to attend college. Suppose that $z = (z_1, z_2)$, where z_1 is a tuition subsidy, and z_2 is proximity to a college (e.g., Kane and Rouse 1993, Card 1995). Larger values of either instrument encourage college attendance, so that $D_i(z)$ is a monotonic function of z and Assumption AM is satisfied. As observed in the previous section, this neither implies nor is implied by Assumption IAM. In Figure 2, we draw two possible indifference curves along which individuals j and k would be on the margin between attending and not attending college.

Suppose that z only takes the values $\{(0,0), (0,1), (1,0), (1,1)\}$ shown in Figure 2 panel A. Then individual j would attend college if and only if they received a tuition

---

[7] This nesting condition is different than the "equivalent monotonicity condition" used by Vytlacil (2002, p. 335), although it shares a superficial resemblance. Vytlacil (2002, p. 336) used the sets $\mathcal{Z}_i$ for his proof of the existence of an equivalent threshold-crossing model.
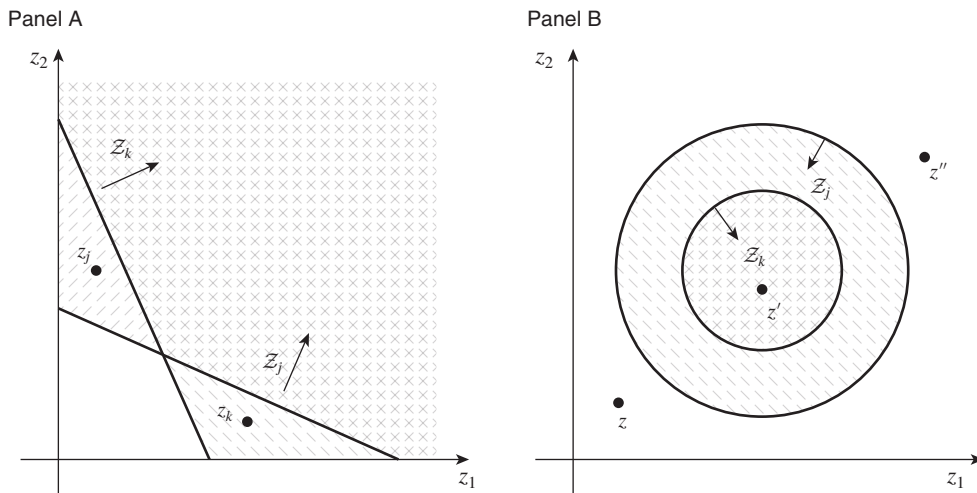
Panel A

Panel B

FIGURE 1. ASSUMPTION IAM NEITHER IMPLIES NOR IS IMPLIED BY MONOTONICITY OF $D_i(z)$ IN z

Notes: Panel A: sets $\mathcal{Z}_j$ and $\mathcal{Z}_k$ are not nested, so Proposition 1 implies that Assumption IAM does not hold. For example, compare z_j and z_k: $D_j(z_j) = 1 > 0 = D_k(z_j)$, while $D_k(z_k) = 1 > 0 = D_j(z_k)$. Yet, $D_i(z)$ is monotone in z for both $i = j,k$. Panel B: if $\mathcal{I} = \{j,k\}$, then Proposition 1 shows that Assumption IAM would hold. However, neither $D_j(z)$ nor $D_k(z)$ are monotone in z. For example, $z \leq z'$, and $z' \leq z''$, but $D_i(z) = D_i(z'') = 0 < D_i(z') = 1$ for $i = j,k$.

subsidy, regardless of whether they lived in close proximity. Individual k would attend college if and only if they lived in close proximity, regardless of whether they received a tuition subsidy. That is, $(1,0)$ is in $\mathcal{Z}_j$ but not $\mathcal{Z}_k$, and $(0,1)$ is in $\mathcal{Z}_k$ but not $\mathcal{Z}_j$, so that Assumption IAM fails. Thus, Assumption IAM does not permit individuals to differ in their responses to different incentives to attend college: all individuals must find either a tuition subsidy or distance to be more compelling. This suggests a strong form of preference homogeneity.

We can sharpen this statement when the instrument is continuous and the net indirect utility function is differentiable. This is shown in the next proposition, which is illustrated in Figure 2 panel B.

PROPOSITION 2: *Suppose that $D_i(z)$ is determined by* (*1*). *Let z^* be a point in the interior of $\mathcal{Z}$, and let $\mathcal{I}(z^*) \equiv \{i \in \mathcal{I} : V_i(z^*) = 0\}$ denote the set of individuals who are indifferent between treatment and nontreatment when faced with z^*. Suppose further that $V_i(z)$ is a continuously differentiable function of z in a neighborhood of z^* for all $i \in \mathcal{I}$. Then Assumption IAM implies that*

$$(2) \qquad \partial_1 V_j(z^*)\partial_2 V_k(z^*) = \partial_1 V_k(z^*)\partial_2 V_j(z^*)$$

for all $j,k \in \mathcal{I}(z^*)$, where $\partial_\ell V_i(z) \equiv (\partial/\partial z_\ell)V_i(z)$ for $\ell = 1,2$.

Proposition 2 says that if Assumption IAM holds, then any two individuals who are indifferent between treatment and nontreatment at z^* must have the same marginal rate of substitution between the two components of the instrument. That is,
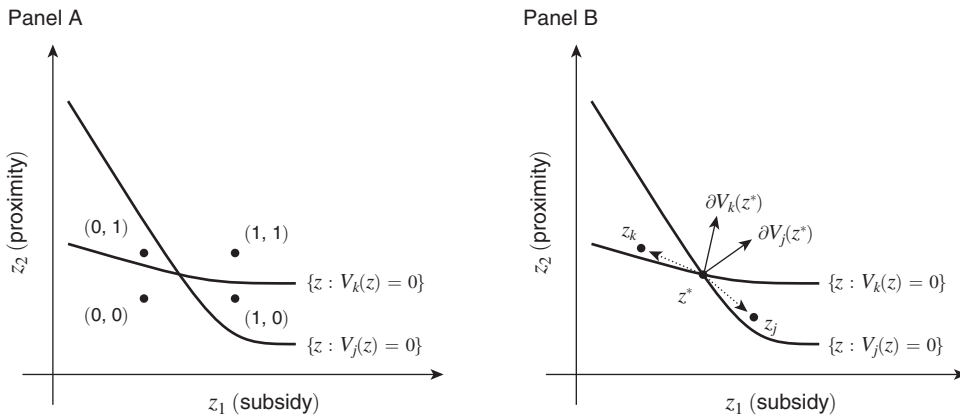
Panel A

Panel B

FIGURE 2. ASSUMPTION IAM REQUIRES HOMOGENOUS CHOICE BEHAVIOR

Notes: Panel A: each instrument takes two binary values. Individual j has $\mathcal{Z}_j = \{(1,0),(1,1)\}$ and individual k has $\mathcal{Z}_k = \{(0,1),(1,1)\}$. Points $(1,0)$ and $(0,1)$ violate the nesting condition in Proposition 1. Panel B: an illustration of Proposition 2. When (3) fails to hold for two individuals $j,k \in \mathcal{I}$, one can find points z_j and z_k which violate the nesting condition in Proposition 1 by taking small steps in the directions of the dotted arrows.

assuming that the second component has an impact at z^* (so that $\partial_2 V_i(z^*) \neq 0$), Assumption IAM implies

$$(3) \qquad \frac{\partial_1 V_j(z^*)}{\partial_2 V_j(z^*)} = \frac{\partial_1 V_k(z^*)}{\partial_2 V_k(z^*)}$$

for *all* individuals j and k who are indifferent at z^*. This is a strong statement about preference homogeneity.

For example, suppose that individual i's net utility from attending college is given by the random coefficient specification

$$(4) \quad V_i(z) = B_{i,0} + B_{i,1}z_1 + z_2 \quad \text{so that} \quad D_i(z) = \mathbf{1}[B_{i,0} + B_{i,1}z_1 + z_2 \geq 0],$$

where $B_{i,1} \geq 0$ controls variation in the taste for subsidies relative to proximity. Proposition 2 shows that Assumption IAM does not hold if $B_{i,1}$ varies with i.[8] Thus, the college attendance decision of every individual is either affected more by tuition subsidies (if $b_1 \equiv B_{i,1} \geq 1$), or by proximity (if $b_1 < 1$), and all individuals trade off these incentives at the same rate. Assumption IAM does not permit heterogeneity in these behavioral responses.

D. *Partial Monotonicity*

Assumption IAM creates unattractive implications for choice behavior because it requires cross-instrument comparisons, such as the comparison between $(0,1)$

---

[8] Heckman, Urzua, and Vytlacil (2006, p. 399) note that Assumption IAM *can* fail in random coefficient specifications like (4). Our analysis shows that it *must* fail when the instruments are continuous.

and $(1, 0)$ in Figure 2 panel A. We can eliminate these comparisons by considering a condition that compares only values of a single component of the instrument, holding all other components fixed. To state such a condition, we divide vectors $z \in \mathcal{Z}$ into their ℓth component, z_ℓ, and all other $(L - 1)$ components, $z_{-\ell}$. We write $z = (z_\ell, z_{-\ell})$ to emphasize the separation of the ℓth component.

Consider the following assumption.[9]

ASSUMPTION PM (Partial Monotonicity): *Take any* $\ell = 1, \ldots, L$, *and let* $(z_\ell, z_{-\ell})$ *and* $(z'_\ell, z_{-\ell})$ *be two points in* $\mathcal{Z}$. *Then either* $D_i(z_\ell, z_{-\ell}) \geq D_i(z'_\ell, z_{-\ell})$ *for all* $i \in \mathcal{I}$, *or* $D_i(z_\ell, z_{-\ell}) \leq D_i(z'_\ell, z_{-\ell})$ *for all* $i \in \mathcal{I}$.

Assumption IAM clearly implies Assumption PM. When $L = 1$, Assumption PM is equivalent to Assumption IAM; when $L > 1$, Assumption PM is strictly weaker. To see this, recall Figure 2 panel A, where we determined that Assumption IAM fails. Holding $z_2 = 0$ fixed, both individuals j and k are weakly induced to treatment by switching z_1 from zero to one. The same is true when the roles of z_1 and z_2 are swapped. If $\mathcal{I}$ consisted of only individuals like j and k, then Assumption PM would be satisfied.

Figure 2 suggests that a simple sufficient condition for Assumption PM is monotonicity in the usual sense of Assumption AM. This is the content of the following proposition.

PROPOSITION 3: *Assumption AM implies Assumption PM.*

Unlike Assumption IAM, Assumption AM can be easily satisfied in random utility models with heterogeneous preferences. For example, if $V_i(z)$ follows the random coefficient specification (4), then it will be satisfied if $B_{i,1}$ is positive for all i. This is easy to interpret and justify: all individuals are more likely to attend college if tuition is lower, even if they differ in their preferences for tuition relative to proximity. More generally, Proposition 3 shows that Assumption PM is satisfied whenever $V_i(z)$ is weakly increasing in z.[10]

Our running example of the returns to college was chosen because it is familiar, not because it favors either Assumption IAM or PM. In online Appendix B, we review three empirical papers from our survey in Section I. These papers are on diverse topics, but each one combines multiple economically distinct IVs using 2SLS, and so falls into column 4 (case A) of Table 1. For each paper, we discuss the substantive assumptions represented by both Assumptions IAM and PM. Our conclusion from that exercise is consistent with our theoretical results: Assumption IAM is an extremely strong condition with multiple instruments, whereas Assumption PM—while not innocuous—is much less controversial.[11]

---

[9] Mountjoy (2019) uses a similar assumption in a setting with multiple unordered treatments.

[10] More generally still, Theorem 4 of Milgrom and Shannon (1994) implies that Assumption PM will be satisfied if $V_i(d, z)$ has the single-crossing property in $(d; z)$.

[11] Coupled with the equivalence of Assumptions PM and IAM when $L = 1$, Proposition 3 may be the source of the common intuition that Assumption IAM follows from Assumption AM. In the examples of online Appendix B, the theoretical rationale underlying each instrument suggests that Assumption AM holds, which implies Assumption PM but not Assumption IAM.

Assumption AM is sufficient but not necessary for Assumption PM. For example, consider the study by Angrist and Evans (1998), who used the sex composition of a family's first two children to instrument for family size. Let D_i indicate whether the family had exactly two, or more than two children. Let $Z_{i,1}$ and $Z_{i,2}$ indicate the sexes of their first two children. Angrist and Evans (1998) present empirical evidence that families for which these two sexes are the same ($Z_i = (0,0)$ or $Z_i = (1,1)$) are more likely to have a third child. Suppose that this is due to a uniform preference ordering in the population, so that for all $i \in \mathcal{I}$

$$(5) \qquad D_i(0,0) \geq D_i(0,1), \qquad D_i(0,0) \geq D_i(1,0),$$

and

$$D_i(1,1) \geq D_i(0,1), \qquad D_i(1,1) \geq D_i(1,0).$$

This is consistent with Assumption PM, but it violates Assumption AM.

Another way to see the difference between Assumptions AM and PM is to consider the interacted random coefficient specification

$$(6) \qquad V_i(z) = B_{i,0} + B_{i,1} z_1 + z_2 + B_{i,2} z_1 z_2,$$

where we suppose for simplicity that $\mathcal{Z} = \{0,1\}^2$. If all individuals $i \in \mathcal{I}$ have $B_{i,1} \geq 0$, $B_{i,2} \leq -1$, but $B_{i,1} \leq -B_{i,2}$, then Assumption AM fails while Assumption PM is satisfied. The reason is due to the strong negative interaction effect (submodularity) between z_1 and z_2 on indirect net utility, which is controlled here by $B_{i,2}$. This implies that $D_i(z_1, z_2)$ is increasing as a function of z_1 when $z_2 = 0$, but decreasing when $z_2 = 1$, and similarly when the roles of z_1 and z_2 are reversed. This violates Assumption AM, even though Assumption PM is satisfied.[12]

III. Interpreting 2SLS under Partial Monotonicity

IA (Theorem 2) showed that under standard instrument exogeneity and relevance conditions, Assumption IAM ensures that the 2SLS estimand can be written as a weighted average of causal effects for complier subpopulations. The weights are convex in that they are nonnegative and sum to one. Their result holds regardless of the number of instruments, as long as the first stage for the 2SLS estimand is fully saturated. However, it crucially depends on Assumption IAM, which we have shown requires a strong form of preference homogeneity with two or more distinct instruments. In this section, we show that the IA result can be partially salvaged under Assumption PM.

---

[12] To see that the configuration above satisfies Assumption PM, note that $B_{i,1} \geq 0$ implies that $D_i(0,0) \leq D_i(1,0)$, $B_{i,1} \leq -B_{i,2}$ implies that $D_i(0,1) \geq D_i(1,1)$, $D_i(0,0) \leq D_i(0,1)$ by virtue of the normalized coefficient on z_2, and $B_{i,2} \leq -1$ implies that $D_i(1,0) \geq D_i(1,1)$.

A. *Potential Outcomes and Exogeneity Condition*

To define the 2SLS estimand, we need to introduce an outcome variable, Y_i. We write potential outcomes for Y_i as $Y_i(1)$ and $Y_i(0)$ to correspond to setting D_i to treatment ($D_i = 1$) and nontreatment states ($D_i = 0$). The notation incorporates the usual exclusion restriction that Z_i has no direct causal effect on Y_i. The observed outcome is $Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0) = Y_i(D_i)$. The observed treatment state is related to the potential treatment states analyzed in the previous section as

$$D_i = \sum_{z \in \mathcal{Z}} \mathbf{1}[Z_i = z] D_i(z) = D_i(Z_i).$$

Throughout the paper, we maintain the following exogeneity condition.

ASSUMPTION E (Exogeneity): $\left(Y_i(0), Y_i(1), \{D_i(z)\}_{z \in \mathcal{Z}} \right) \perp\!\!\!\perp Z_i.$

Assumption E is common in nonparametric IV models, and identical to Condition 1 of IA. For simplicity, we state the condition in terms of full independence, although our analysis of the 2SLS estimand only involves mean outcomes, so only requires a weaker form of Assumption E.[13]

B. *Two Binary Instruments*

We first consider the case in which $Z_i = (Z_{i,1}, Z_{i,2})$ consists of two binary instruments $Z_{i,1} \in \{0,1\}$ and $Z_{i,2} \in \{0,1\}$, with support $\mathcal{Z} = \{0,1\}^2$. We also assume that the direction of Assumption PM is known a priori to be

(7)
$$D_i(0,0) \leq D_i(0,1) \leq D_i(1,1),$$

and

$$D_i(0,0) \leq D_i(1,0) \leq D_i(1,1).$$

This particular ordering is not required, but it helps with the notation and exposition to not have to consider multiple different orderings simultaneously; our more general results in Section IIID relax this condition.[14] The following proposition establishes that under these conditions the population $\mathcal{I}$ can be partitioned into six mutually exclusive and exhaustive groups.

PROPOSITION 4: *If $\mathcal{Z} = \{0,1\}^2$ and Assumption PM is satisfied with the order shown in (7), then each $i \in \mathcal{I}$ belongs to exactly one of the six groups in Table 2.*

---

[13] The weaker form still requires that $\{D_i(z)\}_{z \in \mathcal{Z}} \perp Z_i$, but relaxes full joint independence to the restriction that $E\big[Y_i(d) | \{D_i(z)\}_{z \in \mathcal{Z}}, Z_i = z'\big]$ not depend on z' for both $d = 0$ and $d = 1$.

[14] Also, note that the direction of the comparisons are still identified under Assumption PM, just as they are under Assumption IAM, since Assumption E implies that $\Pr\big[D_i(z) = 1\big] = \Pr[D_i = 1 | Z_i = z]$.

<div align="center">Table 2—Groups under Assumption PM with (7)</div>

G_i (group g)	z^1 $D_i(0,0)$	z^2 $D_i(0,1)$	z^3 $D_i(1,0)$	z^4 $D_i(1,1)$	$\mathcal{C}_g$	$\mathcal{D}_g$
Always-taker (*at*)	1	1	1	1	$\varnothing$	$\varnothing$
Eager complier (*ec*)	0	1	1	1	$\{2\}$	$\varnothing$
Reluctant complier (*rc*)	0	0	0	1	$\{4\}$	$\varnothing$
Never-taker (*nt*)	0	0	0	0	$\varnothing$	$\varnothing$
Z_1 complier (1*c*)	0	0	1	1	$\{3\}$	$\varnothing$
Z_2 complier (2*c*)	0	1	0	1	$\{2,4\}$	$\{3\}$

Notes: The sets $\mathcal{C}_g$ and $\mathcal{D}_g$ are integers k for which group type g acts as a complier or defier when considering the contrast z^{k-1} to z^k. This notation is explained in more detail in Section IIID.

The terminology in Table 2 modifies that of Angrist, Imbens, and Rubin (1996). Always- and never-takers exhibit the same behavior regardless of either instrument. The Z_1 compliers take treatment if and only if $Z_{i,1}$ is switched on, while the Z_2 compliers take treatment if and only if $Z_{i,2}$ is switched on. Eager compliers participate in treatment if either instrument is on, and reluctant compliers only participate if both instruments are on. For any of the six groups, an increase in either instrument weakly increases treatment, which implies that Assumption AM also holds in this case. Assumption IAM will be violated if the population includes both Z_1 and Z_2 compliers, since in this case a change in Z_i from $(0,1)$ to $(1,0)$ would induce Z_1 compliers to enter treatment and Z_2 compliers to exit treatment. Figure 3 shows how realizations of random coefficients would map into these six group types if potential treatment states were generated by equation (4).

As in IA (Theorem 2), we consider the 2SLS estimand with a saturated first stage, and a second stage with D_i and a constant. That is, the 2SLS estimand is formed by using one (a constant), $Z_{i,1}$, $Z_{i,2}$, and $Z_{i,1}Z_{i,2}$ as instruments for one and D_i. Since the first stage is saturated, this 2SLS procedure generates the same coefficient estimate on D_i as the IV estimator that uses the propensity score $p(Z_i) \equiv \Pr[D_i = 1|Z_i]$ as the sole instrument for D_i. Let $\beta_{2sls} = \mathrm{cov}(Y_i, p(Z_i))/\mathrm{cov}(D_i, p(Z_i))$ denote this coefficient. Let $G_i \in \{at, nt, 1c, 2c, ec, rc\}$ denote individual i's group among the six shown in Table 2, and define $\pi_g \equiv \Pr[G_i = g]$ and $\Delta_g \equiv E[Y_i(1) - Y_i(0)|G_i = g]$ to be the population shares and average treatment effects for each group. The following proposition establishes the relationship between these quantities and β_{2sls}.

PROPOSITION 5: *Suppose Z_i has support $\mathcal{Z} \equiv \{0,1\}^2$ and that Assumption PM is satisfied with the ordering in (7). Suppose in addition that Assumption E is satisfied, and that β_{2sls} exists. Then*

$$\beta_{2sls} = \sum_{g \in \{1c, 2c, ec, rc\}} \omega_g \Delta_g,$$

where the ω_g are weights that sum to one. Both ω_{ec} and ω_{rc} are always nonnegative. If $\pi_{1c} \geq \pi_{2c}$, then ω_{1c} is also nonnegative, while the sign of ω_{2c} is given by

$$\mathrm{sgn}(\omega_{2c}) = \mathbf{1}[\pi_{2c} > 0] \times \mathrm{sgn}(\Pr[D_i = 1|Z_{i,2} = 1] - \Pr[D_i = 1|Z_{i,2} = 0]).$$
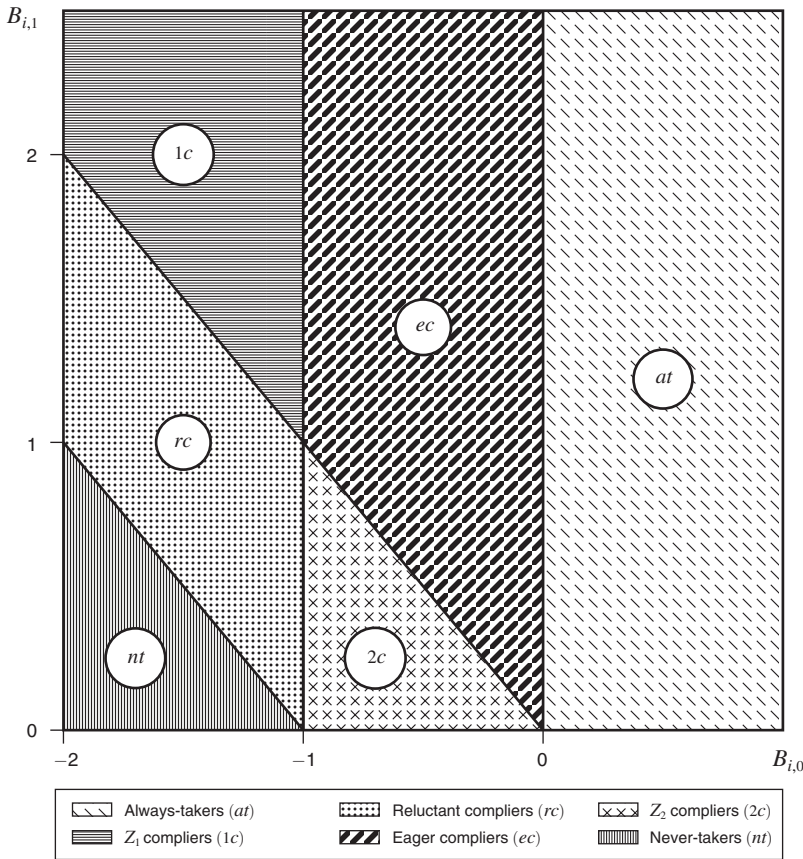
FIGURE 3. CORRESPONDENCE BETWEEN RANDOM COEFFICIENTS AND BEHAVIORAL TYPES

Notes: These regions show how random coefficient realizations relate to behavioral types if treatment were determined by the binary choice model (4) with $B_{i,1} \geq 0$. For example, a Z_2 complier has $D_i(0,0) = 0$, $D_i(1,0) = 0$, $D_i(0,1) = 1$, and $D_i(1,1) = 1$. The first three choices imply that $B_{i,0} < 0$, $B_{i,0} + B_{i,1} < 0$, and $B_{i,0} + 1 > 0$, with the fourth choice implied by the third. The region of such $(B_{i,0}, B_{i,1})$ realizations is shown in crosshatches.

If $\pi_{2c} \geq \pi_{1c}$, then ω_{2c} is nonnegative, and the sign of ω_{1c} is given by

$$\text{sgn}(\omega_{1c}) = \mathbf{1}[\pi_{1c} > 0] \times \text{sgn}\big(\text{Pr}[D_i = 1 | Z_{i,1} = 1] - \text{Pr}[D_i = 1 | Z_{i,1} = 0]\big).$$

Proposition 5 shows that the 2SLS estimand is a linear combination of average treatment effects for the four groups that change treatment status in response to one or both of the instruments. The groups are nonoverlapping, and the weights on the groups (ω_g) sum to unity.[15] This implies that $\beta_{2sls} = \Delta \equiv \Delta_g$ in the unlikely situation that there are constant treatment effects across behavioral groups. In general, however, the 2SLS estimand might not be a positively weighted average of treatment effects for the four groups, because some of the weights might be negative. The weights for reluctant compliers and eager compliers are always nonnegative. If $\pi_{1c} \geq \pi_{2c}$, the weight given to the Z_1 compliers is also nonnegative, but the weight given to the Z_2 compliers can be either positive or negative.

---

[15] Formulas for ω_g are given in the proof.

The intuition is that a shift of Z_i from $(0,1)$ to $(1,0)$ induces Z_1 compliers to enter treatment and Z_2 compliers to exit treatment. If $\pi_{1c} \geq \pi_{2c}$, then the net effect of this shift is still more participation in treatment. However, the Z_2 compliers act as "defiers," and therefore receive negative weight for this binary contrast. Whether the overall weight given to the Z_2 compliers is positive or negative depends on whether this negative weight is outweighed by the positive weight given to the Z_2 compliers in the two other instrument contrasts for which they enter treatment: $Z_i = (0,0)$ to $Z_i = (0,1)$, and $Z_i = (1,0)$ to $Z_i = (1,1)$. If instead $\pi_{2c} \geq \pi_{1c}$, then the roles of the Z_1 and Z_2 compliers are reversed.

C. *When Is the 2SLS Estimand a Positively Weighted Average of Complier Causal Effects?*

Proposition 5 shows that one can check whether the 2SLS estimand is a positively weighted average of causal effects by examining observable relationships between treatment and instruments. Specifically, if both Z_1 and Z_2 compliers are present in the population and $\pi_{1c} \geq \pi_{2c}$, then ω_{2c} will be negative if and only if the coefficient on $Z_{i,2}$ in a regression of D_i on $Z_{i,2}$ (and a constant) is negative. Likewise, if $\pi_{2c} \geq \pi_{1c}$, then ω_{1c} will be negative if and only if the coefficient on $Z_{i,1}$ in a regression of D_i on $Z_{i,1}$ (and a constant) is negative. One can either check both cases, or check only the relevant case, which is identified by the sign of $p(1,0) - p(0,1) = \pi_{1c} - \pi_{2c}$.

Except for π_{at} and π_{nt}, the group shares are not themselves separately identified. This is because there are five linearly independent group shares (after accounting for the fact that they sum to unity), but only four values of the propensity score. As mentioned, the sign of $p(1,0) - p(0,1)$ indicates whether Z_1 or Z_2 compliers are more common. However, it is not possible to determine whether the less common group is entirely absent.

If it were the case that $\pi_{1c} = 0$ or $\pi_{2c} = 0$, then Assumption IAM would hold and the 2SLS estimand would necessarily be a positively weighted average of the remaining three complier groups. The correspondence between complier group shares and Assumption IAM can be seen graphically in Figure 3. Assumption IAM implies the coefficient B_{i1} has a degenerate distribution, collapsing the regions in Figure 3 onto a horizontal line. This line can go through either the region defining Z_1 compliers or the region defining Z_2 compliers but not both, so Assumption IAM implies at least one of the two groups must be absent.

Finding negative weights based on the formulas in Proposition 5 represents a situation in which the unconditional relationship between an instrument and the treatment has a different sign than the ceteris paribus impact of the instrument. This may be rare in practice, since researchers often have prior beliefs regarding the impacts of the instruments (e.g., if each instrument is an encouragement to take treatment), and a researcher may be unlikely to use an instrument if the raw correlation with the treatment contradicts the expected sign. A necessary condition for such a contradiction is that the instruments are negatively correlated.

PROPOSITION 6: *Suppose Z_i consists of two binary instruments that satisfy Assumption PM with the ordering in* (7), *that Assumption E is satisfied, and that β_{2sls} exists. If* $\text{cov}(Z_{i,1}, Z_{i,2}) \geq 0$, *then both ω_{1c} and ω_{2c} are nonnegative.*

The intuition for Proposition 6 follows from the standard omitted variables bias formula for ordinary least squares (OLS) regression. The second instrument serves as an omitted confounder in a bivariate regression of the treatment on a single instrument. Since each instrument increases the likelihood of treatment, a negative bivariate relationship can only arise when the included instrument is negatively correlated with the omitted instrument, creating negative omitted variable bias. An important special case of Proposition 6 is when the instruments are independent, so that $\text{cov}(Z_{i,1}, Z_{i,2}) = 0$ and the 2SLS weights are guaranteed to be positive. The leading scenario in which the instruments would be negatively correlated is when $Z_{i,j} = 1$ tends to imply $Z_{i,k} = 0$ for $j \neq k$. For example, $Z_{i,1}$ and $Z_{i,2}$ may indicate two different arms in an experiment corresponding to different types of encouragement to take the treatment. In this setting, $\text{cov}(Z_{i,1}, Z_{i,2}) < 0$ and negative weights are possible.

In online Appendix C, we use the full sufficient and necessary characterization in Proposition 5 to develop formal statistical tests about the signs of the weights. We consider four tests of the null hypothesis that the weights are positive. In a simulation study, we find that the test based on the procedure of Romano, Shaikh, and Wolf (2014) performs the best, with size only slightly below its nominal level. We also consider a test of the null hypothesis that at least one weight is negative. All five tests can be implemented using the mivcausal Stata package.

D. *Multivalued Instruments*

Suppose that Z_i consists of two or more distinct, discrete instruments, and that its support has K elements total. Label these elements as $\text{supp}(Z_i) \equiv \{z^1, \ldots, z^K\}$ in increasing order of the propensity score, so that $p(z^k) \geq p(z^{k-1})$ for all $k \geq 2$. In the case considered in Section IIIB, $K = 4$ and the instrument values would be ordered as $z^1 = (0,0)$, $z^2 = (0,1)$, $z^3 = (1,0)$, and $z^4 = (1,1)$ if $p(1,0) \geq p(0,1)$, with the roles of z^2 and z^3 switched in the opposite case.

Let $\mathcal{G}$ represent the set of all realizations of $\{D_i(z)\}_{z \in \mathcal{Z}}$ that are consistent with Assumption PM and the observed values of $p(z)$. In Section IIIB, where there were two binary instruments with the ordering (7), $\mathcal{G}$ was composed of the six groups in Table 2. Table 3 displays another example consistent with the ordering (5) in which there are seven groups. As before, let G_i denote individual i's group membership, let π_g denote the proportion of the population in each group g, and let Δ_g denote group g's average treatment effect.

For each $g \in \mathcal{G}$, define the set:

$$\mathcal{C}_g = \left\{ k \in \{2, \ldots, K\} : D_i(z^k) = 1 \text{ and } D_i(z^{k-1}) = 0 \text{ for all } i \text{ with } G_i = g \right\}.$$

This is the set of instrument values k at which individuals in group g are compliers in the sense that they would not take treatment if $Z_i = z^{k-1}$, but would take treatment if $Z_i = z^k$. Similarly, define

$$\mathcal{D}_g = \left\{ k \in \{2, \ldots, K\} : D_i(z^k) = 0 \text{ and } D_i(z^{k-1}) = 1 \text{ for all } i \text{ with } G_i = g \right\}$$

TABLE 3—GROUPS UNDER ASSUMPTION PM WITH (5)

G_i (group g)	z^1 $D_i(1,0)$	z^2 $D_i(0,1)$	z^3 $D_i(1,1)$	z^4 $D_i(0,0)$	$\mathcal{C}_g$	$\mathcal{D}_g$
Always-taker (*at*)	1	1	1	1	$\emptyset$	$\emptyset$
Group 2	0	1	1	1	$\{2\}$	$\emptyset$
Group 3	1	0	1	1	$\{3\}$	$\{2\}$
Group 4	0	0	1	1	$\{3\}$	$\emptyset$
Group 5	0	0	0	1	$\{4\}$	$\emptyset$
Group 6	0	0	1	0	$\{3\}$	$\{4\}$
Never-taker (*nt*)	0	0	0	0	$\emptyset$	$\emptyset$

as the set of instrument values at which individuals in group g act as defiers. For example, in Section IIIB, we had $\mathcal{C}_{1c} = \{3\}$, $\mathcal{D}_{1c} = \emptyset$, $\mathcal{C}_{2c} = \{2,4\}$, and $\mathcal{D}_{2c} = \{3\}$. We also had $\mathcal{C}_{ec} = \{1\}$, $\mathcal{C}_{rc} = \{4\}$, $\mathcal{C}_{at} = \mathcal{C}_{nt} = \emptyset$, and $\mathcal{D}_g = \emptyset$ for each $g \in \{at, nt, ec, rc\}$; recall Table 2. Table 3 reports the sets $\mathcal{C}_g$ and $\mathcal{D}_g$, for an alternative case consistent with (5).

As before, consider the same 2SLS specification used by IA (Theorem 2) with a saturated first stage, and a second stage that contains D_i and a constant. Let β_{2sls} denote the 2SLS estimand corresponding to the coefficient on D_i. The following proposition provides an interpretation of the 2SLS estimand.

PROPOSITION 7: *Suppose Z_i takes K values $\{z^1, \ldots, z^K\}$ labeled so that the propensity score is increasing and suppose that the support of Z_i is rectangular; that is* $\text{supp}(Z_i) = \text{supp}(Z_{i,1}) \times \cdots \times \text{supp}(Z_{i,L})$. *If Assumptions PM and E are satisfied, and if β_{2sls} exists, then*

$$\beta_{2sls} = \sum_{g \in \mathcal{G}: \mathcal{C}_g \neq \emptyset} \omega_g \Delta_g,$$

where ω_g are weights, such that $\sum_{g \in \mathcal{G}: \mathcal{C}_g \neq \emptyset} \omega_g = 1$, and

$$\text{sgn}(\omega_g) = \mathbf{1}[\pi_g > 0]$$
$$\times \text{sgn}\left(\sum_{k=2}^{K} \left(\mathbf{1}[k \in \mathcal{C}_g] - \mathbf{1}[k \in \mathcal{D}_g] \right) \text{cov}\left(D_i, \mathbf{1}[p(Z_i) \geq p(z^k)] \right) \right).$$

Proposition 7 shows that under Assumption PM, the 2SLS estimator produces a weighted average of treatment effects for groups that comply with some instrument change. The weights on each group could be positive or negative, but this can be checked empirically. To do this, one must generate the sets $\mathcal{C}_g$ and $\mathcal{D}_g$ by applying Assumption PM to the set of K instrument values and the observed ordering of the propensity score, as we did in the examples in Tables 2 and 3. The sign of the weight for group g is then determined by the overall sum of $\text{cov}\left(D_i, \mathbf{1}[p(Z_i) \geq p(z^k)] \right)$ for instrument values k at which they comply less the sum of these terms at values k for which they defy.

The additional rectangular support condition in Proposition 7 is necessary to ensure that the contrasts picked up in β_{2sls} are restricted by Assumption PM. For example, if in the special case in Section IIIB the support of Z_i were only

$\{(0,1),(1,0)\}$, then either the Z_1 compliers or Z_2 compliers would always have negative weight. This is intuitive since Assumption PM does not place any direct restrictions on behavior in the contrast between $(0,1)$ and $(1,0)$.

E. *Using Each Instrument Separately*

A natural alternative to combining two or more instruments through 2SLS is to use each instrument separately. This may be unattractive due to a loss of statistical precision. However, it is justified under Assumption PM with the important caveat that all of the other instruments generally must be conditioned on as control variables in both the first and the second stage of the 2SLS estimator.[16] This is necessary because Assumption E is stronger than the exogeneity condition in the traditional constant effects linear IV model. Whereas the traditional condition only requires Z_i to be exogenous with respect to the outcome process, Assumption E also requires Z_i to be independent of $\{D_i(z)\}_{z \in \mathcal{Z}}$, and thus exogenous with respect to potential treatments as well.

To see why conditioning is necessary, consider again the case with two binary instruments. Using the first instrument without conditioning on the second requires considering marginal potential treatment states with respect to $Z_{i,1}$. These are related to the joint potential treatment states by

$$(8) \qquad D_{i,1}(z_1) \equiv (1 - Z_{i,2}) D_i(z_1, 0) + Z_{i,2} D_i(z_1, 1).$$

To apply the IA LATE interpretation using only $Z_{i,1}$ as the instrument would require that $Z_{i,1}$ is independent of $(D_{i,1}(0), D_{i,1}(1))$. We can see from (8) that this is unlikely to hold outside of the case where $Z_{i,1}$ and $Z_{i,2}$ are themselves independent. However, Assumption E does imply that $Z_{i,1}$ and $(D_{i,1}(0), D_{i,1}(1))$ will be independent *conditional* on $Z_{i,2}$. Similarly, Assumption PM implies that Assumption IAM holds for the marginal potential treatments *conditional* on $Z_{i,2}$. See Mogstad, Torgovitsky, and Walters (2020) for a complete development of these ideas.

Under Assumption PM, using one instrument while conditioning on the others preserves the interpretation of the 2SLS estimand as a positively weighted average of treatment effects because the results of IA apply for each instrument holding the others fixed. The resulting parameters generally identify treatment effects for mixtures of complier groups. In the example of Section IIIB, using $Z_{i,1}$ as the instrument when $Z_{i,2} = 0$ identifies a weighted average of effects for Z_1 compliers and eager compliers, while using $Z_{i,1}$ as the instrument when $Z_{i,2} = 1$ identifies a weighted average of effects for Z_1 compliers and reluctant compliers. The labels are reversed when $Z_{i,2}$ is used as the instrument conditional on $Z_{i,1}$. Since the relative shares of each group in the population are not identified, it is not possible to use these parameters to solve for the average treatment effect for any specific complier group.

In Table 4, we report on the ways that researchers use multiple instruments separately using the same sample of papers from Section I. Of the 122 papers in Table 1,

---

[16]The consequences of not controlling for other instruments have been discussed elsewhere, e.g., CHV (footnote 6).

we found 20 with multiple IVs that fit a separate 2SLS model using a single IV for at least one specification reported in the main body of the paper. Column 2 of Table 4 restricts these papers to those that do not condition on the other omitted IVs when fitting such a model, while column 3 further restricts column 2 to papers that never used all IVs together in any specification. As shown in column 2, 90 percent of the papers that use multiple IVs separately do not control for the omitted instruments. Unless the multiple IVs are independent, these papers are reporting an estimated quantity that likely does not have a causal interpretation.

IV. Estimating the Returns to College

We illustrate the results of the previous sections by estimating the returns to college attendance.[17]

A. *Sample and Instruments*

CHV used multiple IVs to estimate the returns to college in a sample of White men from the National Longitudinal Survey of Youth (NLSY) 1979 Cohort (Bureau of Labor Statistics 2019). We use their sample (Carneiro, Heckman, and Vytlacil 2019) and focus on the two instruments that are the strongest predictors of college attendance: average log earnings in a youth's county of residence at age 17, and the presence of a four-year college in a youth's county of residence at age 14.[18] We control for permanent differences in earnings across counties so that the local earnings instrument can be interpreted as capturing idiosyncratic fluctuations in local labor market opportunities at the time of the college enrollment decision. For these two instruments, Assumption PM holds if all individuals view college as more attractive when local labor market opportunities are weaker and when a college is nearby. This seems plausible, since higher opportunity costs and longer distances are likely undesirable for everyone. In contrast, Assumption IAM requires all individuals to place the same relative weight on these economically distinct types of incentives.

Table 5 reports descriptive statistics for key variables in the CHV sample. The outcome variable Y_i is the log of individual i's average hourly wage from 1989 to 1993, and the treatment D_i equals one for individuals that attended college. Throughout our analysis we control for a vector X_i of covariates that includes AFQT (Armed Forces Qualification Test) scores, mother's years of education, number of siblings, urban residence at age 14, year of birth indicators, permanent local earnings (defined as average log earnings in the age 17 county of residence from 1973 to 2000), and average log earnings in the county of residence in 1991.[19] The latter two variables account for general differences in earnings across counties.

---

[17] As an additional example, Mogstad, Torgovitsky, and Walters (2021) provide simulated data, estimation programs, and results based on a data generating process calibrated to our empirical application.

[18] These two instruments were also used in several earlier studies of the returns to education, e.g., Card (1995), Cameron and Heckman (1998), Kling (2001), and Cameron and Taber (2004).

[19] We follow CHV and use a version of AFQT that corrects for differences in years of schooling at the time of the test. CHV also include labor market experience as an additional control variable in the second stage of their estimation procedure. Because experience may partially mediate the effects of college enrollment, we instead control for age (captured by the year of birth indicators) and interpret the effects of college attendance inclusive of any impacts on experience.

	Papers that ever use IVs separately (1)	Papers that do not control for omitted IVs (2)	Papers that always use IVs separately and never control for omitted IVs (3)
Number	20	18	9

Note: This table includes the subset of papers from Table 1 that use multiple IVs and fit a separate 2SLS model using a single IV in at least one specification in the main body of the paper.

The bottom rows of Table 5 display descriptive statistics for the two instruments. To parallel the discussion of Section IIIB we create a binary version of the local earnings instrument, labeled $Z_{i,1}$, which equals one if the residual from a regression of local earnings on X_i is in the bottom quartile of the sample.[20] By construction, this definition means that 25 percent of youths experienced low local earnings at age 17. Fifty-three percent of individuals lived near a four-year college at age 14, which we label with a binary indicator, $Z_{i,2}$.

B. *2SLS Estimates*

We begin by reporting ordinary least squares (OLS) and 2SLS estimates of the returns to college. Following CHV, we divide all college coefficients by four to express returns on an annual basis. Column 1 of Table 6 displays the coefficient from a bivariate OLS regression of log wages on college attendance, which gives an annual return of 9 percent. Column 2 shows that controlling for the covariates and instruments reduces the college premium to 4 percent per year.

An extensive literature argues that OLS estimates of the returns to schooling may be biased by unobserved differences in ability (e.g., Card 2001). Motivated by this concern, columns 3–7 of Table 6 report 2SLS estimates of the system:

$$(9) \qquad D_i = \psi + \Pi_1 Z_{i,1} + \Pi_2 Z_{i,2} + X_i'\delta + \eta_i,$$

$$(10) \qquad Y_i = \alpha + \beta D_i + X_i'\gamma + \epsilon_i.$$

We compare 2SLS estimates from models excluding the instruments from the second stage one at a time, adding the other to the set of controls X_i in equation (10) as in Section IIIE, as well as a model that excludes both instruments from the second stage simultaneously.

The first stage estimates show that both instruments boost college attendance. Column 3 of Table 6 reports OLS estimates of equation (9), and column 4 shows estimates from a saturated specification that adds an interaction between the two instruments. The interaction term is small and statistically insignificant, and its value implies that the partial effect of each instrument is always positive, indicating that

---

[20] In results not reported here, we found that other thresholds for discretizing the local earnings instrument yielded similar results but were less predictive of college attendance.

TABLE 5—SUMMARY STATISTICS

	Mean (1)	Standard deviation (2)
Average log wage, 1989–1993	2.378	0.499
Attended college	0.495	0.500
Corrected AFQT score	0.449	0.952
Mother's years of schooling	12.102	2.335
Number of siblings	2.927	1.909
Urban residence at age 14	0.744	0.436
Year of birth	1959.759	2.340
Permanent local log earnings	10.283	0.188
Local log earnings in 1991	10.293	0.165
Low local earnings at age 17	0.250	0.433
Nearby four-year college at age 14	0.525	0.500
Sample size	1,747	

Notes: This table reports means and standard deviations of key variables for White men in the NLSY 1979 sample. Column 1 displays mean characteristics, and column 2 shows standard deviations. AFQT score is corrected for differences in years of schooling at the time individuals took the test. Permanent local log earnings equals average log earnings from 1973 to 2000 for an individual's county of residence at age 17. Low local earnings at age 17 is an indicator equal to one if the residual from a regression of average log earnings for an individual's county at age 17 on the other covariates falls below the twenty-fifth percentile in the sample.

the ordering of the propensity score is consistent with (7) from our discussion of two binary instruments in Section IIIB. In view of the insignificant interaction term, we maintain the more parsimonious additively separable model from column 3 going forward.

This specification shows that youths experiencing low local earnings at age 17 are 10 percentage points more likely to attend college conditional on the controls, and this estimate is highly statistically significant. In the framework of Section IIIB, this implies that 10 percent of the population complies with $Z_{i,1}$, including eager compliers, reluctant compliers, and Z_1 compliers. Youths that live near a four-year college at age 14 are 6 percentage points more likely to attend college, an estimate that is also statistically significant. The larger first stage for the local earnings instrument implies that the share of Z_1 compliers exceeds the share of Z_2 compliers $(\pi_{1c} > \pi_{2c})$. However, the levels of these shares are not identified, so we have no empirical basis for assuming that $\pi_{2c} = 0$.

Columns 5 and 6 of Table 6 reveal that the two instruments generate very different estimates of the returns to college. Treating the low local earnings variable as the excluded instrument yields a positive and significant 2SLS estimate of 19.8 percent per year, while excluding the nearby college variable produces an insignificant estimated return of negative 2.4 percent. Though controlling for omitted instruments can be important for the interpretation of IV estimates (as discussed in Section IIIE), in this case we found that the 2SLS estimates were insensitive to including the other instrument as a maintained control. The instrument-specific estimates in columns 5 and 6 are outside the typical range of estimated returns to college attendance (usually around 10 percent; see Card 1999), but both estimates are also statistically imprecise.

TABLE 6—OLS AND 2SLS ESTIMATES OF THE RETURNS TO COLLEGE

| | OLS | | 2SLS | | | | |
| | | | First stage | | Second stage excluding: | | |
	No controls (1)	Controls (2)	No interaction (3)	Interaction (4)	Local earnings (5)	Nearby college (6)	Both (7)
College attendance (return per year)	0.085 (0.006)	0.042 (0.006)			0.198 (0.065)	−0.024 (0.100)	0.137 (0.049)
Low local earnings at 17		0.103 (0.033)	0.104 (0.022)	0.124 (0.033)		0.092 (0.051)	
Nearby four-year college at 14		0.000 (0.027)	0.061 (0.023)	0.069 (0.025)	−0.054 (0.033)		
Low local earnings × nearby college		−0.067 (0.046)		−0.036 (0.045)			
Corrected AFQT score		0.096 (0.014)	0.232 (0.011)	0.232 (0.011)	−0.049 (0.062)	0.157 (0.094)	0.008 (0.047)
Mother's years of schooling		0.013 (0.005)	0.030 (0.004)	0.030 (0.004)	−0.006 (0.010)	0.021 (0.014)	0.001 (0.008)
Number of siblings		−0.005 (0.005)	−0.018 (0.005)	−0.018 (0.005)	0.006 (0.008)	−0.010 (0.009)	0.002 (0.007)
Urban residence at 14		0.020 (0.024)	0.031 (0.024)	0.032 (0.024)	−0.001 (0.029)	0.027 (0.028)	0.002 (0.027)
Permanent local log earnings		0.073 (0.073)	−0.081 (0.067)	−0.083 (0.067)	0.127 (0.088)	0.056 (0.068)	0.060 (0.070)
Local log earnings in 1991		0.727 (0.078)	0.224 (0.069)	0.228 (0.069)	0.581 (0.104)	0.780 (0.115)	0.641 (0.090)
First stage F-statistic for excluded instruments					21.970	7.182	15.451
p-value: overidentification test							0.080
Sample size	1,747	1,747	1,747	1,747	1,747	1,747	1,747

Notes: This table reports ordinary least squares (OLS) and two-stage least squares (2SLS) estimates of the effects of college attendance on log wages for White men in the NLSY 1979. Columns 1 and 2 show coefficients from OLS regressions of log hourly wages on college attendance. Columns 3–7 display coefficients from two-stage least squares models instrumenting college attendance with low earnings in an individual's county at 17 and the presence of a four-year public college in an individual's county at 14. Columns 3 and 4 show first stage specifications with and without the interaction of the two instruments. Column 5 displays the estimated second stage using the low local earnings variable as the excluded instrument, column 6 shows the second stage using the nearby college variable as the excluded instrument, and column 7 displays the second stage excluding both instruments. The second stage estimates are based on the additively separable first stage in column 3. Columns 2–7 also control for cohort indicators. Returns to college are expressed annually by dividing the college coefficient by four. Robust standard errors in parentheses.

As can be seen in column 7, excluding both instruments from the second stage yields a more reasonable and precisely estimated return of 13.7 percent per year. At the same time, the standard rationale for this procedure relies on the strong restrictions on choice behavior underlying Assumption IAM. Moreover, the overidentification test for the combined model is rejected at marginal significance levels ($p = 0.08$), suggesting heterogeneity in returns across subpopulations affected by each instrument. This raises the possibility that violations of Assumption IAM could contaminate the combined 2SLS estimates.

C. *Assessing IA Monotonicity*

We explore violations of Assumption IAM by comparing estimates of marginal treatment effects (MTEs) for the two instruments (Heckman and Vytlacil 1999, 2005, 2007). The MTEs are based on the model

$$(11) \qquad\qquad D_i(z) \;=\; \mathbf{1}\big[p(z,X_i) \;\geq\; U_i\big],$$

$$(12) \qquad\qquad \big(Y_i(1), Y_i(0), U_i\big) \perp\!\!\!\perp Z_i \,|\, X_i,$$

where U_i is distributed uniformly, conditional on X_i and Z_i. Vytlacil (2002) showed that the model defined by (11) and (12) is equivalent to the framework of IA. In particular, the separable threshold-crossing representation of potential treatment choices in equation (11) is equivalent to Assumption IAM, with the homogeneous ordering of preferences over the instruments represented by the one-dimensional unobservable U_i.

Heckman and Vytlacil (2005) define the MTE as the average treatment effect at a percentile of the unobserved cost U_i and value of the covariates X_i, that is,

$$\mathrm{MTE}(u,x) \;\equiv\; E\big[Y_i(1) \;-\; Y_i(0)\,|\, U_i \;=\; u, X_i \;=\; x\big].$$

They show that if there is continuous variation in the probability of treatment, then the MTEs are identified by derivatives of average outcomes with respect to the propensity score: $\mathrm{MTE}(u,x) \;=\; \partial E\big[Y_i\,|\,p(Z_i, X_i) \;=\; u, X_i \;=\; x\big]/\partial u$.

Assumption IAM (or the equivalent (11)) implies that these derivatives should be the same regardless of which instrument generates variation in the propensity score. This is no longer the case when Assumption IAM is not true. The reason can be seen from Figure 2 panel B and Proposition 2, which showed that the marginal rate of substitution between two instruments must be constant under Assumption IAM. If the marginal rate of substitution is heterogeneous, so that Assumption IAM is not satisfied, then a small change in the propensity score due to the first instrument induces different individuals to enter treatment than a small change due to the second instrument. If these individuals have different treatment effects, then the resulting MTEs will also be different. Comparing MTEs generated by different instruments therefore provides a test of Assumption IAM.[21]

Since MTEs are not nonparametrically identified when the instruments are discrete, we use parametric models for $p(Z_i, X_i)$ and $E\big[Y_i\,|\,p(Z_i, X_i), X_i\big]$ to estimate MTE curves separately for each instrument. This amounts to extrapolating from the LATEs identified by each instrument to compare treatment effects at a common value of the propensity score, which must be the same if Assumption IAM holds.[22] We estimate MTEs by first fitting a logit model for $p(Z_i, X_i)$, then running local regressions of Y_i on X_i and a quadratic function of the estimated propensity score, weighted with a

---

[21] This observation is based on Mogstad, Torgovitsky, and Walters (2020), who develop MTE methods for multiple IVs that enable testing Assumption IAM and aggregating information across instruments under Assumption PM.

[22] See Brinch, Mogstad, and Wiswall (2017); Mogstad, Santos, and Torgovitsky (2018); Mogstad and Torgovitsky (2018); and Kline and Walters (2019) for related discussion of extrapolation based on discrete instruments.

triangle kernel function maximized at the target value of u. We allow for covariate heterogeneity by interacting the linear propensity score term with elements of X_i. The MTE estimates are derivatives of the fitted conditional mean function evaluated at the sample average of the covariates. Paralleling the 2SLS models in columns 4 and 5 of Table 6, we exclude each instrument separately while including the other as a control variable, thereby producing separate MTE estimates based on the separate variation in each instrument. We focus on MTEs for values of u from 0.2 to 0.8 because estimates at extreme values of u were imprecise. Inference is conducted based on 1,000 bootstrap replications of the entire procedure.

The results reveal that the two instruments generate different MTE schedules. Panel A of Figure 4 shows that the local earnings instrument yields a relatively flat MTE curve, suggesting limited variation in treatment effects as a function of unobserved responsiveness to labor market opportunity costs. In contrast, the MTE curve for the nearby college instrument is sharply downward sloping, implying substantial treatment effect heterogeneity along the distance dimension. As a result, MTEs based on the nearby college instrument tend to be larger at low values of u and smaller at high values of u compared to those based on the local earnings instrument. The bootstrap confidence intervals displayed in panel B show that the differences in MTE schedules are statistically significant at low and high values of u, and a joint test rejects equality of the two sets of MTEs at conventional levels ($p = 0.046$). Subject to the modeling restrictions required to estimate MTEs with discrete instruments, these results indicate that Assumption IAM does not hold jointly for the local earnings and nearby college instruments.

D. *Implications for Combined* 2*SLS Estimates*

Our results so far suggest that Assumption IAM may be violated for the combination of instruments considered here. At the same time, estimates based on the individual instruments are imprecise, so it may still be useful to combine these instruments via 2SLS. How do violations of Assumption IAM affect the interpretation of the combined 2SLS results in Table 6?

We answer this question by applying Propositions 5 and 6 to the CHV sample. Table 7 reports coefficients from regressions of D_i on each instrument separately along with the coefficient from a regression of $Z_{i,2}$ on $Z_{i,1}$. These models also control for X_i. Column 1 demonstrates that controlling for the covariates (but not the other instrument), the correlation between each instrument and the treatment is positive and statistically significant. The formulas in Proposition 5 then imply that the weights for each complier group must be positive under Assumption PM even if both Z_1 and Z_2 compliers are present. Similarly, column 2 shows that the partial correlation between the two instruments is also positive. By Proposition 6, this implies that the 2SLS weights are positive even if Assumption IAM is violated. The joint distribution of the two instruments therefore turns out to be sufficient to yield positive weights in this case.

The final rows of Table 7 use these results in formal statistical tests for positive and negative 2SLS weights. We consider tests of both the null hypothesis that the weights are all positive, and the complimentary hypothesis that at least one weight
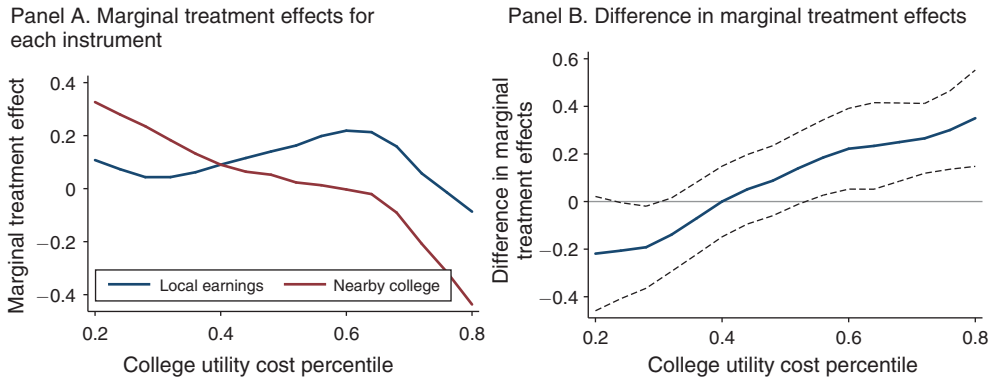
Panel A. Marginal treatment effects for each instrument

Panel B. Difference in marginal treatment effects

Figure 4. Marginal Treatment Effects of College Attendance Based on Local Earnings and Nearby College Instruments

Notes: This figure plots estimates of marginal treatment effects (MTEs) of a year of college attendance on log wages using low average log earnings in an individual's county at age 17 (local earnings) and the presence of a college in an individual's county at age 14 (nearby college) as instruments. Panel A plots separate MTE estimates based on each instrument, and panel B plots the difference in estimated MTEs with 90 percent confidence intervals based on 1,000 bootstrap replications. MTE estimates are estimated derivatives of the conditional mean log wage with respect to the propensity score evaluated at the mean of the controls. The predicted probability of college attendance comes from a logit regression of college attendance on the two instruments and controls. The conditional mean of log wages is estimated based on local regressions of log wages on a quadratic function of the estimated propensity score and controls. Local regressions are evaluated at increments of 0.04 from 0.2 to 0.8 using a triangle kernel and bandwidth of 0.32. The controls in both steps include quadratic functions of AFQT, mother's education, number of siblings, permanent local earnings, and local earnings in 1991, along with urban status and cohort indicators. AFQT and mother's education are interacted with the instruments in the logit model and with the linear propensity score term in the outcome regressions. The MTE estimates using a given instrument include the other instrument and its interactions with the linear propensity score term, AFQT, and mother's education as additional controls.

is negative.[23] As expected given the strong positive correlations in Table 7, the null hypothesis of negative weights is rejected at conventional levels ($p = 0.001$), while a test of the null of positive weights does not reject and generates a high p-value of 1.0. These findings show that in the CHV application, the combined 2SLS estimand retains an interpretation as a positively weighted average of treatment effects when Assumption IAM is replaced by the weaker Assumption PM. This has been emphasized elsewhere as an attractive property of estimands in settings with treatment effect heterogeneity (e.g., Angrist and Pischke 2009). Our results demonstrate that if estimating a positive weighted average of treatment effects is the researcher's goal, then they may be justified in combining instruments via 2SLS without relying on the onerous homogeneity restrictions implied by IA monotonicity.

V. Conclusion

The IA monotonicity condition is a cornerstone of modern IV analysis. It is appealed to often, but rarely justified explicitly. As we have shown, it will not hold when there are multiple IVs without severe restrictions on choice heterogeneity.

---

[23] The tests, which were implemented using our companion Stata package mivcausal, are described in more detail in online Appendix C. The p-value reported for the null of positive weights is from the test based on Romano, Shaikh, and Wolf (2014).

Table 7—Testing for Positive 2SLS Weights

	College attendance (1)	Nearby college (2)
Low local earnings at 17	0.110 (0.022)	0.102 (0.024)
Nearby college at 14	0.071 (0.023)	1.000 —
p-value: positive weights	1.000	
p-value: negative weights	0.001	

Notes: This table displays regressions of the variable listed in each column on the variable listed in each row in the NLSY sample. All models control for covariates. Robust standard errors are in parentheses. The first *p*-value comes from a test of the null hypothesis that the 2SLS weights are all positive, and the second comes from a test of the null hypothesis that at least one weight is negative.

This creates a dilemma for using IV methods to aggregate findings into a larger body of knowledge: each instrument is associated with a different set of complier groups, but combining multiple IVs together using IA monotonicity requires assuming that these groups are effectively identical in terms of their choice behavior.

In this paper, we have clarified the implications of IA monotonicity and considered combining instruments under a weaker condition called partial monotonicity. Partial monotonicity is not a strong assumption about choice behavior. It is satisfied under the usual mathematical interpretation of monotonicity that each instrument encourages all individuals either toward or away from treatment. We have shown that it still preserves the interpretation of the 2SLS estimand as a positively weighted average of causal effects for complier groups, except in rare cases. The conditions for positive weights turn out to hold in our application to estimating the returns to college attendance, even while the IA monotonicity condition seems to be violated.

Our results provide theoretical and empirical guidance for researchers who wish to use 2SLS with multiple IVs. Under partial monotonicity, a set of positively weighted average treatment effects can be identified by using each instrument separately, treating the others as control variables. In some cases it may be attractive to use multiple instruments together to gain statistical precision, in which case our test procedures allow analysts to test for negative weights. When the test results indicate that the weights are positive, the resulting 2SLS estimate is interpretable as a positively weighted average of treatment effects for a set of underlying complier groups. When the test suggests negative weights, the combined 2SLS estimand may not have a causal interpretation, but the instrument-specific estimates retain their interpretation as positively weighted averages of treatment effects and can be reported separately or aggregated as desired.

It is important to recognize that the treatment effect parameter identified by 2SLS may not answer an interesting scientific question or policy counterfactual even when the weights are positive. This point motivates our companion paper (Mogstad, Torgovitsky, and Walters 2020), in which we develop a framework for aggregating multiple IVs to conduct inference about specific target parameters that do answer

well-posed counterfactual questions. The framework generalizes the approach of Mogstad, Santos, and Torgovitsky (2018) to replace IA monotonicity with partial monotonicity. The key idea is that even under partial monotonicity, each instrument still carries identifying content about unobserved quantities that are instrument invariant, such as the average treatment effect. Harmonizing this content across instruments allows for information to flow between instruments and to aggregate into a greater whole. This provides a way to combine exogenous variation from multiple sources for policy evaluation, while still maintaining plausible conditions on choice behavior.

APPENDIX

PROOF OF PROPOSITION 1:

($\Rightarrow$) Suppose that the nesting statement is not true. Then there exist $j, k \in \mathcal{I}$ such that $\mathcal{Z}_j \nsubseteq \mathcal{Z}_k$ and $\mathcal{Z}_k \nsubseteq \mathcal{Z}_j$. Since the empty set $\emptyset$ is a subset of every set (including itself), this implies that both $\mathcal{Z}_j$ and $\mathcal{Z}_k$ are not empty. Thus, there exists a $z_j \in \mathcal{Z}_j$ such that $z_j \notin \mathcal{Z}_k$, and there exists a $z_k \in \mathcal{Z}_k$ such that $z_k \notin \mathcal{Z}_j$. By the definition of these sets, this means that

(13) $$D_j(z_j) = 1 > 0 = D_j(z_k)$$

and

$$D_k(z_j) = 0 < 1 = D_k(z_k).$$

Thus, $D_i(z_j) \geq D_i(z_k)$ for some $i = j$, but $D_i(z_j) < D_i(z_k)$ for some other $i = k$. This violates Assumption IAM. By contraposition, it follows that Assumption IAM implies the nesting statement.

($\Leftarrow$) Conversely, if Assumption IAM is not true, then there exist $j, k \in \mathcal{I}$ and z_j, $z_k \in \mathcal{Z}$ such that (13) holds. By definition, (13) implies that $z_j \in \mathcal{Z}_j$, but $z_j \notin \mathcal{Z}_k$, and that $z_k \in \mathcal{Z}_k$, but $z_k \notin \mathcal{Z}_j$. That is, $\mathcal{Z}_k \nsubseteq \mathcal{Z}_j$, and $\mathcal{Z}_j \nsubseteq \mathcal{Z}_k$. It follows that the nesting statement also implies Assumption IAM. ∎

PROOF OF PROPOSITION 2:

Suppose to the contrary that there exist $j, k \in \mathcal{I}(z^*)$ for which (2) does not hold. Then the matrix

$$\partial V_{jk}(z^*) \equiv \begin{bmatrix} \partial_1 V_j(z^*) & \partial_2 V_j(z^*) \\ \partial_1 V_k(z^*) & \partial_2 V_k(z^*) \end{bmatrix} \equiv \begin{bmatrix} \partial V_j(z^*) \\ \partial V_k(z^*) \end{bmatrix}$$

is invertible. Thus, the span of the rows of $\partial V_{jk}(z^*)$ is $\mathbb{R}^2$, so there exists a unit vector $v^* \in \mathbb{R}^2$ such that $\partial V_j(z^*) v^* > 0$, while $\partial V_k(z^*) v^* < 0$. Taking a Taylor series expansion at $z^* + \epsilon v^*$ for sufficiently small $\epsilon > 0$, we have that

$$V_j(z^* + \epsilon v^*) \approx V_j(z^*) + \epsilon [\partial V_j(z^*) v^*] > 0,$$

while

$$V_k(z^* + \epsilon v^*) \approx V_k(z^*) + \epsilon[\partial V_k(z^*)v^*] < 0,$$

since $V_j(z^*) = V_k(z^*) = 0$. On the other hand, an ϵ step in the direction $-v^*$ yields

$$V_j(z^* - \epsilon v^*) < 0 \quad \text{while} \quad V_k(z^* - \epsilon v^*) > 0.$$

Using the random utility model, (1), we have that

$$D_j(z^* + \epsilon v^*) = 1 > D_j(z^* - \epsilon v^*) = 0$$

and

$$D_k(z^* + \epsilon v^*) = 0 < D_k(z^* - \epsilon v^*) = 1,$$

which shows that Assumption IAM is violated. This establishes the result through contraposition. ∎

PROOF OF PROPOSITION 3:

Take any $(z_\ell, z_{-\ell})$ and $(z'_\ell, z_{-\ell})$ in $\mathcal{Z}$. Since $z_\ell, z'_\ell \in \mathbb{R}$, either $z_\ell \geq z'_\ell$ or $z'_\ell \geq z_\ell$. Suppose that the first case holds. Then $(z_\ell, z_{-\ell}) \geq (z'_\ell, z_{-\ell})$, so $D_i(z_\ell, z_{-\ell}) \geq D_i(z'_\ell, z_{-\ell})$ for all i, as required by Assumption PM. ∎

PROOF OF PROPOSITION 4:

That an individual i cannot belong to more than one of the six groups can be verified by inspection. To see that i must belong to at least one of these groups, note that under (7), either

$$(14) \qquad\qquad D_i(0,0) \leq D_i(0,1) \leq D_i(1,0) \leq D_i(1,1)$$

or

$$(15) \qquad\qquad D_i(0,0) \leq D_i(1,0) \leq D_i(0,1) \leq D_i(1,1).$$

If i satisfies (14) then by Table 2, their group is $G_i \in \{at, nt, 1c, ec, rc\}$, that is, something other than a Z_2 complier. If i satisfies (15), then their group is something other than a Z_1 complier. In either case, they must belong to one of the six groups listed in Table 2. ∎

PROOF OF PROPOSITION 5:

Label the instrument pairs as $z^1 = (0,0)$, $z^2 = (0,1)$, $z^3 = (1,0)$, and $z^4 = (1,1)$, and denote their associated probabilities as $q^k \equiv \Pr[Z_i = z^k]$ for $k = 1, 2, 3, 4$. We will prove the result for the case with $\pi_{1c} \geq \pi_{2c}$, so that the propensity score $p(z^k) \equiv \Pr[D_i = 1 \,|\, Z_i = z^k]$ is increasing in k, due to Assumption PM and (7). A symmetric proof applies to the case with $\pi_{2c} \geq \pi_{1c}$.

Theorem 2 in IA shows that the 2SLS estimand is given by a convex weighted average of three Wald (1940) estimands, which we write as

$$(16) \qquad \beta_{2sls} = \lambda_2 w_2 + \lambda_3 w_3 + \lambda_4 w_4,$$

where the Wald estimands are given by

$$w_k \equiv \frac{E[Y_i \,|\, Z_i = z^k] - E[Y_i \,|\, Z_i = z^{k-1}]}{p(z^k) - p(z^{k-1})},$$

and the weights are defined by

$$\lambda_k \equiv \frac{\left(p(z^k) - p(z^{k-1}) \right) \sum_{\ell=k}^{4} q^\ell \left(p(z^\ell) - E[p(Z_i)] \right)}{\sum_{j=2}^{4} \left[\left(p(z^j) - p(z^{j-1}) \right) \sum_{\ell=j}^{4} q^\ell \left(p(z^\ell) - E[p(Z_i)] \right) \right]}.$$

Theorem 1 of IA shows that each Wald estimand, w_k, gives the average treatment effect for individuals who change treatment status in response to a change in the instrument from z^{k-1} to z^k. Using the group definitions in Proposition 4, this implies that w_2 represents the average treatment effect for both the Z_2 compliers and eager compliers. Similarly, w_4 reflects the average treatment effect for the Z_2 compliers and the reluctant compliers. So,

$$w_2 = \left(\frac{\pi_{2c}}{\pi_{2c} + \pi_{ec}} \right) \Delta_{2c} + \left(\frac{\pi_{ec}}{\pi_{2c} + \pi_{ec}} \right) \Delta_{ec},$$

and

$$w_4 = \left(\frac{\pi_{2c}}{\pi_{2c} + \pi_{rc}} \right) \Delta_{2c} + \left(\frac{\pi_{rc}}{\pi_{2c} + \pi_{rc}} \right) \Delta_{rc}.$$

However, w_3 is different, since a shift from $z^2 \equiv (0,1)$ to $z^3 \equiv (1,0)$ creates two-way flows. In particular, such a shift induces Z_1 compliers to take treatment, but Z_2 compliers to exit treatment. Using a minor modification of the argument in IA, it follows that

$$w_3 = \frac{E[(Y_i(1) - Y_i(0))(D_i(1,0) - D_i(0,1))]}{p(z^3) - p(z^2)}$$

$$= \left(\frac{\pi_{1c}}{p(z^3) - p(z^2)} \right) \Delta_{1c} - \left(\frac{\pi_{2c}}{p(z^3) - p(z^2)} \right) \Delta_{2c}$$

$$= \left(\frac{\pi_{1c}}{\pi_{1c} - \pi_{2c}} \right) \Delta_{1c} - \left(\frac{\pi_{2c}}{\pi_{1c} - \pi_{2c}} \right) \Delta_{2c},$$

where the last equality used $p(z^3) - p(z^2) = (\pi_{at} + \pi_{1c} + \pi_{ec}) - (\pi_{at} + \pi_{2c} + \pi_{ec}) \ \pi_{1c} - \pi_{2c}$.

Substituting the expressions for the Wald estimands into (16), we have

$$\beta_{2sls} = \sum_{g \in \{1c,2c,ec,rc\}} \omega_g \Delta_g,$$

where

$$\omega_{ec} \equiv \frac{\lambda_2 \pi_{ec}}{\pi_{ec} + \pi_{2c}},$$

$$\omega_{1c} \equiv \frac{\lambda_3 \pi_{1c}}{\pi_{1c} - \pi_{2c}},$$

$$\omega_{rc} \equiv \frac{\lambda_4 \pi_{rc}}{\pi_{rc} + \pi_{2c}}, \quad \text{and}$$

$$\omega_{2c} \equiv \left(\frac{\lambda_2 \pi_{2c}}{\pi_{ec} + \pi_{2c}} \right) - \left(\frac{\lambda_3 \pi_{2c}}{\pi_{1c} - \pi_{2c}} \right) + \left(\frac{\lambda_4 \pi_{2c}}{\pi_{rc} + \pi_{2c}} \right).$$

It is straightforward to verify that ω_{ec}, ω_{1c}, and ω_{rc} are each nonnegative, and that

$$\sum_{g \in \{1c,2c,ec,rc\}} \omega_g = \lambda_2 + \lambda_3 + \lambda_4 = 1.$$

For ω_{2c}, note first that

$$\pi_{ec} + \pi_{2c} = p(z^2) - p(z^1),$$

and that, similarly,

$$\pi_{1c} - \pi_{2c} = p(z^3) - p(z^2) \quad \text{and} \quad \pi_{rc} + \pi_{2c} = p(z^4) - p(z^3).$$

Substituting this observation and the definition of λ_k into the expression for ω_{2c} and simplifying, we have

$$\omega_{2c} = \pi_{2c} \times \frac{q^2 \left(p(z^2) - E[p(Z_i)] \right) + q^4 \left(p(z^4) - E[p(Z_i)] \right)}{\sum_{j=2}^{4} \left[\left(p(z^j) - p(z^{j-1}) \right) \sum_{\ell=j}^{4} q^\ell \left(p(z^\ell) - E[p(Z_i)] \right) \right]}.$$

The denominator of this expression is always positive and π_{2c} is always nonnegative. For the numerator, notice that since $Z_{i,2} = 1$ if and only if $Z_i \in \{z^2, z^4\}$,

$$q^2 \left(p(z^2) - E[p(Z_i)] \right) + q^4 \left(p(z^4) - E[p(Z_i)] \right) = E \left[Z_{i,2} \left(p(Z_i) - E[p(Z_i)] \right) \right]$$

$$= E \left[Z_{i,2} \left(D_i - E[D_i] \right) \right]$$

$$\equiv \text{cov}(D_i, Z_{i,2}),$$

where the second equality follows by iterating expectations. Thus, the sign of ω_{2c} is the same as that of $\text{cov}(D_i, Z_{i,2})$, which is in turn the same as the sign of $E[D_i | Z_{i,2} = 1] - E[D_i | Z_{i,2} = 0]$, since

$$\text{cov}(D_i, Z_{i,2}) = \left(E[D_i | Z_{i,2} = 1] - E[D_i | Z_{i,2} = 0]\right)\Pr[Z_{i,2} = 1]\Pr[Z_{i,2} = 0]. \blacksquare$$

PROOF OF PROPOSITION 6:

Since $Z_{i,1}$ and $Z_{i,2}$ are binary,

$$\Pr[D_i = 1 | Z_{i,2} = 1] - \Pr[D_i = 1 | Z_{i,2} = 0]$$

$$= p(1,1)\Pr[Z_{i,1} = 1 | Z_{i,2} = 1] + p(0,1)\Pr[Z_{i,1} = 0 | Z_{i,2} = 1]$$

$$- p(1,0)\Pr[Z_{i,1} = 1 | Z_{i,2} = 0] - p(0,0)\Pr[Z_{i,1} = 0 | Z_{i,2} = 0].$$

Assumption PM with (7) together with Assumption E imply that

$$p(1,1) \equiv \Pr[D_i = 1 | Z_{i,1} = 1, Z_{i,2} = 1]$$

$$= \Pr[D_i(1,1) = 1] \geq \Pr[D_i(0,1) = 1] = p(0,1),$$

and similarly that $p(1,0) \geq p(0,0)$ and $p(0,1) \geq p(0,0)$. If $\text{cov}(Z_{i,1}, Z_{i,2}) \geq 0$, then also $\Pr[Z_{i,1} = 1 | Z_{i,2} = 1] \geq \Pr[Z_{i,1} = 1] \geq \Pr[Z_{i,1} = 1 | Z_{i,2} = 0]$. Thus,

$$\Pr[D_i = 1 | Z_{i,2} = 1] - \Pr[D_i = 1 | Z_{i,2} = 0]$$

$$\geq [p(1,1) - p(1,0)]\Pr[Z_{i,1} = 1]$$

$$+ [p(0,1) - p(0,0)](1 - \Pr[Z_{i,1} = 1]) \geq 0.$$

By Proposition 5, this implies that $\omega_{2c} \geq 0$. A symmetric argument shows that $\Pr[D_i = 1 | Z_{i,1} = 1] - \Pr[D_i = 1 | Z_{i,1} = 0] \geq 0$, so that $\omega_{1c} \geq 0$ as well. $\blacksquare$

PROOF OF PROPOSITION 7:

By Theorem 2 in IA,

$$(17) \qquad\qquad \beta_{2sls} = \sum_{k=2}^{K} \lambda_k w_k,$$

where the Wald estimands are

$$w_k \equiv \frac{E[Y_i | Z_i = z^k] - E[Y_i | Z_i = z^{k-1}]}{p(z^k) - p(z^{k-1})},$$

and the weights are

$$(18) \qquad \lambda_k \equiv \frac{\left(p(z^k) - p(z^{k-1})\right)\sum_{\ell=k}^{K}q^\ell\left(p(z^\ell) - E[p(Z_i)]\right)}{\sum_{j=2}^{K}\left[\left(p(z^j) - p(z^{j-1})\right)\sum_{\ell=j}^{K}q^\ell\left(p(z^\ell) - E[p(Z_i)]\right)\right]}.$$

By Assumption E,

$$(19) \qquad w_k = \frac{E\left[Y_i\left(D_i(z^k)\right) - Y_i\left(D_i(z^{k-1})\right)\right]}{p(z^k) - p(z^{k-1})}$$

$$= \frac{\sum_{g\in\mathcal{G}}E\left[Y_i\left(D_i(z^k)\right) - Y_i\left(D_i(z^{k-1})\right)\mid G_i = g\right]\pi_g}{p(z^k) - p(z^{k-1})}$$

$$= \frac{\sum_{g:k\in\mathcal{C}_g}\Delta_g\pi_g - \sum_{g:k\in\mathcal{D}_k}\Delta_g\pi_g}{p(z^k) - p(z^{k-1})},$$

since $Y_i\left(D_i(z^k)\right) - Y_i\left(D_i(z^{k-1})\right) = 0$ except when $k \in \mathcal{C}_{G_i}$ or $k \in \mathcal{D}_{G_i}$. Substituting (19) into (17),

$$(20) \quad \beta_{2sls} = \sum_{k=2}^{K}\lambda_k\left(\frac{\sum_{g:k\in\mathcal{C}_g}\Delta_g\pi_g - \sum_{gk\in\mathcal{D}_g}\Delta_g\pi_g}{p(z^k) - p(z^{k-1})}\right)$$

$$= \sum_{k=2}^{K}\lambda_k\left(\frac{\sum_{g\in\mathcal{G}}\left(\mathbf{1}[k \in \mathcal{C}_g] - \mathbf{1}[k \in \mathcal{D}_g]\right)\Delta_g\pi_g}{p(z^k) - p(z^{k-1})}\right)$$

$$= \sum_{g\in\mathcal{G}}\left(\pi_g\sum_{k=2}^{K}\left(\mathbf{1}[k \in \mathcal{C}_g] - \mathbf{1}[k \in \mathcal{D}_g]\right)\left(\frac{\lambda_k}{p(z^k) - p(z^{k-1})}\right)\right)\Delta_g$$

$$\equiv \sum_{g\in\mathcal{G}}\omega_g\Delta_g.$$

Substituting the definition of λ_k from (18) and simplifying,

$$(21) \quad \omega_g = \pi_g\sum_{k=2}^{K}\left(\mathbf{1}[k \in \mathcal{C}_g] - \mathbf{1}[k \in \mathcal{D}_g]\right)$$

$$\times \left(\frac{\sum_{\ell=k}^{K}q^\ell\left(p(z^\ell) - E[p(Z_i)]\right)}{\sum_{j=2}^{K}\left[\left(p(z^j) - p(z^{j-1})\right)\sum_{\ell=j}^{K}q^\ell\left(p(z^\ell) - E[p(Z_i)]\right)\right]}\right)$$

$$= \pi_g\sum_{k=2}^{K}\left(\mathbf{1}[k \in \mathcal{C}_g] - \mathbf{1}[k \in \mathcal{D}_g]\right)\left(\frac{\mathrm{cov}\left(D_i, \mathbf{1}[p(Z_i) \geq p(z^k)]\right)}{\mathrm{var}(p(Z_i))}\right),$$

where in the numerator we used

$$\text{cov}\left(D_i, \mathbf{1}\left[p(Z_i) \geq p(z^k)\right]\right) = E\left[\mathbf{1}\left[p(Z_i) \geq p(z^k)\right](D_i - E[D_i])\right]$$

$$= E\left[\mathbf{1}\left[p(Z_i) \geq p(z^k)\right]\left(p(Z_i) - E[p(Z_i)]\right)\right]$$

$$= \sum_{\ell=k}^{K} q^\ell \left(p(z^\ell) - E[p(Z_i)]\right),$$

and in the denominator we used

$$\text{var}(p(Z_i)) = E\left[D_i\left(p(Z_i) - E[p(Z_i)]\right)\right]$$

$$= \sum_{\ell=1}^{K} p(z^\ell)\left(p(z^\ell) - E[p(Z_i)]\right)q^\ell$$

$$= \sum_{\ell=1}^{K}\left(\sum_{j=2}^{K}\mathbf{1}[j \leq \ell]\left(p(z^j) - p(z^{j-1})\right)\right)\left(p(z^\ell) - E[p(Z_i)]\right)q^\ell$$

$$= \sum_{j=2}^{K}\left[\left(p(z^j) - p(z^{j-1})\right)\sum_{\ell=j}^{K}\left(p(z^\ell) - E[p(Z_i)]\right)q^\ell\right],$$

where the third equality follows from a telescoping sum identity,[24] together with the fact that $\sum_{\ell=1}^{K}\left(p(z^\ell) - E[p(Z_i)]\right)q^\ell = 0$. Examining the expression for the weights in (21), we have that

$$\text{sgn}(\omega_g) = \mathbf{1}[\pi_g > 0]$$

$$\times \text{sgn}\left(\sum_{k=2}^{K}\left(\mathbf{1}[k \in \mathcal{C}_g] - \mathbf{1}[k \in \mathcal{D}_g]\right)\text{cov}\left(D_i, \mathbf{1}\left[p(Z_i) \geq p(z^k)\right]\right)\right).$$

It remains to show that $\omega_g = 0$ when $\mathcal{C}_g = \emptyset$, so that only groups that comply with at least one instrument contrast receive weight in the 2SLS estimand. To see that this is so, suppose to the contrary that there is a group g with $\pi_g > 0$ for which $\mathcal{C}_g = \emptyset$, while $\omega_g \neq 0$. Given the structure of ω_g, such a group must have $\mathcal{D}_g \neq \emptyset$. That is, this group must defy at some instrument contrast, even though they do not comply at any other instrument contrasts. We will prove that such a "pure defier" group cannot exist under Assumption PM by establishing a contradiction.

Let $j_0 \in \mathcal{D}_g$ be the instrument contrast at which the "pure defier" group g defies. By definition, $D_i(z^{j_0}) = 0$, while $D_i(z^{j_0-1}) = 1$. Since $\mathcal{C}_g = \emptyset$, it follows that for any i with $G_i = g$,

$$(22) \qquad\qquad D_i(z^j) = \mathbf{1}[j < j_0].$$

---

[24] In particular, that $a^\ell = a^1 + \sum_{j=2}^{K}\mathbf{1}[j \leq \ell](a^j - a^{j-1})$ for any scalars $\{a^\ell\}_{\ell=1}^{K}$.

In particular, $D_i(z^1) = 1$, while $D_i(z^K) = 0$.

To proceed, it will be helpful to use the following terminology. We call two vectors z^j and z^k *pm-comparable* if they differ in only one component. That is, z^j and z^k are pm-comparable if there exists an $\ell' \in \{1, \ldots, L\}$ such that $z_\ell^j = z_\ell^k$ for all $\ell \neq \ell'$. If z^j and z^k are pm-comparable, then Assumption PM requires that either $D_i(z^j) \leq D_i(z^k)$ for all $i \in \mathcal{I}$ or that $D_i(z^j) \geq D_i(z^k)$ for all $i \in \mathcal{I}$. Moreover, as we show in Lemma 1, if $j \leq k$, then there cannot exist a group g^* with $\pi_{g^*} > 0$ for which individuals i in group g^* have $D_i(z^j) > D_i(z^k)$. We will now use this result to show that the existence of the "pure defier" group g defined above creates a contradiction.

Let z^{j_1} be the vector whose first component is the same as that of the largest propensity-score instrument value, z^K, while all other components are the same as the smallest, z^1. That is,

$$z^{j_1} \equiv (z_1^K, z_{-1}^1).$$

Then z^{j_1} and z^1 are pm-comparable, and $z^{j_1} \in \text{supp}(Z_i)$, which we have assumed is rectangular. Since $D_i(z^1) = 1$ for any i with $G_i = g$ and $p(z^1)$ is the smallest propensity score value, it follows from Lemma 1 that $D_i(z^{j_1}) = 1$ for these individuals as well. Thus by (22), it must be that $j_1 < j_0$.

Now let z^{j_2} be the same as z^{j_1} except with its second component replaced by z_2^K. That is,

$$z^{j_2} \equiv (z_2^K, z_{-2}^{j_1}) \equiv (z_1^K, z_2^K, z_3^1, \ldots, z_L^1).$$

Then z^{j_2} is pm-comparable to z^{j_1}, and $z^{j_2} \in \text{supp}(Z_i)$. If it were the case that $j_2 \geq j_0$, then $p(z^{j_2}) \geq p(z^{j_0}) \geq p(z^{j_1})$, so that Lemma 1 would imply that $D_i(z^{j_2}) = 1$ for individuals with $G_i = g$. At the same time, (22) would imply that $D_i(z^{j_2}) = 0$ for these individuals, yielding a contradiction. Thus, it must be that $j_2 < j_0$.

Continuing in this way, we find a sequence of vectors $z^{j_1}, z^{j_2}, z^{j_3}, \ldots, z^{j_L}$ that each differ from z^K in one component less than its predecessor, and such that $j_\ell < j_0$ for each ℓ. This process ends once we reach j_L, at which point $z^{j_L} = z^K$ is the instrument value corresponding to the largest propensity score value. However, this implies a contradiction, because $j_L < j_0$ while at the same time $j_L = K \geq j_0$. We conclude that a "pure defier" group cannot exist under Assumption PM, and therefore that the sum in (20) only needs to be indexed over $g \in \mathcal{G}$ for which $\mathcal{C}_g \neq \emptyset$. ∎

LEMMA 1: *Suppose that Assumption E holds. Let $\mathcal{G}$ denote the set of all realizations of $\{D_i(z)\}_{z \in \mathcal{Z}}$ that are consistent with Assumption PM. Suppose that z and z' are pm-comparable and that $p(z) \leq p(z')$. Then there does not exist a group $g^* \in \mathcal{G}$ such that $\pi_{g^*} > 0$ and $D_i(z) > D_i(z')$ for all i with $G_i = g^*$.*

PROOF:

Since z and z' are pm-comparable, Assumption PM requires that $D_i(z) \leq D_i(z')$ for all $i \in \mathcal{I}$, or $D_i(z) \geq D_i(z')$ for all $i \in \mathcal{I}$. If such a group g^* did exist, then the latter case would need to hold. However, this would imply that

$$
\begin{aligned}
p(z) &\equiv \Pr[D_i = 1 \,|\, Z_i = z] \\
&= \Pr[D_i(z) = 1] \\
&= \Pr[D_i(z) = 1 \,|\, G_i = g^*]\pi_{g^*} + \Pr[D_i(z) = 1 \,|\, G_i \neq g^*](1 - \pi_{g^*}) \\
&> \Pr[D_i(z') = 1 \,|\, G_i = g^*]\pi_{g^*} + \Pr[D_i(z') = 1 \,|\, G_i \neq g^*](1 - \pi_{g^*}) \\
&= \Pr[D_i = 1 \,|\, Z_i = z'] \equiv p(z'),
\end{aligned}
$$

which contradicts the assumption that $p(z) \leq p(z')$. $\blacksquare$

REFERENCES

Angrist, Joshua D., and William N. Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* 88 (3): 450–77.

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): 431–42.

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): 444–55.

Angrist, Joshua D., and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106 (4): 979–1014.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton and Oxford: Princeton University Press.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes. 2006. "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments." *American Economic Review* 96 (4): 988–1012.

Bitler, Marianne P., Hilary Hoynes, and Thurston Domina. 2014. "Experimental Evidence on Distributional Effects of Head Start." Unpublished.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall. 2017. "Beyond LATE with a Discrete Instrument." *Journal of Political Economy* 125 (4): 985–1039.

Bureau of Labor Statistics. 2019. "National Longitudinal Survey of Youth 1979 Cohort." United States Department of Labor. https://www.bls.gov/nls/nlsy79.htm.

Cameron, Stephen V., and James J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *Journal of Political Economy* 106 (2): 262–333.

Cameron, Stephen V., and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." *Journal of Political Economy* 112 (1): 132–82.

Card, David. 1995. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspects of Labour Market Behavior: Essays in Honour of John Vanderkamp*, edited by Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, 201–22. Toronto: University of Toronto Press.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, Vol. 3A, edited by Orley C. Ashenfelter and David Card, 1801–63. Amsterdam: North-Holland.

Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (5): 1127–60.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101 (6): 2754–81.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2019. "Replication Data for: Estimating Marginal Returns to Education." *American Economic Association* [publisher]. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E112467V1.

Carneiro, Pedro, and Sokbae Lee. 2009. "Estimating Distributions of Potential Outcomes Using Local Instrumental Variables with an Application to Changes in College Enrollment and Wage Inequality." *Journal of Econometrics* 149 (2): 191–208.

Carneiro, Pedro, Michael Lokshin, and Nithin Umapathi. 2016. "Average and Marginal Returns to Upper Secondary Schooling in Indonesia." *Journal of Applied Econometrics* 32 (1): 16–36.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg. 2018. "Who Benefits from Universal Childcare? Estimating Marginal Returns to Early Childcare Attendance." *Journal of Political Economy* 126 (6): 2356–2409.

Doyle, Joseph J., Jr. 2007. "Child Protection and Child Outcomes: Measuring the Effects of Foster Care." *American Economic Review* 97 (5): 1583–1610.

Felfe, Christina, and Rafael Lalive. 2018. "Does Early Child Care Affect Children's Development?" *Journal of Public Economics* 159: 33–53.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77 (3): 953–73.

French, Eric, and Jae Song. 2014. "The Effect of Disability Insurance Receipt on Labor Supply." *American Economic Journal: Economic Policy* 6 (2): 291–337.

Havnes, Tarjei, and Magne Mogstad. 2015. "Is Universal Child Care Leveling the Playing Field?" *Journal of Public Economics* 127: 100–114.

Heckman, James J. 2001. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109 (4): 673–748.

Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *Review of Economics and Statistics* 88 (3): 389–432.

Heckman, James J., and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73 (3): 669–738.

Heckman, James J., and Edward J. Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences* 96 (8): 4730–34.

Heckman, James J., and Edward J. Vytlacil. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments." In *Handbook of Econometrics*, Vol. 6B, edited by James J. Heckman and Edward E. Leamer, 4875–5143. Amsterdam: North-Holland.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.

Kane, Thomas J., and Cecilia E. Rouse. 1993. "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?" NBER Working Paper 4268.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *Quarterly Journal of Economics* 131 (3): 1057–1111.

Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *Quarterly Journal of Economics* 131 (4): 1795–1848.

Kline, Patrick, and Christopher R. Walters. 2019. "On Heckits, LATE, and Numerical Equivalence." *Econometrica* 87 (2): 677–96.

Kling, Jeffrey R. 2001. "Interpreting Instrumental Variables Estimates of the Returns to Schooling." *Journal of Business and Economic Statistics* 19 (3): 358–64.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2013. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." *American Economic Review* 103 (5): 1797–1829.

Milgrom, Paul, and Chris Shannon. 1994. "Monotone Comparative Statics." *Econometrica* 62 (1): 157–80.

Moffitt, Robert. 2008. "Estimating Marginal Treatment Effects in Heterogeneous Populations." *Annales d'Économie et de Statistique* 91–92: 239–61.

Mogstad, Magne, Andres Santos, and Alexander Torgovitsky. 2018. "Using Instrumental Variables for Inference about Policy Relevant Treatment Parameters." *Econometrica* 86 (5): 1589–1619.

Mogstad, Magne, and Alexander Torgovitsky. 2018. "Identification and Extrapolation of Causal Effects with Instrumental Variables." *Annual Review of Economics* 10: 577–613.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R. Walters. 2020. "Policy Evaluation with Multiple Instrumental Variables." Unpublished.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R. Walters. 2021. "Replication Data for: The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables." *American Economic Association* [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/135041V1.

Mountjoy, Jack. 2019. "Community Colleges and Upward Mobility." Unpublished.

Nybom, Martin. 2017. "The Distribution of Lifetime Earnings Returns to College." *Journal of Labor Economics* 35 (4): 903–52.

Romano, Joseph P., Azeem M. Shaikh, and Michael Wolf. 2014. "A Practical Two-Step Method for Testing Moment Inequalities." *Econometrica* 82 (5): 1979–2002.

Vytlacil, Edward. 2002. "Independence, Monotonicity, and Latent Index Models: An Equivalence Result." *Econometrica* 70 (1): 331–41.

Wald, Abraham. 1940. "The Fitting of Straight Lines if Both Variables Are Subject to Error." *Annals of Mathematical Statistics* 11 (3): 284–300.

Walters, Christopher R. 2018. "The Demand for Effective Charter Schools." *Journal of Political Economy* 126 (6): 2179–2223.