When Should We (Not) Interpret Linear IV Estimands as LATE?
Online Appendix

Tymon Słoczyński

October 5, 2024

Appendices

Appendix A Proofs

Proof of Theorem 3.2. Lemma 3.1 states that $\beta_{2SLS} = \frac{E[\sigma^2(X)\cdot\tau(X)]}{E[\sigma^2(X)]}$. It remains to show that $\sigma^2(X) = [\pi(X)]^2 \cdot \text{Var}[Z \mid X]$. Indeed, it follows from the definition of $\sigma^2(X)$, equation (4), and iterated expectations that $\sigma^2(X) = [\omega(X)]^2 \cdot \text{Var}[Z \mid X]$. Then, it follows from Lemma 2.1 that $\sigma^2(X) = [\pi(X)]^2 \cdot \text{Var}[Z \mid X]$ because $[\omega(X)]^2 = [\pi(X)]^2$ under Assumptions IV and WM.

Proof of Theorem 3.3. Let R and T be generic notation for two random variables, where T is binary and R is arbitrarily discrete or continuous. The following lemma, due to Angrist (1998), will be useful for what follows.

Lemma A.1 (Angrist, 1998)**.** *Suppose that* $E[T \mid X]$ *is linear in X. Then,* ξ*, the coefficient on T in the linear projection of R on T and X can be written as*

$$\xi = \frac{E[\text{Var}[T \mid X] \cdot \xi(X)]}{E[\text{Var}[T \mid X]]},$$

where $\xi(X) = E[R \mid T = 1, X] - E[R \mid T = 0, X]$.

Recall that β_{IV} is equal to the ratio of the reduced-form and first-stage coefficients on Z. It follows that we can apply Lemma A.1 separately to these two coefficients, and thereby obtain the following expression for the estimand of interest:

$$\beta_{IV} = \frac{\frac{E[\text{Var}[Z|X]\cdot\phi(X)]}{E[\text{Var}[Z|X]]}}{\frac{E[\text{Var}[Z|X]\cdot\omega(X)]}{E[\text{Var}[Z|X]]}}, \tag{A.1}$$

where

$$\phi(x) = E[Y \mid Z = 1, X = x] - E[Y \mid Z = 0, X = x] \tag{A.2}$$

is the conditional reduced-form slope coefficient and $\omega(x)$ is as defined in equation (5). Upon rearrangement, we obtain

$$\begin{aligned}\beta_{IV} &= \frac{E[\text{Var}[Z \mid X] \cdot \phi(X)]}{E[\text{Var}[Z \mid X] \cdot \omega(X)]} \\ &= \frac{E[\text{Var}[Z \mid X] \cdot \omega(X) \cdot \beta(X)]}{E[\text{Var}[Z \mid X] \cdot \omega(X)]},\end{aligned} \tag{A.3}$$

where the second equality uses the definition of $\beta(x)$ in equation (6). See also Walters (2018) for a similar argument. Finally, we know from Lemma 2.1 that $\beta(x) = \tau(x)$ and $\omega(x) = c(x) \cdot \pi(x)$ under

Assumptions IV and WM. This completes the proof because β_{IV} can now be written as

$$\beta_{\mathrm{IV}} = \frac{\mathrm{E}\left[c(X) \cdot \pi(X) \cdot \mathrm{Var}\left[Z \mid X\right] \cdot \tau(X)\right]}{\mathrm{E}\left[c(X) \cdot \pi(X) \cdot \mathrm{Var}\left[Z \mid X\right]\right]}. \tag{A.4}$$

Alternative Proof of Theorem 3.3. The following proof of Theorem 3.3 uses Kolesár (2013)'s result in Remark 3.4. Let us begin by restating the representation of two-step IV estimands in equation (13):

$$\beta_{\mathrm{TSIV}} = \int \sum_{j=1}^{J_x-1} \frac{\theta_j(x)}{\int \sum_{j=1}^{J_x-1} \theta_j(x)\,\mathrm{d}F^X(x)}\, \tau(p_{j,x}; x)\,\mathrm{d}F^X(x). \tag{A.5}$$

The notation is the same as in Remark 3.4. The representation in equation (A.5) is appropriate for any two-step IV estimand (*e.g.*, 2SLS) which uses $Z_{\mathrm{G}} = z_{\mathrm{G}}(X, Z)$ as instruments, as long as D is binary, Z is discrete, and the relevant assumptions are satisfied. When Z is binary and $Z_{\mathrm{G}} = Z$, we get $\tau(p_{1,x}; x) = \tau(x)$ and $\beta_{\mathrm{TSIV}} = \beta_{\mathrm{IV}}$, and we can use equation (A.5) to write

$$\begin{aligned} \beta_{\mathrm{IV}} &= \int \frac{\theta_1(x)}{\int \theta_1(x)\,\mathrm{d}F^X(x)}\, \tau(x)\,\mathrm{d}F^X(x) \\ &= \frac{\mathrm{E}\left[\theta_1(X) \cdot \tau(X)\right]}{\mathrm{E}\left[\theta_1(X)\right]}, \end{aligned} \tag{A.6}$$

where

$$\begin{aligned} \theta_1(x) &= (p_{2,x} - p_{1,x}) \cdot \mathrm{P}\left[P > p_{1,x} \mid X = x\right] \cdot \mathrm{E}\left[\tilde{P}^L \mid X = x, P > p_{1,x}\right] \\ &= \left|\mathrm{E}\left[D \mid Z = 1, X = x\right] - \mathrm{E}\left[D \mid Z = 0, X = x\right]\right| \cdot \mathrm{P}\left[P > p_{1,x} \mid X = x\right] \\ &\quad \cdot \mathrm{E}\left[\tilde{P}^L \mid X = x, P > p_{1,x}\right] \\ &= \pi(x) \cdot \Big(1[\omega(x) > 0] \cdot \mathrm{P}\left[Z = 1 \mid X = x\right] \cdot \mathrm{E}\left[\tilde{P}^L \mid X = x, Z = 1\right] \\ &\quad + 1[\omega(x) < 0] \cdot \mathrm{P}\left[Z = 0 \mid X = x\right] \cdot \mathrm{E}\left[\tilde{P}^L \mid X = x, Z = 0\right]\Big). \end{aligned} \tag{A.7}$$

Next, if $Z_{\mathrm{G}} = Z$, we get $\tilde{P}^L = \mathrm{L}\left[D \mid Z, X\right] - \mathrm{L}\left[D \mid X\right]$. If we write $\mathrm{L}\left[D \mid Z, X\right] = Z\delta + X\zeta$, then $\mathrm{L}\left[D \mid X\right] = \mathrm{L}\left[Z \mid X\right]\delta + X\zeta$, which implies that, under Assumption PS, $\tilde{P}^L = (Z - \mathrm{L}\left[Z \mid X\right])\delta = (Z - \mathrm{E}\left[Z \mid X\right])\delta$. It follows that $\mathrm{E}\left[\tilde{P}^L \mid X, Z = 1\right] = (1 - \mathrm{E}\left[Z \mid X\right])\delta = \mathrm{P}\left[Z = 0 \mid X\right] \cdot \delta$ and $\mathrm{E}\left[\tilde{P}^L \mid X, Z = 0\right] = (0 - \mathrm{E}\left[Z \mid X\right])\delta = -\mathrm{P}\left[Z = 1 \mid X\right] \cdot \delta$, and further that

$$\begin{aligned} \theta_1(x) &= \pi(x) \cdot \Big(1[\omega(x) > 0] \cdot \mathrm{P}\left[Z = 1 \mid X = x\right] \cdot \mathrm{P}\left[Z = 0 \mid X = x\right] \cdot \delta \\ &\quad - 1[\omega(x) < 0] \cdot \mathrm{P}\left[Z = 0 \mid X = x\right] \cdot \mathrm{P}\left[Z = 1 \mid X = x\right] \cdot \delta\Big) \\ &= \pi(x) \cdot c(x) \cdot \mathrm{Var}\left[Z \mid X = x\right] \cdot \delta, \end{aligned} \tag{A.8}$$

which finally implies that

$$\beta_{\mathrm{IV}} = \frac{\mathrm{E}\left[c(X) \cdot \pi(X) \cdot \mathrm{Var}\left[Z \mid X\right] \cdot \delta \cdot \tau(X)\right]}{\mathrm{E}\left[c(X) \cdot \pi(X) \cdot \mathrm{Var}\left[Z \mid X\right] \cdot \delta\right]} = \frac{\mathrm{E}\left[c(X) \cdot \pi(X) \cdot \mathrm{Var}\left[Z \mid X\right] \cdot \tau(X)\right]}{\mathrm{E}\left[c(X) \cdot \pi(X) \cdot \mathrm{Var}\left[Z \mid X\right]\right]}. \tag{A.9}$$

This completes the proof.

Proof of Corollary 3.4. Recall that Assumption SM is a special case of Assumption WM where the existence of compliers but no defiers is postulated at all covariate values and the existence of defiers but no compliers everywhere else (*i.e.* on an empty set). Thus, it follows from Theorem 3.3 that, under Assumptions IV, SM, and PS, $\beta_{\mathrm{IV}} = \frac{\mathrm{E}[c(X) \cdot \pi(X) \cdot \mathrm{Var}[Z|X] \cdot \tau(X)]}{\mathrm{E}[c(X) \cdot \pi(X) \cdot \mathrm{Var}[Z|X]]}$ and $c(X) = 1$ a.s.

Reordered IV. Remark 3.6 suggests using $Z_{\mathrm{R}} = 1[\omega(X) > 0] \cdot Z + 1[\omega(X) < 0] \cdot (1 - Z)$ as a new, "reordered" instrument in a noninteracted specification. This instrument is binary and takes the value 1 if either $Z = 1$ and $\omega(X) > 0$ or $Z = 0$ and $\omega(X) < 0$; it also takes the value 0 if either $Z = 0$ and $\omega(X) > 0$ or $Z = 1$ and $\omega(X) < 0$. It follows that Z_{R} takes the value 1 for this value of the original instrument that encourages treatment conditional on X and the value 0 otherwise. When we construct the linear IV estimand using Z_{R} rather than Z, we obtain

$$\beta_{\mathrm{RIV}} = \left[\left(\mathrm{E}\left[Q_{\mathrm{R}}' W\right]\right)^{-1} \mathrm{E}\left[Q_{\mathrm{R}}' Y\right]\right]_1, \tag{A.10}$$

where $Q_{\mathrm{R}} = (Z_{\mathrm{R}}, X)$ and, as before, $W = (D, X)$. Formally, we establish the following result.

Corollary A.2 (Reordered IV)**.** *Suppose that Assumptions IV and WM hold. Suppose further that* $\mathrm{E}\left[Z_{\mathrm{R}} \mid X\right] = X\alpha_{\mathrm{R}}$. *Then*

$$\beta_{\mathrm{RIV}} = \frac{\mathrm{E}\left[\pi(X) \cdot \mathrm{Var}\left[Z \mid X\right] \cdot \tau(X)\right]}{\mathrm{E}\left[\pi(X) \cdot \mathrm{Var}\left[Z \mid X\right]\right]}.$$

Proof. The assumption that the conditional mean of the instrument is linear in X underlies the proof of Theorem 3.3, including equation (A.3). Under this assumption, we can use equation (A.3) to write

$$\beta_{\mathrm{RIV}} = \frac{\mathrm{E}\left[\mathrm{Var}\left[Z_{\mathrm{R}} \mid X\right] \cdot \omega_{\mathrm{R}}(X) \cdot \beta_{\mathrm{R}}(X)\right]}{\mathrm{E}\left[\mathrm{Var}\left[Z_{\mathrm{R}} \mid X\right] \cdot \omega_{\mathrm{R}}(X)\right]}, \tag{A.11}$$

where

$$\omega_{\mathrm{R}}(x) = \mathrm{E}\left[D \mid Z_{\mathrm{R}} = 1, X = x\right] - \mathrm{E}\left[D \mid Z_{\mathrm{R}} = 0, X = x\right] \tag{A.12}$$

and

$$\beta_{\mathrm{R}}(x) = \frac{\phi_{\mathrm{R}}(x)}{\omega_{\mathrm{R}}(x)}, \tag{A.13}$$

where

$$\phi_{\mathrm{R}}(x) = \mathrm{E}[Y \mid Z_{\mathrm{R}} = 1, X = x] - \mathrm{E}[Y \mid Z_{\mathrm{R}} = 0, X = x]. \tag{A.14}$$

Then, it is important to see that $\omega_{\mathrm{R}}(x) = \omega(x)$ and $\phi_{\mathrm{R}}(x) = \phi(x)$ if $\omega(x) > 0$, $\omega_{\mathrm{R}}(x) = -\omega(x)$ and $\phi_{\mathrm{R}}(x) = -\phi(x)$ if $\omega(x) < 0$, and consequently $\beta_{\mathrm{R}}(x) = \beta(x)$ regardless of the sign of $\omega(x)$. We can also write $\omega_{\mathrm{R}}(x) = c(x) \cdot \omega(x)$, $\phi_{\mathrm{R}}(x) = c(x) \cdot \phi(x)$, and $\mathrm{Var}[Z_{\mathrm{R}} \mid X = x] = \mathrm{Var}[Z \mid X = x]$ regardless of the sign of $\omega(x)$. It follows that

$$\beta_{\mathrm{RIV}} = \frac{\mathrm{E}[\mathrm{Var}[Z \mid X] \cdot c(X) \cdot \omega(X) \cdot \beta(X)]}{\mathrm{E}[\mathrm{Var}[Z \mid X] \cdot c(X) \cdot \omega(X)]}. \tag{A.15}$$

To complete this proof, note that, under Assumptions IV and WM, we know from Lemma 2.1 that $\beta(x) = \tau(x)$ and $\omega(x) = c(x) \cdot \pi(x)$. Also, $[c(x)]^2 = 1$ because $c(x) \in \{-1, 1\}$. Thus, it follows that

$$\begin{aligned} \beta_{\mathrm{RIV}} &= \frac{\mathrm{E}\left[\mathrm{Var}[Z \mid X] \cdot [c(X)]^2 \cdot \pi(X) \cdot \tau(X)\right]}{\mathrm{E}\left[\mathrm{Var}[Z \mid X] \cdot [c(X)]^2 \cdot \pi(X)\right]} \\ &= \frac{\mathrm{E}[\mathrm{Var}[Z \mid X] \cdot \pi(X) \cdot \tau(X)]}{\mathrm{E}[\mathrm{Var}[Z \mid X] \cdot \pi(X)]}. \end{aligned} \tag{A.16}$$

This completes the proof.

Appendix B Simulations

Table B.1: Simulation Results for $K = 250$, “Strong” IV, and No Monotonicity Violations

	N = 3,000			N = 10,000			N = 50,000		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–0.670	–0.669	3.113	–0.670	–0.669	12.66	–0.669	–0.669	163.1
IV	0.005	0.000	0.121	–0.001	–0.001	0.128	0.000	0.000	0.325
2SLS	–0.371	–0.372	1.000	–0.179	–0.180	1.000	–0.044	–0.044	1.000
MB2SLS	0.261	0.228	1.140	0.025	0.023	0.218	0.004	0.004	0.368
JIVE	2.900	0.700	5.1e+04	0.393	0.383	5.073	0.055	0.055	1.546
IJIVE	–0.049	–0.054	0.252	–0.006	–0.010	0.182	0.000	0.000	0.357
UJIVE	0.024	0.013	0.327	0.001	–0.002	0.186	0.000	0.000	0.357
FEJIV	0.028	0.010	0.419	0.001	0.000	0.188	0.000	0.000	0.357

B. Pretest for Weak Identification	N = 3,000	N = 10,000	N = 50,000
Average $\widetilde{F}$	11.30	33.99	173.25
$q_{0.05}$	8.04	28.50	161.83
$q_{0.95}$	14.82	39.74	184.88

Notes: The underlying data-generating process is described in Section 3.3.3. “OLS” is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. “IV” is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.2: Simulation Results for K = 250, “Strong” IV, and Moderate Monotonicity Violations

	N = 3,000			N = 10,000			N = 50,000		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.051	–1.050	3.204	–1.051	–1.051	11.67	–1.051	–1.051	151.0
IV	0.190	0.180	0.303	0.172	0.173	0.499	0.173	0.173	4.551
2SLS	–0.575	–0.578	1.000	–0.297	–0.299	1.000	–0.074	–0.074	1.000
MB2SLS	0.183	0.143	0.439	–0.017	–0.020	0.148	–0.004	–0.005	0.305
JIVE	68.98	1.944	1.5e+07	0.561	0.552	3.882	0.081	0.080	1.273
IJIVE	–0.048	–0.061	0.226	–0.005	–0.006	0.154	0.000	–0.001	0.306
UJIVE	0.062	0.047	0.315	0.006	0.005	0.158	0.001	0.000	0.307
FEJIV	0.099	0.077	0.363	0.013	0.010	0.157	0.001	0.000	0.307
B. Pretest for Weak Identification									
Average $\widetilde{F}$		11.48			33.45			162.81	
$q_{0.05}$		8.20			28.52			152.88	
$q_{0.95}$		15.08			38.69			172.65	

Notes: The underlying data-generating process is described in Section 3.3.3. “OLS” is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. “IV” is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.3: Simulation Results for $K = 250$, "Strong" IV, and Large Monotonicity Violations

	N = 3,000			N = 10,000			N = 50,000		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.058	–1.059	3.132	–1.060	–1.060	11.29	–1.059	–1.059	147.4
IV	0.729	0.546	4.686	0.564	0.551	4.983	0.545	0.541	43.22
2SLS	–0.586	–0.588	1.000	–0.305	–0.307	1.000	–0.077	–0.078	1.000
MB2SLS	0.038	0.024	0.251	–0.047	–0.050	0.151	–0.011	–0.013	0.271
JIVE	–11.89	2.337	1.4e+06	0.527	0.515	3.309	0.077	0.075	1.091
IJIVE	–0.058	–0.073	0.234	–0.005	–0.007	0.148	0.000	–0.002	0.262
UJIVE	0.056	0.030	0.327	0.007	0.006	0.153	0.000	–0.001	0.262
FEJIV	0.099	0.074	0.357	0.014	0.012	0.153	0.000	–0.001	0.262
B. Pretest for Weak Identification									
Average $\widetilde{F}$		11.61			32.16			155.20	
$q_{0.05}$		8.07			27.72			146.33	
$q_{0.95}$		15.66			36.54			165.50	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.4: Simulation Results for $K = 250$, "Strong" IV, and Monotonicity Violations with Weak Cells

	$N = 3{,}000$			$N = 10{,}000$			$N = 50{,}000$		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.127	–1.128	2.354	–1.125	–1.126	6.550	–1.125	–1.125	72.16
IV	0.811	0.629	24.01	0.678	0.600	4.107	0.641	0.633	26.21
2SLS	–0.721	–0.726	1.000	–0.427	–0.429	1.000	–0.120	–0.119	1.000
MB2SLS	0.146	0.079	0.600	–0.072	–0.076	0.167	–0.018	–0.018	0.241
JIVE	–3.500	–3.108	257.9	1.185	1.076	9.265	0.128	0.128	1.235
IJIVE	–0.058	–0.108	0.429	–0.011	–0.020	0.174	–0.001	–0.002	0.231
UJIVE	0.164	0.058	0.996	0.008	–0.002	0.183	–0.001	–0.001	0.232
FEJIV	0.220	0.116	1.757	0.019	0.008	0.180	–0.001	–0.001	0.232
B. Pretest for Weak Identification									
Average $\widetilde{F}$		7.64			20.47			99.71	
$q_{0.05}$		5.09			16.36			92.88	
$q_{0.95}$		10.82			24.61			107.47	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.5: Simulation Results for K = 20, "Weak" IV, and No Monotonicity Violations

	N = 3,000			N = 10,000		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–0.737	–0.736	1.875	–0.737	–0.736	5.975
IV	0.070	–0.004	1.368	0.024	0.010	0.902
2SLS	–0.426	–0.419	1.000	–0.199	–0.204	1.000
MB2SLS	0.696	–0.009	193.4	0.053	0.021	1.569
JIVE	–1.586	–1.084	1.2e+03	1.092	0.559	3.0e+03
IJIVE	0.384	–0.055	440.5	0.058	0.023	1.671
UJIVE	–1.492	–0.039	8.0e+04	0.065	0.026	1.755
FEJIV	–1.280	–0.012	2.1e+03	0.063	0.027	1.700
B. Pretest for Weak Identification						
Average $\widetilde{F}$		2.80			7.84	
$q_{0.05}$		0.28			4.14	
$q_{0.95}$		6.15			12.49	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.6: Simulation Results for $K = 20$, "Weak" IV, and Moderate Monotonicity Violations

	$N = 3{,}000$			$N = 10{,}000$		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.134	–1.132	1.949	–1.133	–1.132	6.075
IV	2.716	0.186	8.0e+03	0.277	0.199	2.317
2SLS	–0.663	–0.652	1.000	–0.312	–0.318	1.000
MB2SLS	0.295	–0.095	386.4	0.051	–0.011	1.540
JIVE	–2.849	–1.689	1.6e+03	–3.561	0.768	6.6e+04
IJIVE	1.833	–0.072	3.4e+03	0.092	0.031	1.801
UJIVE	2.511	–0.061	1.2e+04	0.102	0.035	1.899
FEJIV	1.033	–0.027	1.1e+03	0.099	0.033	1.769
B. Pretest for Weak Identification						
Average $\widetilde{F}$		2.62			7.75	
$q_{0.05}$		0.24			3.90	
$q_{0.95}$		5.50			12.03	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.7: Simulation Results for $K = 20$, "Weak" IV, and Large Monotonicity Violations

	N = 3,000			N = 10,000		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.138	–1.137	1.951	–1.138	–1.137	6.109
IV	0.602	0.074	506.6	0.802	0.331	67.86
2SLS	–0.663	–0.649	1.000	–0.324	–0.332	1.000
MB2SLS	1.209	–0.122	1.1e+03	0.025	–0.031	1.331
JIVE	–0.965	–1.781	8.5e+03	0.440	0.752	3.1e+03
IJIVE	–0.203	–0.099	183.9	0.084	0.003	1.653
UJIVE	0.093	–0.090	264.9	0.094	0.011	1.736
FEJIV	0.369	–0.003	239.8	0.090	0.014	1.645
B. Pretest for Weak Identification						
Average $\widetilde{F}$		2.54			7.48	
$q_{0.05}$		0.02			3.80	
$q_{0.95}$		5.97			11.62	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.8: Simulation Results for $K = 20$, "Weak" IV, and Monotonicity Violations with Weak Cells

	$N = 3{,}000$			$N = 10{,}000$		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.152	–1.151	1.633	–1.149	–1.149	4.847
IV	1.409	0.139	1.3e+03	–0.872	0.323	7.7e+03
2SLS	–0.740	–0.749	1.000	–0.371	–0.371	1.000
MB2SLS	–0.276	–0.290	104.8	0.059	–0.005	1.658
JIVE	–0.889	–1.696	1.8e+03	0.803	0.969	7.9e+03
IJIVE	–0.672	–0.202	892.2	0.131	0.053	2.075
UJIVE	–0.751	–0.212	377.9	0.148	0.067	2.243
FEJIV	0.265	–0.122	351.7	0.138	0.064	2.130
B. Pretest for Weak Identification						
Average $\widetilde{F}$		2.18			6.09	
$q_{0.05}$		–0.02			2.32	
$q_{0.95}$		4.66			9.85	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.9: Simulation Results for $K = 20$, "Strong" IV, and No Monotonicity Violations

	$N = 3{,}000$			$N = 10{,}000$		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–0.669	–0.669	29.70	–0.669	–0.668	105.8
IV	0.003	–0.001	0.949	0.002	0.002	0.983
2SLS	–0.049	–0.052	1.000	–0.014	–0.015	1.000
MB2SLS	0.009	0.005	1.053	0.004	0.003	1.021
JIVE	0.072	0.063	1.679	0.023	0.022	1.226
IJIVE	0.004	0.001	1.027	0.003	0.002	1.020
UJIVE	0.007	0.002	1.047	0.004	0.002	1.023
FEJIV	0.007	0.001	1.033	0.004	0.002	1.020
B. Pretest for Weak Identification						
Average $\widetilde{F}$		42.75			131.58	
$q_{0.05}$		31.36			113.22	
$q_{0.95}$		55.10			154.91	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.10: Simulation Results for $K = 20$, "Strong" IV, and Moderate Monotonicity Violations

	N = 3,000			N = 10,000		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.027	–1.027	32.72	–1.025	–1.025	114.2
IV	0.213	0.196	3.545	0.206	0.200	6.702
2SLS	–0.084	–0.089	1.000	–0.021	–0.022	1.000
MB2SLS	–0.003	–0.008	0.958	0.004	0.003	1.012
JIVE	0.099	0.094	1.562	0.036	0.033	1.237
IJIVE	0.002	0.002	0.980	0.007	0.005	1.025
UJIVE	0.006	0.005	0.999	0.007	0.006	1.028
FEJIV	0.006	0.003	0.986	0.007	0.007	1.025
B. Pretest for Weak Identification						
Average $\widetilde{F}$		40.58			124.89	
$q_{0.05}$		31.14			107.54	
$q_{0.95}$		51.84			140.35	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.11: Simulation Results for $K = 20$, "Strong" IV, and Large Monotonicity Violations

	N = 3,000			N = 10,000		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.032	–1.033	31.52	–1.032	–1.030	117.0
IV	0.397	0.333	13.87	0.357	0.339	22.19
2SLS	–0.087	–0.090	1.000	–0.026	–0.026	1.000
MB2SLS	–0.008	–0.012	0.950	–0.001	–0.002	0.987
JIVE	0.096	0.086	1.532	0.031	0.030	1.182
IJIVE	0.002	0.001	0.986	0.004	0.003	1.002
UJIVE	0.006	0.004	1.007	0.004	0.003	1.005
FEJIV	0.007	0.002	0.989	0.004	0.003	1.002
B. Pretest for Weak Identification						
Average $\widetilde{F}$		39.01			119.92	
$q_{0.05}$		28.85			104.08	
$q_{0.95}$		49.38			136.68	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Table B.12: Simulation Results for $K = 20$, "Strong" IV, and Monotonicity Violations with Weak Cells

	$N = 3{,}000$			$N = 10{,}000$		
A. Estimator Performance	Bias	Median Bias	MSE	Bias	Median Bias	MSE
OLS	–1.062	–1.064	25.17	–1.061	–1.062	91.49
IV	0.374	0.325	7.865	0.340	0.333	14.39
2SLS	–0.101	–0.101	1.000	–0.032	–0.030	1.000
MB2SLS	–0.004	–0.011	0.987	–0.001	0.001	0.990
JIVE	0.123	0.101	1.732	0.038	0.042	1.206
IJIVE	0.006	–0.002	1.022	0.004	0.007	1.005
UJIVE	0.011	0.002	1.052	0.004	0.008	1.008
FEJIV	0.012	0.004	1.008	0.004	0.008	1.007
B. Pretest for Weak Identification						
Average $\widetilde{F}$		33.37			101.53	
$q_{0.05}$		24.50			85.67	
$q_{0.95}$		42.08			117.25	

Notes: The underlying data-generating process is described in Section 3.3.3. "OLS" is the OLS estimator in the regression of the outcome on the treatment indicator and group indicators. "IV" is the IV estimator in the noninteracted specification. The remaining estimators are based on the interacted specification and are described in Section 3.3. JIVE, IJIVE, and UJIVE are computed after dropping all groups with fewer than two observations in either (X, Z) combination. FEJIV is computed after dropping all groups with fewer than three observations in either (X, Z) combination. The pretest for weak identification follows Mikusheva and Sun (2022); see also the Stata implementation in Sun (2023). Bias and median bias are reported as the proportion of the target parameter. MSE is normalized by the MSE of 2SLS. Results are based on 1,000 replications. Pretest results are based on 250 replications.

Appendix C Review of Applications of Instrumental Variables

Table C.1: Negative First Stages with Alternative Binarizations

		Z_1		Z_2		Z_3	
		LPM	Probit	LPM	Probit	LPM	Probit
D_1	Average Share	0.293	0.250	0.246	0.157	0.281	0.263
	Weighted Average Share	0.321	0.304	0.277	0.253	0.315	0.327
D_2	Average Share	0.272	0.220	0.218	0.176	0.199	0.187
	Weighted Average Share	0.327	0.314	0.285	0.280	0.263	0.259
D_3	Average Share	0.379	0.341	0.276	0.190	0.243	0.223
	Weighted Average Share	0.379	0.382	0.316	0.283	0.293	0.307

Notes: The table reports summary statistics on the fraction of observations for which $\hat{\mathrm{E}}\left[D_j \mid Z_k = 1, X = x\right] - \hat{\mathrm{E}}\left[D_j \mid Z_k = 0, X = x\right]$ is negative. "Average Share" treats every applicable regression equally. "Weighted Average Share" weights by the inverse of the number of applicable regressions associated with a given paper. D_j and Z_k are defined as either the original endogenous explanatory variable and instrument (if they are binary) or indicators for whether these variables are above the jth and kth quartile, respectively, subject to a normalization that Z_k is always associated with a positive estimated coefficient in the linear first stage. Sampling weights and clustered standard errors are used in line with the original papers.

Table C.2: First-Stage Heterogeneity with Alternative Binarizations

		Z_1		Z_2		Z_3	
		LPM	Probit	LPM	Probit	LPM	Probit
D_1	Rejected Papers	25/25	18/20	22/25	16/18	25/25	15/19
	Average Share of Rejections	0.807	0.749	0.781	0.765	0.855	0.752
D_2	Rejected Papers	24/25	19/22	22/25	19/21	24/25	18/21
	Average Share of Rejections	0.789	0.756	0.715	0.749	0.789	0.768
D_3	Rejected Papers	24/25	18/20	24/25	18/21	24/25	19/20
	Average Share of Rejections	0.824	0.733	0.779	0.744	0.817	0.747

Notes: The table reports results of Wald tests that the coefficients on the interaction terms in regressions of D_j on Z_k, X, and Z_kX are jointly equal to zero. "Rejected Papers" reports the number of papers for which the Bonferroni p-value is less than or equal to 0.05. "Average Share of Rejections" reports the average share (across papers) of regressions associated with a given paper for which the corresponding Holm p-value is less than or equal to 0.05. D_j and Z_k are defined as either the original endogenous explanatory variable and instrument (if they are binary) or indicators for whether these variables are above the jth and kth quartile, respectively, subject to a normalization that Z_k is always associated with a positive estimated coefficient in the linear first stage. Sampling weights and clustered standard errors are used in line with the original papers.

Table C.3: Detailed Results on Negative First Stages and First-Stage Heterogeneity

	Average Share of Negative First Stages		First-Stage Heterogeneity			
			Bonferroni p-Value		Share of Rejections	
	LPM	Probit	LPM	Probit	LPM	Probit
Acemoglu *et al.* (2008)	0.546	0.554	0	0	1	1
Albouy (2012)	0.147	0.122	0.115	0	0	0.226
Alesina and Zhuravskaya (2011)	0.145	0.159	0.001	0	0.064	0.011
Ananat (2011)	0.020	0.223	0.308	1	0	0
Autor *et al.* (2013)	0.009	0.036	0	0	0.894	0.894
Bazzi and Clemens (2013)	0.303	0.242	0	1	0.750	0
Becker *et al.* (2011)	0.068	0.100	0	0	0.125	0.450
Bleakley and Chin (2010)	0.206	0.029	0	0	1	1
Brown and Laschever (2012)	0.134	0.143	0	0	0.375	0.375
Chalfin (2015)	0.342	N/A	0	0	1	1
Chodorow-Reich *et al.* (2012)	0.123	0.105	1	N/A	0	N/A
Chou *et al.* (2010)	0.514	0.503	0	0	1	1
Collins and Shester (2013)	0.343	0.341	0	0	1	1
Decarolis (2015)	0.595	0.568	0	0	0.667	1
Dinkelman (2011)	0.306	0.271	0	0	0.158	1
Draca *et al.* (2011)	0.250	0.269	0	0	1	1
Guryan and Kearney (2010)	0.427	N/A	0	N/A	1	N/A
Hornung (2014)	0.116	0.151	0	0	0.909	0.889
Hunt and Gauthier-Loiselle (2010)	0.245	0.335	0	0	1	1
James (2015)	0.555	0.616	0	0	1	1
Kraay (2014)	0.435	0.439	0	0	1	1
Lipscomb *et al.* (2013)	0	0	0.018	0.002	1	1
Moser *et al.* (2014)	0.671	0.626	0	N/A	1	N/A
Oreopoulos (2006)	0.463	0.434	0	0	0.941	0.875
Saiz and Wachter (2011)	0.152	0.165	0	N/A	1	N/A
Number of Regressions	988	930	988	899	988	899

Notes: "Average Share of Negative First Stages" reports the average fraction of observations for which $\hat{E}[D \mid Z = 1, X = x] - \hat{E}[D \mid Z = 0, X = x]$ is negative. "First-Stage Heterogeneity" reports results of Wald tests that the coefficients on the interaction terms in regressions of D on Z, X, and ZX are jointly equal to zero. "Bonferroni p-Value" reports the product of the smallest p-value associated with a given paper and the number of applicable regressions in that paper. "Share of Rejections" reports the fraction of applicable regressions associated with a given paper for which the corresponding Holm p-value is less than or equal to 0.05. D and Z are defined as either the original endogenous explanatory variable and instrument (if they are binary) or indicators for whether these variables are above their medians, subject to a normalization that Z is always associated with a positive estimated coefficient in the linear first stage. Sampling weights and clustered standard errors are used in line with the original papers.

Appendix D Reanalysis of Stevenson (2018)

Table D.1: Alternative Estimates of the Effects of Pretrial Detention on Conviction and Incarceration Length

	Specification #1		Specification #2		Specification #3	
A. Effects on Conviction	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$
MB2SLS	0.1610***	0.0360	0.1860***	0.0326	0.0751***	0.0271
JIVE	0.2127	0.1969	0.3762	0.2783	0.3251	0.3863
B. Effects on Incarceration Length	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$	$\hat{\beta}$	$\hat{\sigma}_{\hat{\beta}}$
MB2SLS	134***	47	140***	41	55	43
JIVE	–313	280	–225	359	1,279*	732
Number of Groups	431		563		981	
Number of Observations	327,560		325,915		319,573	

Notes: The data are Stevenson (2018)'s sample of 331,971 arrests in Philadelphia. The outcomes are conviction (Panel A) or incarceration length (Panel B), defined as the maximum days of an incarceration sentence. The treatment is pretrial detention. The instrument is whether a given case was heard by Judge C. Each specification is based on a division of the sample into a number of mutually exclusive groups, with a separate group for each combination of values of selected variables. Specification #1 uses the offense type and race (Black, White, or other) of the defendant. Specification #2 uses the offense type, race, and gender (male or female) of the defendant. Specification #3 uses the offense type, race and gender of the defendant, and three time periods considered by Stevenson (2018). Groups with fewer than three observations in either (G, Z) combination are dropped. MB2SLS and JIVE are based on the interacted specification and are described in Section 3.3.
*Statistically different from zero at the 10% level; **at the 5% level; ***at the 1% level.

Table D.2: Bootstrap p-Values for the Comparison of the Noninteracted and Interacted Specifications

	Specification #1	Specification #2	Specification #3
A. Effects on Conviction			
2SLS	0.496	0.486	0.301
MB2SLS	0.829	0.956	0.377
JIVE	0.818	0.119	0.195
UJIVE	0.619	0.671	0.229
B. Effects on Incarceration Length			
2SLS	0.051	0.019	0.030
MB2SLS	0.049	0.021	0.017
JIVE	0.001	0.001	0.031
UJIVE	0.032	0.016	0.016
Number of Groups	431	563	981
Number of Observations	327,560	325,915	319,573

Notes: This table revisits the estimates in Tables 8 and D.1. The data are Stevenson (2018)'s sample of 331,971 arrests in Philadelphia. The outcomes are conviction (Panel A) or incarceration length (Panel B), defined as the maximum days of an incarceration sentence. The treatment is pretrial detention. The instrument is whether a given case was heard by Judge C. Each specification is based on a division of the sample into a number of mutually exclusive groups, with a separate group for each combination of values of selected variables. Specification #1 uses the offense type and race (Black, White, or other) of the defendant. Specification #2 uses the offense type, race, and gender (male or female) of the defendant. Specification #3 uses the offense type, race and gender of the defendant, and three time periods considered by Stevenson (2018). Groups with fewer than three observations in either (G, Z) combination are dropped. Each p-value is calculated using a bootstrap test of the equality of the estimands in the noninteracted and interacted specifications (with 250 bootstrap replications). The noninteracted specification is estimated using IV. The estimators of the interacted specification, listed in the first column, are described in Section 3.3.

Table D.3: Bootstrap p-Values for the Comparison of 2SLS and Other Estimators of the Interacted Specification

	Specification #1	Specification #2	Specification #3
A. Effects on Conviction			
MB2SLS	0.008	0.000	0.450
JIVE	0.018	0.000	0.000
UJIVE	0.188	0.057	0.176
B. Effects on Incarceration Length			
MB2SLS	0.404	0.490	0.030
JIVE	0.000	0.000	0.000
UJIVE	0.010	0.108	0.006
Number of Groups	431	563	981
Number of Observations	327,560	325,915	319,573

Notes: This table revisits the estimates in Tables 8 and D.1. The data are Stevenson (2018)'s sample of 331,971 arrests in Philadelphia. The outcomes are conviction (Panel A) or incarceration length (Panel B), defined as the maximum days of an incarceration sentence. The treatment is pretrial detention. The instrument is whether a given case was heard by Judge C. Each specification is based on a division of the sample into a number of mutually exclusive groups, with a separate group for each combination of values of selected variables. Specification #1 uses the offense type and race (Black, White, or other) of the defendant. Specification #2 uses the offense type, race, and gender (male or female) of the defendant. Specification #3 uses the offense type, race and gender of the defendant, and three time periods considered by Stevenson (2018). Groups with fewer than three observations in either (G, Z) combination are dropped. Each p-value is calculated using a bootstrap test of the equality of the probability limits of 2SLS and other estimators in the interacted specifications (with 250 bootstrap replications). Other estimators of the interacted specification, listed in the first column, are described in Section 3.3.

Appendix References

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared (2008): "Income and Democracy," *American Economic Review*, 98, 808–842.

Albouy, David Y. (2012): "The Colonial Origins of Comparative Development: An Empirical Investigation: Comment," *American Economic Review*, 102, 3059–3076.

Alesina, Alberto and Ekaterina Zhuravskaya (2011): "Segregation and the Quality of Government in a Cross Section of Countries," *American Economic Review*, 101, 1872–1911.

Ananat, Elizabeth Oltmans (2011): "The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality," *American Economic Journal: Applied Economics*, 3, 34–66.

Angrist, Joshua D. (1998): "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants," *Econometrica*, 66, 249–288.

Autor, David H., David Dorn, and Gordon H. Hanson (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–2168.

Bazzi, Samuel and Michael A. Clemens (2013): "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth," *American Economic Journal: Macroeconomics*, 5, 152–186.

Becker, Sascha O., Erik Hornung, and Ludger Woessmann (2011): "Education and Catch-up in the Industrial Revolution," *American Economic Journal: Macroeconomics*, 3, 92–126.

Bleakley, Hoyt and Aimee Chin (2010): "Age at Arrival, English Proficiency, and Social Assimilation Among US Immigrants," *American Economic Journal: Applied Economics*, 2, 165–192.

Brown, Kristine M. and Ron A. Laschever (2012): "When They're Sixty-Four: Peer Effects and the Timing of Retirement," *American Economic Journal: Applied Economics*, 4, 90–115.

Chalfin, Aaron (2015): "The Long-Run Effect of Mexican Immigration on Crime in US Cities: Evidence from Variation in Mexican Fertility Rates," *American Economic Review: Papers & Proceedings*, 105, 220–225.

Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Gui Woolston (2012): "Does State Fiscal Relief During Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act," *American Economic Journal: Economic Policy*, 4, 118–145.

Chou, Shin-Yi, Jin-Tan Liu, Michael Grossman, and Ted Joyce (2010): "Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan," *American Economic Journal: Applied Economics*, 2, 33–61.

Collins, William J. and Katharine L. Shester (2013): "Slum Clearance and Urban Renewal in the United States," *American Economic Journal: Applied Economics*, 5, 239–273.

Decarolis, Francesco (2015): "Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design?" *American Economic Review*, 105, 1547–1580.

Dinkelman, Taryn (2011): "The Effects of Rural Electrification on Employment: New Evidence from South Africa," *American Economic Review*, 101, 3078–3108.

Draca, Mirko, Stephen Machin, and Robert Witt (2011): "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks," *American Economic Review*, 101, 2157–2181.

Guryan, Jonathan and Melissa S. Kearney (2010): "Is Lottery Gambling Addictive?" *American Economic Journal: Economic Policy*, 2, 90–110.

Hornung, Erik (2014): "Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia," *American Economic Review*, 104, 84–122.

Hunt, Jennifer and Marjolaine Gauthier-Loiselle (2010): "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics*, 2, 31–56.

James, Alexander (2015): "US State Fiscal Policy and Natural Resources," *American Economic Journal: Economic Policy*, 7, 238–257.

Kolesár, Michal (2013): "Estimation in an Instrumental Variables Model with Treatment Effect Heterogeneity." Unpublished.

Kraay, Aart (2014): "Government Spending Multipliers in Developing Countries: Evidence from Lending by Official Creditors," *American Economic Journal: Macroeconomics*, 6, 170–208.

Lipscomb, Molly, A. Mushfiq Mobarak, and Tania Barham (2013): "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil," *American Economic Journal: Applied Economics*, 5, 200–231.

Mikusheva, Anna and Liyang Sun (2022): "Inference with Many Weak Instruments," *Review of Economic Studies*, 89, 2663–2686.

Moser, Petra, Alessandra Voena, and Fabian Waldinger (2014): "German Jewish Émigrés and US Invention," *American Economic Review*, 104, 3222–3255.

Oreopoulos, Philip (2006): "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter," *American Economic Review*, 96, 152–175.

Saiz, Albert and Susan Wachter (2011): "Immigration and the Neighborhood," *American Economic Journal: Economic Policy*, 3, 169–188.

Stevenson, Megan T. (2018): "Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes," *Journal of Law, Economics, and Organization*, 34, 511–542.

Sun, Liyang (2023): "manyweakiv: Weak-Instrument Robust Test for Linear IV Regressions with Many Instruments," https://github.com/lsun20/manyweakiv.

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