On-Line Appendix for
Nearly Collinear Robust Procedures for 2SLS Estimation

Alwyn Young, August 2023

A Maximum R^2 vs Condition Numbers

In a footnote in the paper I note that one minus the maximum partial R^2 (net of any fixed effects) in the regression of the instruments on each other explains as much of the variation in the computational sensitivity of 2SLS results in Stata as the standard and Skeel condition numbers of the matrix of demeaned instruments, but when included alongside these measures in the regression it is less frequently statistically significant. Tables A1 and A2 below substantiate these claims. I use the condition number of the matrix of demeaned instruments, because as the results in the paper suggest demeaning is used in Stata's 2SLS commands. Gould (2018) recommends improving the accuracy of matrix calculations by demeaning and rescaling by diagonal values. Since the algorithms Stata uses to invert matrices in 2SLS estimation are not visible to users, I examine the condition numbers for the matrix of demeaned instruments and for the same matrix rescaled by the diagonal values. With $\widetilde{\mathbf{Z}}_{\sim c}$ denoting the matrix of demeaned values of instruments (net of the constant term and absorbed fixed effects) and $\mathbf{D}_{\mathbf{A}}^{-1/2}$ a diagonal matrix made up of the square root of the inverses of the diagonal elements of $\mathbf{A}$, I consider the two matrices

$$\text{(A1a)}\ \widetilde{\mathbf{Z}}'_{\sim c}\widetilde{\mathbf{Z}}_{\sim c},\ \&\ \text{(A1b)}\ \mathbf{D}^{-1/2}_{\widetilde{\mathbf{Z}}'_{\sim c}\widetilde{\mathbf{Z}}_{\sim c}}(\widetilde{\mathbf{Z}}'_{\sim c}\widetilde{\mathbf{Z}}_{\sim c})\mathbf{D}^{-1/2}_{\widetilde{\mathbf{Z}}'_{\sim c}\widetilde{\mathbf{Z}}_{\sim c}},$$

whose condition numbers (the ratios of the largest to the smallest eigenvalues) are referred to as CN_a and CN_b below. I also examine the Skeel condition number, which for a matrix $\mathbf{A}$ is the maximum rowsum of $|\mathbf{A}^{-1}||\mathbf{A}|$ and is denoted by SCN_a or SCN_b below. As proven in the paper, CN_a, CN_b , SCN_a, and SCN_b are all bounded from below by $(1\text{-}R^{2max})^{-1}$.

In Table A1 the dependent variables are the $\log_{10}$ coefficient of variations of estimated 2SLS coefficients across 50 permutations of variable order for the 10 collinearity increasing rotations of each of the 837 2SLS specifications in 28 papers, as described in the paper. In panel (A) the coefficients of variation are for regressions run using Stata's built-in routines *ivregress* and *xtivreg* (for specifications that have large numbers of absorbed fixed effects), in panel (B) for the same regressions run using the user written commands *ivreg2* and *xtivreg2*, in panel (C) for my computations for the same regressions using method D, demeaned variables and matrix inverses in Mata, and in panel (D) for my computations using method D, demeaned variables and solvers of linear equation systems in Mata. Sub-panel (i) presents results in which other than the collinearity or conditioning measure the regression in the table only contains a constant term, (ii) in which the regression contains paper fixed effects, and (iii) in which it contains paper x regression fixed effects, so that all of the identification comes from variation induced by the 10 collinearity-increasing rotations of the included instruments of each sample regression. Results are presented separately for the coefficients of variation of the coefficients on instrumented variables and on the included instruments (including the constant term when there are no absorbed fixed effects). Standard errors (in parentheses) are clustered at the 28 paper level with adjustment for bias and p-values (in brackets) with effective degrees of freedom corrections, both as implemented by *edfreg*. These adjustments account for the bias and excess volatility of the standard error estimate brought about by uneven leverage and generally increase standard errors and raise p-values. As seen in the table, the R^2s attained with $\log_{10}(1 - R^{2Max})$ in these regressions are generally much greater than those found using $\log_{10}(CN_a)$, somewhat greater than those found using $\log_{10}(SCN_a)$, and on par with those found using $\log_{10}(CN_b)$ or $\log_{10}(SCN_b)$ (i.e. the condition numbers for the rescaled matrices of inner-products).

Table A1. Determinants of Log_{10} Coefficient of Variation by Collinearity Measure
(10 rotations each of 837 2SLS regression specifications in 28 papers)

	Coefficients on Instrumented Variable ($\hat{\boldsymbol{\beta}}_1$)					Coefficients on Included Instruments ($\hat{\boldsymbol{\beta}}_2$)				
	$1\text{-}R^{2Max}$	CN_a	CN_b	SCN_a	SCN_b	$1\text{-}R^{2Max}$	CN_a	CN_b	SCN_a	SCN_b
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(A) Dependent variable: coefficient of variation of *ivregress* & *xtivreg*										
(i) no fixed effects										
$\log_{10}$ collinearity	-1.87	1.02	1.63	1.55	1.59	-1.94	1.22	1.65	1.84	1.65
	(.050)	(.209)	(.070)	(.099)	(.076)	(.039)	(.299)	(.059)	(.118)	(.085)
	[.000]	[.002]	[.000]	[.000]	[.000]	[.000]	[.005]	[.000]	[.000]	[.000]
constant	-15.4	-15.5	-16.9	-17.5	-17.3	-15.0	-16.3	-17.1	-19.1	-17.9
	(.254)	(1.04)	(.401)	(.570)	(.455)	(.149)	(1.69)	(.351)	(.767)	(.524)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.9186	.5118	.9077	.8154	.8866	.9548	.5131	.9474	.8648	.9309
(ii) paper fixed effects										
$\log_{10}$ collinearity	-1.88	.969	1.67	1.60	1.67	-1.98	1.09	1.77	1.89	1.81
	(.031)	(.190)	(.065)	(.108)	(.072)	(.052)	(.230)	(.044)	(.079)	(.033)
	[.000]	[.004]	[.000]	[.000]	[.000]	[.000]	[.007]	[.000]	[.000]	[.000]
R^2	.9588	.6904	.9550	.9066	.9533	.9655	.7034	.9688	.9292	.9687
(iii) paper x regression fixed effects										
$\log_{10}$ collinearity	-2.00	2.07	1.82	2.02	1.83	-2.10	2.22	1.90	2.09	1.93
	(.033)	(.196)	(.030)	(.075)	(.030)	(.034)	(.201)	(.013)	(.076)	(.018)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.006]	[.000]	[.000]	[.000]
R^2	.9895	.8774	.9918	.9802	.9919	.9757	.8511	.9796	.9656	.9795
N (i)-(iii)	(1) - (5) = 8338					(6) - (10) = 382576				
(B) Dependent variable: coefficient of variation of *ivreg2* & *xtivreg2*										
(i) no fixed effects										
$\log_{10}$ collinearity	-.799	.393	.708	.652	.696	-.872	.503	.745	.830	.744
	(.069)	(.120)	(.064)	(.088)	(.062)	(.055)	(.159)	(.067)	(.073)	(.079)
	[.000]	[.013]	[.000]	[.000]	[.000]	[.000]	[.017]	[.000]	[.000]	[.000]
constant	-13.5	-13.3	-14.2	-14.3	-14.4	-13.3	-13.5	-14.2	-15.2	-14.6
	(.266)	(.629)	(.332)	(.480)	(.341)	(.204)	(.930)	(.366)	(.524)	(.462)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.6285	.2822	.6476	.5439	.6409	.7799	.3528	.7758	.7146	.7647
(ii) paper fixed effects										
$\log_{10}$ collinearity	-.796	.368	.719	.668	.720	-.961	.476	.866	.912	.888
	(.054)	(.101)	(.038)	(.089)	(.038)	(.083)	(.128)	(.068)	(.087)	(.059)
	[.000]	[.015]	[.000]	[.000]	[.000]	[.000]	[.017]	[.000]	[.000]	[.000]
R^2	.8534	.6357	.8626	.8096	.8622	.8695	.5793	.8799	.8323	.8829
(iii) paper x regression fixed effects										
$\log_{10}$ collinearity	-.952	.986	.868	.971	.876	-1.11	1.15	1.00	1.10	1.02
	(.029)	(.103)	(.029)	(.043)	(.030)	(.017)	(.113)	(.021)	(.046)	(.021)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.9724	.8775	.9753	.9671	.9754	.9201	.7727	.9252	.9090	.9251
N (i)-(iii)	(1) - (5) = 8369					(6) - (10) = 382628				

Notes: Reported numbers = coefficient estimate, standard error estimate (in parentheses) clustered at the 28 paper level and adjusted for bias, & p-value [in brackets] with effective degrees of freedom corrections (last two using Stata command *edfreg*). N = number of observations; some are dropped because the coefficient of variation is zero.

Table A1 - continued

	(A) Coefficients on Instrumented Variable ($\hat{\boldsymbol{\beta}}_1$)					(B) Coefficients on Included Instruments ($\hat{\boldsymbol{\beta}}_2$)				
	$1\text{-}R^{2Max}$	CN_a	CN_b	SCN_a	SCN_b	$1\text{-}R^{2Max}$	CN_a	CN_b	SCN_a	SCN_b
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	(C) Dependent variable: coefficient of variation of method D using demeaned variables & matrix inverses									
	(i) no fixed effects									
$\log_{10}$ collinearity	-.258	.122	.234	.218	.232	-.764	.464	.654	.739	.653
	(.034)	(.049)	(.027)	(.043)	(.028)	(.049)	(.112)	(.066)	(.053)	(.076)
	[.000]	[.039]	[.000]	[.000]	[.000]	[.000]	[.005]	[.000]	[.000]	[.000]
constant	-15.5	-15.4	-15.8	-15.8	-15.8	-15.2	-15.6	-16.0	-16.9	-16.4
	(.088)	(.235)	(.108)	(.185)	(.124)	(.118)	(.626)	(.267)	(.304)	(.350)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.4438	.1798	.4788	.4062	.4819	.8058	.4028	.8037	.7603	.7934
	(ii) paper fixed effects									
$\log_{10}$ collinearity	-.247	.093	.222	.198	.222	-.873	.452	.786	.835	.807
	(.056)	(.055)	(.045)	(.068)	(.045)	(.072)	(.105)	(.058)	(.074)	(.047)
	[.002]	[.148]	[.001]	[.020]	[.001]	[.000]	[.010]	[.000]	[.000]	[.000]
R^2	.5757	.3993	.5781	.5258	.5767	.8413	.5371	.8523	.8078	.8570
	(iii) paper x regression fixed effects									
$\log_{10}$ collinearity	-.357	.418	.326	.373	.329	-1.02	1.12	.927	1.03	.942
	(.029)	(.047)	(.025)	(.028)	(.025)	(.021)	(.091)	(.010)	(.034)	(.013)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.9271	.8704	.9301	.9328	.9307	.8851	.7426	.8922	.8847	.8934
N (i)-(iii)	(1) - (5) = 8305					(6) - (10) = 382391				
	(D) Dependent variable: coefficient of variation of method D using demeaned variables & linear solvers									
	(i) no fixed effects									
$\log_{10}$ collinearity	-.281	.143	.247	.238	.243	-.900	.564	.779	.874	.781
	(.021)	(.042)	(.022)	(.029)	(.024)	(.027)	(.146)	(.020)	(.051)	(.030)
	[.000]	[.011]	[.000]	[.000]	[.000]	[.000]	[.007]	[.000]	[.000]	[.000]
constant	-15.6	-15.5	-15.8	-15.9	-15.9	-15.2	-15.8	-16.2	-17.3	-16.7
	(.063)	(.206)	(.100)	(.135)	(.118)	(.114)	(.779)	(.108)	(.285)	(.148)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.5298	.2555	.5383	.4935	.5338	.8673	.4628	.8832	.8250	.8782
	(ii) paper fixed effects									
$\log_{10}$ collinearity	-.287	.130	.253	.239	.253	-.938	.519	.841	.896	.860
	(.032)	(.043)	(.030)	(.042)	(.031)	(.031)	(.100)	(.027)	(.040)	(.021)
	[.000]	[.029]	[.000]	[.001]	[.000]	[.000]	[.005]	[.000]	[.000]	[.000]
R^2	.6241	.4240	.6168	.5804	.6141	.8930	.6474	.8987	.8616	.8988
	(iii) paper x regression fixed effects									
$\log_{10}$ collinearity	-.371	.419	.338	.382	.341	-1.03	1.11	.935	1.03	.949
	(.027)	(.049)	(.024)	(.031)	(.025)	(.021)	(.089)	(.009)	(.033)	(.012)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.9293	.8564	.9313	.9284	.9319	.9120	.7943	.9176	.9075	.9179
N (i)-(iii)	(1) - (5) = 8342					(6) - (10) = 382586				

Notes: As above.

Table A2. Determinants of Log_{10} Coefficient of Variation by Collinearity Measure
(10 rotations each of 837 2SLS regression specifications in 28 papers)

	ivregress & *xtivreg*			*ivreg2* & *xtivreg2*			method D with demeaned variables & matrix inverses			method D with demeaned variables & linear solvers		
(A) Dependent variable: coefficient of variation of coefficients on instrumented variable ($\hat{\boldsymbol{\beta}}_1$)												
(i) no fixed effects (but includes a constant term)												
$\log_{10}$ $1\text{-}R^{2Max}$	-1.16 (.215) [.000]	-1.33 (.175) [.000]		-.168 (.307) [.596]	-.319 (.249) [.235]		.066 (.185) [.730]	-.007 (.142) [.960]		-.098 (.098) [.340]	-.131 (.071) [.101]	
$\log_{10}$ CN_b	.634 (.211) [.013]		2.99 (.555) [.001]	.565 (.262) [.058]		.779 (.626) [.251]	.290 (.144) [.073]		.031 (.280) [.913]	.163 (.087) [.090]		.242 (.154) [.156]
$\log_{10}$ SCN_b		.488 (.166) [.017]	-1.35 (.543) [.039]		.433 (.198) [.058]	-.070 (.582) [.907]		.226 (.107) [.064]	.201 (.267) [.474]		.135 (.064) [.064]	.005 (.155) [.975]
(ii) paper fixed effects												
$\log_{10}$ $1\text{-}R^{2Max}$	-1.29 (.299) [.001]	-1.47 (.348) [.001]		.119 (.220) [.598]	.122 (.250) [.634]		-.058 (.291) [.845]	-.094 (.333) [.782]		-.311 (.141) [.048]	-.387 (.170) [.043]	
$\log_{10}$ CN_b	.531 (.285) [.088]		1.85 (.698) [.020]	.825 (.183) [.001]		.500 (.407) [.242]	.170 (.228) [.469]		.353 (.258) [.195]	-.022 (.134) [.874]		.543 (.311) [.105]
$\log_{10}$ SCN_b		.367 (.335) [.296]	-.178 (.730) [.811]		.829 (.214) [.002]	.220 (.401) [.593]		.138 (.268) [.615]	-.132 (.258) [.617]		-.-090 (.159) [.584]	-.291 (.325) [.388]
(iii) paper x regression fixed effects												
$\log_{10}$ $1\text{-}R^{2Max}$	-.552 (.179) [.009]	-.238 (.258) [.373]		-.154 (.123) [.232]	.054 (.172) [.761]		-.050 (.132) [.714]	.070 (.186) [.712]		-.101 (.123) [.427]	-.009 (.173) [.959]	
$\log_{10}$ CN_b	1.32 (.176) [.000]		.784 (.227) [.004]	.730 (.116) [.000]		.407 (.231) [.100]	.281 (.120) [.037]		.018 (.165) [.914]	.247 (.117) [.057]		.024 (.186) [.899]
$\log_{10}$ SCN_b		1.62 (.247) [.000]	1.04 (.213) [.000]		.925 (.158) [.000]	.466 (.211) [.045]		.393 (.170) [.039]	.311 (.168) [.085]		.333 (.163) [.064]	.317 (.185) [.109]

Notes: Samples sizes and notes as in Table A1 above.

Table A2 reruns the specifications of Table A1 with pairs of the collinearity and conditioning measures entered alongside each other, focusing on the "b" measures for conditioning numbers as these have the highest R^2s in Table A1. Although highly statistically significant when entered alone in the regressions of Table A1, because of their collinearity when entered as pairs either one or both of the measures are often rendered statistically insignificant. However, as noted in the footnote in the paper, the conditioning measures appear to have the edge over 1- R^{2Max}. In panel (A), with coefficients of variation of instrumented coefficients as the dependent variable, the number of specifications in which 1- R^{2Max} is statistically significant at the .05 level in 24 head to head races with CN_b or SCN_b (7), is less than the 10 times (5 each)

Table A2 - continued

	(B) Dependent variable: coefficient of variation of coefficients on included instruments ($\hat{\boldsymbol{\beta}}_2$)											
	ivregress & *xtivreg*			*ivreg2* & *xtivreg2*			method D with demeaned variables & matrix inverses			method D with demeaned variables & linear solvers		
	(ii) no fixed effects (but includes a constant term)											
$\log_{10}$ 1-R^{2Max}	-1.21	-1.43		-.510	-.597		-.416	-.495		-.200	-.354	
	(.165)	(.107)		(.277)	(.228)		(.200)	(.145)		(.088)	(.051)	
	[.000]	[.000]		[.088]	[.021]		[.057]	[.004]		[.040]	[.000]	
$\log_{10}$ CN_b	.637		2.84	.315		1.11	.303		.911	.610		.699
	(.147)		(.417)	(.253)		(.580)	(.203)		(.389)	(.075)		(.231)
	[.001]		[.000]	[.234]		[.081]	[.158]		[.038]	[.000]		[.011]
$\log_{10}$ SCN_b		.452	-1.20		.244	-.369		.238	-.259		.484	.080
		(.103)	(.429)		(.215)	(.600)		(.159)	(.412)		(.047)	(.229)
		[.001]	[.016]		[.276]	[.550]		[.157]	[.541]		[.000]	[.732]
	(ii) paper fixed effects											
$\log_{10}$ 1-R^{2Max}	-.661	-.638		.050	.281		.022	.292		-.184	-.145	
	(.346)	(.398)		(.267)	(.339)		(.330)	(.432)		(.300)	(.341)	
	[.094]	[.144]		[.857]	[.429]		[.949]	[.517]		[.557]	[.682]	
$\log_{10}$ CN_b	1.19		.956	.910		-.256	.805		-.486	.679		.398
	(.304)		(.305)	(.202)		(.591)	(.263)		(.709)	(.265)		(.369)
	[.005]		[.015]	[.003]		[.677]	[.017]		[.514]	[.036]		[.315]
$\log_{10}$ SCN_b		1.24	.837		1.14	1.15		1.07	1.30		.730	.454
		(.361)	(.289)		(.282)	(.571)		(.367)	(.692)		(.309)	(.366)
		[.008]	[.021]		[.003]	[.080]		[.018]	[.099]		[.044]	[.251]
	(iii) paper x regression fixed effects											
$\log_{10}$ 1-R^{2Max}	-.343	.068		-.119	.141		-.029	.402		-.024	.325	
	(.136)	(.191)		(.089)	(.117)		(.106)	(.223)		(.136)	(.182)	
	[.034]	[.731]		[.217]	[.259]		[.788]	[.105]		[.866]	[.107]	
$\log_{10}$ CN_b	1.60		1.04	.898		.534	.901		.054	.914		.330
	(.126)		(.199)	(.080)		(.216)	(.098)		(.623)	(.128)		(.414)
	[.000]		[.001]	[.000]		[.037]	[.000]		[.933]	[.000]		[.446]
$\log_{10}$ SCN_b		1.99	.879		1.15	.478		1.31	.887		1.24	.614
		(.179)	(.198)		(.110)	(.214)		(.206)	(.635)		(.171)	(.421)
		[.000]	[.002]		[.000]	[.054]		[.000]	[.198]		[.000]	[.181]

Notes: Sample sizes and notes as in Table A1 above.

that these measures are significant at the same level in these comparisons. Similarly, in panel (B) where the dependent variable is the coefficient of variation of coefficients on included instruments, in 24 head to head races with CN_b or SCN_b the number of times 1- R^{2Max} is statistically significant at the .05 level (7 again) is well below the 20 times (10 each) CN_b and SCN_b are .05 significant in these comparisons.

B Determinants of Coefficient of Variation

In the paper I indicate that the coefficient of variation of 2SLS estimates for instrumented coefficients is increasing in the influence conditioning on the covariates has on the instrumented point estimate, but is not robustly significantly related to factors such as the strength of the first stage or the number of observations or instruments. This appendix substantiates that claim. In Table B1 below the dependent variable is the $\log_{10}$ coefficient of variation of estimated 2SLS coefficients across 50 permutations of variable order for the 10 collinearity increasing rotations of each of the 837 2SLS specifications in 28 papers, as described in the paper. $\log_{10}(1\text{-}R^{2Max})$ measures the collinearity induced by the random rotation of the included instruments, with R^{2Max} denoting the maximum partial R^2 (net of any fixed effects) of the regression of one instrument on the others. As a measure of the importance of conditioning on the included instruments, I use the proportional change in the estimated coefficient brought about by removing these, i.e. $\log_{10}|(\hat{\beta}_1 - \hat{\beta}_{1\sim\mathbf{X}_{1\sim c}})/\hat{\beta}_1|$, where $\hat{\beta}_1$ & $\hat{\beta}_{1\sim\mathbf{X}_{1\sim c}}$ are the estimated coefficients on the instrumented endogenous variable with and without the included instruments (other than the constant term or absorbed fixed effects) in the regression. Other regressors are the $\log_{10}$ number of rotated included instruments, number of observations and 1st stage heteroskedasticity robust or clustered (if the authors did so in their regression) F-statistic. The coefficients of variation of estimated coefficients are based in panel (A) on Stata's built-in routines *ivregress* and *xtivreg* (for specifications that have large numbers of absorbed fixed effects), in panel (B) the user written commands *ivreg2* and *xtivreg2*, in panel (C) method D with demeaned variables and matrix inverses in Mata, and in panel (D) method D with demeaned variables and solvers of linear equation systems in Mata. Sub-panel (i) includes a constant term in the regression and sub-panel (ii) paper fixed effects.[1] Standard errors (in parentheses) are clustered at the 28 paper level with corrections for bias brought about by high leverage points and p-values (in brackets) adjusted for effective degrees of freedom based upon the volatility of standard error estimates created by these leverage points, both using the command *edfreg*.

As shown in the left-hand columns of Table B, $\log_{10}(1\text{-}R^{2Max})$ and $\log_{10}|(\hat{\beta}_1 - \hat{\beta}_{1\sim\mathbf{X}_{1\sim c}})/\hat{\beta}_1|$ by themselves explain more than 90% of the variation in the log coefficient of variation of the coefficients of instrumented variables calculated using Stata's built in commands and between 50 and 75% of the variation for coefficients calculated using the alternative user routines or method D in Mata. Not surprisingly, the importance of conditioning on these instruments for the estimated coefficient on the endogenous variable has no robust relevance for the variation of the coefficients on the included instruments themselves, as shown in the right-hand columns. The number of included instruments, number of observations, and 1st stage F of each regression specification are sometimes significant at the .05 level, but not robustly so, as they are easily rendered insignificant with the inclusion of paper fixed effects or substitution of a different measure of variation.

[1]As regression characteristics other than the R^{2Max} are fixed across rotations of the included instruments, a specification with paper x regression fixed effects as in Appendix A cannot be used.

Table B1. Determinants of Log_{10} Coefficient of Variation
(11782 observations for 1179 2SLS specifications in 29 papers)

	Coefficients on Instrumented Variable ($\hat{\boldsymbol{\beta}}_1$)				Coefficients on Included Instruments ($\hat{\boldsymbol{\beta}}_2$)			
	(A) Dependent variable: coefficient of variation of *ivregress* & *xtivreg*							
	(i) without paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-1.84	-1.82	-1.84	-1.84	-1.94	-1.92	-1.91	-1.93
	(.049)	(.050)	(.048)	(.051)	(.039)	(.034)	(.039)	(.043)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.496	.471	.452	.503	.007			
	(.101)	(.078)	(.061)	(.107)	(.063)			
	[.001]	[.000]	[.000]	[.001]	[.915]			
$\log_{10}$ # of included instruments		.504				.332		
		(.131)				(.101)		
		[.005]				[.007]		
$\log_{10}$ # of observations			.508				.389	
			(.258)				(.161)	
			[.078]				[.032]	
$\log_{10}$ 1st stage F				-.315				-.165
				(.183)				(.153)
				[.133]				[.316]
constant	-15.3	-15.9	-16.7	-14.9	-15.0	-15.6	-16.2	-14.9
	(.255)	(.312)	(.624)	(.292)	(.153)	(.264)	(.430)	(.251)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
R^2	.9309	.9366	.9410	.9356	.9548	.9563	.9581	.9557
	(ii) with paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-1.89	-1.90	-1.89	-1.89	-1.99	-1.98	-1.98	-1.98
	(.027)	(.023)	(.027)	(.025)	(.052)	(.052)	(.052)	(.054)
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.512	.510	.512	.503	.084			
	(.058)	(.057)	(.057)	(.057)	(.039)			
	[.000]	[.000]	[.000]	[.000]	[.078]			
$\log_{10}$ # of included instruments		-.105				.310		
		(.431)				(.186)		
		[.815]				[.157]		
$\log_{10}$ # of observations			.108				.158	
			(.233)				(.189)	
			[.656]				[.447]	
$\log_{10}$ 1st stage F				-.575				-.416
				(.201)				(.280)
				[.035]				[.186]
R^2	.9664	.9665	.9665	.9710	.9656	.9655	.9655	.9669

Notes: Reported numbers = coefficient estimate, standard error estimate clustered at 28 paper level () and adjusted for bias, & p-value with effective degrees of freedom corrections [] (last two based on *edfreg*). R^{2Max} = maximum partial (net of any fixed effects) R^2 found in the regression of the instruments on each other; $\hat{\beta}_1$ & $\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}$ = coefficient on instrumented regressor with and without included instruments (other than constant term and absorbed fixed effects). Sample sizes as in Table A1 above.

Table B1 - continued

	Coefficients on Instrumented Variable ($\hat{\boldsymbol{\beta}}_1$)				Coefficients on Included Instruments ($\hat{\boldsymbol{\beta}}_2$)			
	(B) Dependent variable: coefficient of variation of *ivreg2* & *xtivreg2*							
	(i) without paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-.763 (.067) [.000]	-.745 (.066) [.000]	-.752 (.064) [.000]	-.761 (.067) [.000]	-.867 (.055) [.000]	-.866 (.057) [.000]	-.843 (.069) [.000]	-.874 (.053) [.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.718 (.113) [.000]	.701 (.112) [.001]	.673 (.110) [.000]	.722 (.115) [.000]	.244 (.123) [.085]			
$\log_{10}$ # of included instruments		.367 (.149) [.039]				.093 (.227) [.690]		
$\log_{10}$ # of observations			.534 (.282) [.088]				.380 (.250) [.154]	
$\log_{10}$ 1st stage F				-.150 (.112) [.228]				.081 (.140) [.581]
constant	-13.3 (.298) [.000]	-13.7 (.375) [.000]	-14.8 (.724) [.000]	-13.1 (.290) [.000]	-13.2 (.212) [.000]	-13.4 (.510) [.000]	-14.4 (.720) [.000]	-13.4 (.194) [.000]
R^2	.7246	.7359	.7666	.7286	.7882	.7803	.7926	.7807
	(ii) with paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-.810 (.058) [.000]	-.785 (.056) [.000]	-.810 (.057) [.000]	-.807 (.058) [.000]	-.964 (.082) [.000]	-.945 (.080) [.000]	-.959 (.081) [.000]	-.962 (.084) [.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.474 (.070) [.000]	.487 (.063) [.000]	.474 (.069) [.000]	.466 (.069) [.000]	.084 (.034) [.051]			
$\log_{10}$ # of included instruments		.654 (.192) [.014]				.834 (.205) [.010]		
$\log_{10}$ # of observations			.191 (.175) [.306]				.384 (.172) [.087]	
$\log_{10}$ 1st stage F				-.534 (.208) [.050]				-.384 (.293) [.236]
R^2	.8776	.8803	.8779	.8923	.8702	.8713	.8704	.8744

Notes: As above.

Table B1 - continued

	Coefficients on Instrumented Variable ($\hat{\boldsymbol{\beta}}_1$)				Coefficients on Included Instruments ($\hat{\boldsymbol{\beta}}_2$)			
	(C) Dependent variable: coefficient of variation of method D using matrix inverses							
	(i) without paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-.245 (.033) [.000]	-.234 (.035) [.000]	-.245 (.034) [.000]	-.243 (.032) [.000]	-.764 (.049) [.000]	-.759 (.047) [.000]	-.766 (.048) [.000]	-.762 (.050) [.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.258 (.030) [.000]	.248 (.023) [.000]	.259 (.029) [.000]	.261 (.039) [.000]	-.008 (.054) [.886]			
$\log_{10}$ # of included instruments		.226 (.092) [.038]				.087 (.182) [.644]		
$\log_{10}$ # of observations			-.016 (.064) [.813]				-.023 (.105) [.831]	
$\log_{10}$ 1st stage F				-.153 (.081) [.103]				-.065 (.097) [.524]
constant	-15.4 (.113) [.000]	-15.7 (.148) [.000]	-15.4 (.147) [.000]	-15.2 (.131) [.000]	-15.2 (.120) [.000]	-15.4 (.391) [.000]	-15.1 (.360) [.000]	-15.1 (.127) [.000]
R^2	.5285	.5576	.5288	.5573	.8058	.8064	.8059	.8066
	(ii) with paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-.257 (.060) [.002]	-.259 (.073) [.006]	-.257 (.060) [.002]	-.256 (.059) [.002]	-.874 (.071) [.000]	-.863 (.074) [.000]	-.871 (.071) [.000]	-.873 (.072) [.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.313 (.052) [.000]	.312 (.050) [.000]	.313 (.052) [.000]	.310 (.053) [.000]	.037 (.028) [.245]			
$\log_{10}$ # of included instruments		-.055 (.387) [.891]				.528 (.446) [.291]		
$\log_{10}$ # of observations			.038 (.165) [.824]				.357 (.175) [.108]	
$\log_{10}$ 1st stage F				-.168 (.089) [.116]				.034 (.108) [.759]
R^2	.6484	.6485	.6485	.6584	.8415	.8423	.8424	.8414

Notes: As above

Table B1 - continued

	Coefficients on Instrumented Variable ($\hat{\boldsymbol{\beta}}_1$)				Coefficients on Included Instruments ($\hat{\boldsymbol{\beta}}_2$)			
	(D Dependent variable: coefficient of variation of method D using linear solvers							
	(i) without paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-.269 (.018) [.000]	-.263 (.018) [.000]	-.270 (.018) [.000]	-.267 (.018) [.000]	-.901 (.027) [.000]	-.873 (.017) [.000]	-.888 (.024) [.000]	-.899 (.027) [.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.234 (.019) [.000]	.228 (.023) [.000]	.237 (.020) [.000]	.237 (.029) [.000]	-.039 (.066) [.572]			
$\log_{10}$ # of included instruments		.128 (.075) [.127]				.431 (.060) [.000]		
$\log_{10}$ # of observations			-.045 (.051) [.404]				.159 (.090) [.104]	
$\log_{10}$ 1st stage F				-.137 (.067) [.084]				-.055 (.113) [.639]
constant	-15.5 (.077) [.000]	-15.7 (.143) [.000]	-15.4 (.139) [.000]	-15.3 (.087) [.000]	-15.3 (.110) [.000]	-16.0 (.103) [.000]	-15.7 (.312) [.000]	-15.2 (.140) [.000]
R^2	.5996	.6091	.6016	.6228	.8675	.8778	.8696	.8677
	(ii) with paper fixed effects							
$\log_{10}$ 1-R^{2Max}	-.297 (.033) [.000]	-.309 (.038) [.000]	-.297 (.033) [.000]	-.296 (.033) [.000]	-.940 (.030) [.000]	-.938 (.034) [.000]	-.938 (.031) [.000]	-.938 (.031) [.000]
$\log_{10}\left\lvert\frac{\hat{\beta}_1-\hat{\beta}_{1\sim\mathbf{x}_{1\sim c}}}{\hat{\beta}_1}\right\rvert$	.326 (.045) [.000]	.321 (.047) [.000]	.326 (.045) [.000]	.324 (.045) [.000]	.045 (.028) [.161]			
$\log_{10}$ # of included instruments		-.314 (.224) [.210]				.040 (.376) [.920]		
$\log_{10}$ # of observations			-.086 (.106) [.443]				.092 (.280) [.760]	
$\log_{10}$ 1st stage F				-.157 (.089) [.139]				.055 (.117) [.653]
R^2	.7031	.7072	.7034	.7118	.8932	.8930	.8931	.8931

Notes: As above.

C Table 1 using the *ivregress* Command

The notes to Table 1 in the paper indicate that I follow the Oreopoulos (2006) code and use Stata's older *ivreg* command, but that results are nearly identical using the newer *ivregress* command. Table C1 shows this using the summary statistics for the range across permutations of data and variable order (panels c and d in Table 1 in the paper). As noted in the paper, this similarity only exists with frequency weights, and not with aweights, as with aweights *ivregress* does systematically worse.

Table 1 in the paper follows Oreopoulos' public use code for his UK regressions, using frequency weights [fw] instead of the more appropriate aweights [aw], where the weights are the number of observations used to produce the cell means that constitute his data. Frequency and aweights normally yield the same point estimates, but in nearly collinear data using Stata's built-in routines they do not. Moreover, when the weights are switched from frequency to aweights, the similarity between the volatility and bias of *ivregress* and *ivreg* ends, as *ivregress* (which has superseded *ivreg*) has worse average and worst case outcomes, as shown later in Tables 2 and 4 in the paper.

Table C1. Instrumented Effect of a Year's Education on ln UK Labour Income (Oreopoulos 2006)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
table/row/column	2/1/4	2/1/5	2/1/6	2/2/4	2/2/5	2/2/6	2/3/4	2/3/5	2/3/6	4/6/2	4/7/2	4/6/3	4/7/3	4/8/2	4/9/2
(a) replicated coefficient range in 10000 random permutations of data order: Intel Xeon W-2175 CPU															
using *ivreg* (as in Oreopoulos 2006 and reported in Table 1 in the paper)															
min	.091	.094	.100	.124	.177	.177	.036	.129	.127	.108	.054	-.056	-.032	.091	.100
5^{th} percentile	.101	.106	.110	.127	.179	.178	.038	.133	.131	.108	.054	-.056	-.032	.098	.109
95^{th} percentile	.122	.126	.129	.131	.182	.179	.043	.139	.137	.108	.054	-.055	-.032	.117	.129
max	.138	.144	.142	.133	.184	.179	.046	.144	.141	.108	.054	-.055	-.031	.141	.144
using *ivregress*															
min	.091	.094	.100	.124	.177	.177	.036	.129	.127	.108	.054	-.056	-.032	.091	.100
5^{th} percentile	.101	.106	.110	.127	.179	.178	.038	.133	.131	.108	.054	-.056	-.032	.098	.109
95^{th} percentile	.122	.126	.129	.131	.182	.179	.043	.139	.137	.108	.054	-.055	-.032	.117	.129
max	.138	.144	.142	.133	.184	.179	.046	.144	.141	.108	.054	-.055	-.031	.141	.144
(b) replicated coefficient range in 10000 random permutations of variable order: Intel Xeon W-2175 CPU															
using *ivreg* (as in Oreopoulos 2006 and reported in Table 1 in the paper)															
min	.091	-.018	-.007	.123	.082	.164	.021	.104	.055	.108	.053	-.056	-.035	.006	.012
5^{th} percentile	.093	.078	.067	.125	.161	.176	.027	.122	.113	.108	.053	-.056	-.033	.061	.069
95^{th} percentile	.176	.194	.298	.140	.196	.187	.057	.158	.172	.108	.054	-.055	-.031	.271	.287
max	.208	27.9	25.0	.141	5.80	2.81	.064	.264	13.3	.109	.056	-.054	-.027	8.80	30.0
using *ivregress*															
min	.091	-.018	-.007	.123	.082	.163	.021	.104	.055	.108	.053	-.056	-.035	.006	.012
5^{th} percentile	.093	.078	.067	.125	.161	.176	.027	.122	.113	.108	.053	-.056	-.033	.061	.069
95^{th} percentile	.176	.194	.298	.140	.196	.184	.057	.158	.172	.108	.054	-.055	-.031	.271	.287
max	.208	27.9	25.0	.141	5.80	.701	.064	.264	13.3	.109	.056	-.054	-.027	8.80	30.0

Notes: As in Table 1 in the paper.

D Rescaling/Standardizing Variables

As noted in a footnote in the paper, rescaling variables so that the matrix of inner-products is the identity matrix is sometimes recommended (e.g. Gould 2018) and ensures that the condition number of the K x K matrix is less than or equal to K times the minimum condition number attainable by any form of rescaling (van der Sluis 1969). However, it may worsen rather than improve the condition number and does nothing to reduce the dimensionality of matrix calculations. Tables D1 and D2 below show how it works out in practice, comparing results on the average and maximum coefficients of variation and bias found in 50 permutations of the variable order of the 10 collinearity increasing rotations of each regression in the 28 paper sample. The tables report results using methods A-D (as described in the paper) using the original data, demeaned data, rescaled data so that the matrix of inner-products is the identity matrix, and demeaned and rescaled (i.e. standardized) data. As noted in the paper and shown in these tables, relative to the original data, on average and in terms of worst case (maximal) outcomes rescaling alone achieves much less than demeaning, and when applied in combination with demeaning does not improve on what is achieved by demeaning alone.

Table D1. Coefficient of Variation using Different Methods
(across 50 permutations of variable order in 10 collinearity increasing rotations of instruments for each of 837 regressions in 28 papers)

	$\hat{\boldsymbol{\beta}}_1$ - coefficient on instrumented variable				$\hat{\boldsymbol{\beta}}_2$ - coefficients on included instruments			
	mean		max		mean		max	
	invert	solve	invert	solve	invert	solve	invert	solve
				(a) original data				
method A	.10	1.1e-08	926	4.6e-06	.30	1.2e-07	4128	3.9e-03
method B	2.6e-09	4.5e-09	1.5e-06	1.5e-06	4.6e-08	5.8e-08	2.5e-03	2.7e-03
method C	8.6e-10	3.9e-09	6.7e-07	1.2e-06	8.6e-09	3.0e-08	3.5e-04	1.6e-03
method D	2.4e-12	4.7e-14	2.0e-09	4.2e-11	3.5e-09	2.4e-08	1.9e-04	1.4e-03
				(b) demeaned				
method A	7.3e-03	1.3e-09	9.2	2.8e-06	3.4e-02	3.5e-09	637	1.1e-04
method B	2.4e-10	4.1e-10	3.8e-07	5.0e-07	6.5e-09	1.6e-09	3.1e-03	5.8e-05
method C	7.4e-11	3.4e-10	1.1e-07	3.4e-07	2.3e-10	9.9e-10	2.1e-05	4.3e-05
method D	1.8e-14	1.8e-14	7.1e-12	6.1e-12	1.1e-10	4.8e-10	1.8e-05	3.6e-05
				(c) standardized				
method A	5.1e-02	8.3e-09	87	4.1e-06	.44	9.3e-08	15082	3.4e-03
method B	2.4e-09	3.5e-09	1.4e-06	1.4e-06	4.3e-08	3.8e-08	2.5e-03	1.6e-03
method C	7.3e-10	2.8e-09	6.4e-07	8.0e-07	6.4e-09	2.2e-08	2.4e-04	1.1e-03
method D	2.8e-12	4.7e-14	2.7e-09	5.6e-11	2.7e-09	1.5e-08	1.9e-04	1.2e-03
				(d) demeaned and standardized				
method A	2.9e-03	1.1e-09	3.5	2.2e-06	4.0e-02	3.0e-09	3580	9.2e-05
method B	2.4e-10	3.1e-10	3.9e-07	3.9e-07	6.2e-09	1.2e-09	3.5e-03	5.6e-05
method C	1.2e-10	2.7e-10	3.1e-07	3.2e-07	3.3e-10	7.8e-10	2.7e-05	4.6e-05
method D	1.8e-14	1.8e-14	6.1e-12	7.2e-12	1.5e-10	3.8e-10	2.6e-05	3.9e-05

Notes: As in Tables 2 - 4 in the paper.

Table D2. Relative Bias using Different Methods
(across 50 permutations of variable order in 10 collinearity increasing rotations of instruments for each of 837 regressions in 28 papers)

	$\hat{\beta}_1$ - coefficient on instrumented variable				$\hat{\beta}_2$ - coefficients on included instruments			
	mean		max		mean		max	
	invert	solve	invert	solve	invert	solve	invert	solve
				(a) original data				
method A	2.9e-02	1.7e-09	24	6.0e-07	.12	1.3e-08	27770	7.1e-04
method B	5.8e-10	6.0e-10	2.2e-07	2.6e-07	6.8e-09	3.6e-09	1.9e-04	4.4e-04
method C	1.0e-09	1.2e-09	8.1e-07	5.0e-07	1.2e-08	1.2e-08	6.5e-04	4.7e-04
method D	4.0e-13	1.1e-14	5.4e-10	2.8e-12	5.4e-09	3.9e-09	5.6e-04	4.0e-04
				(b) demeaned				
method A	1.0e-03	2.2e-10	.96	6.0e-07	1.7e-02	5.3e-10	3990	8.0e-06
method B	6.8e-11	6.1e-11	3.7e-08	3.9e-08	2.1e-09	3.1e-10	1.4e-04	8.1e-06
method C	1.6e-10	1.4e-10	2.6e-07	2.4e-07	4.4e-10	4.5e-10	2.0e-05	9.4e-06
method D	9.3e-15	9.5e-15	3.8e-12	3.8e-12	1.9e-10	1.6e-10	1.7e-05	5.3e-06
				(c) standardized				
method A	2.4e-02	1.0e-09	17	1.3e-06	8.9e-02	9.5e-09	61401	3.8e-04
method B	3.8e-10	7.1e-10	2.7e-07	2.7e-07	6.4e-09	8.6e-09	5.0e-04	4.0e-04
method C	5.8e-10	9.5e-10	9.9e-07	8.7e-07	4.4e-09	5.2e-09	2.5e-04	2.0e-04
method D	2.5e-13	1.7e-14	5.4e-10	2.9e-11	2.9e-09	2.7e-09	2.3e-04	2.4e-04
				(d) demeaned and standardized				
method A	5.3e-04	1.5e-10	1.2	2.2e-06	3.7e-03	4.4e-10	238	1.4e-05
method B	4.2e-11	4.3e-11	2.5e-08	2.5e-08	1.0e-09	2.4e-10	4.4e-05	1.5e-05
method C	1.6e-10	1.6e-10	6.3e-07	7.4e-07	4.2e-10	4.1e-10	2.2e-05	7.1e-06
method D	1.1e-14	1.1e-14	3.8e-12	3.8e-12	1.5e-10	1.5e-10	1.7e-05	5.3e-06

Notes: As in Tables 2 - 4 in the paper. Bias evaluated using 100 digit precision computations using the Advanpix Toolbox for Matlab, as described in the paper.

References

Gould, William. 2018. *The Mata Book: A Book for Serious Programmers and Those Who Want to Be*. College Station, TX: Stata Press.

van der Sluis, A. 1969. Condition Numbers and Equilibration of Matrices. *Numerische Mathematik* 14: 14-23. https://doi.org/10.1007/BF02165096.

Young, Alwyn. 2016. Improved, Nearly Exact, Statistical Inference using Effective Degrees of Freedom Corrections. Manuscript, London School of Economics.