Example: Small-Sample Properties of IV and OLS Estimators

Considerable technical analysis is required to characterize the finite-sample distributions of IV estimators analytically. However, simple numerical examples provide a picture of the situation. Consider a regression $y = x\beta + \varepsilon$ where there is a single right-hand-side variable, and a single instrument w, and assume x, w, and ε have the simple joint distribution given in the table below, where λ is the correlation of x and w, ρ is the correlation of x and ε, and $|\lambda| + |\rho| < 1$. The interpretation of the second row of the table, for example, is that $(x,w,\epsilon) = (1,1,-1)$ and $(x,w,\epsilon) = (-1,-1,1)$ each occur with probability $(1-\rho+\lambda)/8$:

x	w	ε	Prob
±1	±1	±1	$(1+\rho+\lambda)/8$
±1	±1	∓1	$(1-\rho+\lambda)/8$
±1	∓1	±1	$(1+\rho-\lambda)/8$
±1	∓1	∓1	$(1-\rho-\lambda)/8$

The random variables (x,w,ϵ) have mean zero, variance one, and $Ex\varepsilon = \rho$, $Exw = \lambda$, and $Ew\varepsilon = 0$. Their products have the joint distribution

xw	xε	wε	Prob
1	1	1	$(1+\rho+\lambda)/4$
1	-1	-1	$(1-\rho+\lambda)/4$
-1	1	-1	$(1+\rho-\lambda)/4$
-1	-1	1	$(1-\rho-\lambda)/4$

This implies $P(x\varepsilon=1) = (1+\rho)/2$. Then, in a sample of size n, $n((b_{OLS} - \beta) + 1)/2$ has an exact distribution that is binomial with n draws and probability $(1+\rho)/2$. Then $n^{1/2}(b_{OLS} - \beta)$ has mean $n^{1/2}\rho$ and variance $(1-\rho^2)$. Thus, $n\cdot MSE = n\cdot(\text{Variance} + \text{Bias}^2) = 1 + (n-1)\rho^2$. The asymptotic theory for the IV estimator establishes that $n^{1/2}(b_{IV} - \beta)$ is approximately normal with mean zero and $n\cdot MSE = 1/\lambda^{2}$, equal to the asymptotic variance $Ew^2/(Exw)^2$ This suggests that the larger n, ρ, and λ, the more likely that IV will be better than OLS.

We compare b_{OLS} and b_{IV} for samples of various sizes drawn from the distribution above, for different values of ρ and λ. The following tables summarize the results of 1000 replications of each sample. In these tables, Bias is the mean (in 1000 samples) of $n^{1/2}(b_{OLS} - \beta)$ or $n^{1/2}(b_{IV} - \beta)$, while MSE is the mean (in 1000 samples) of $n(b_{OLS} - \beta)^2$ or $n(b_{IV} - \beta)^2$, where these moments for b_{IV} are calculated conditioned on the event that b_{IV} exists. The IV Pct. Finite column gives the proportion of the replications where b_{IV} exists; this is always less than one for this data generation process when n is even, but it converges toward one rapidly, so that for sample sizes above 40, it is negligible. The IV Pct. Better column gives the proportion of replications where b_{IV} is closer than b_{OLS} to the true β. Because of the thick tail for values of bIV, the sample sizes where IV Pct. Better exceeds 50 are smaller than the sample sizes where the sample expectation of MSE for IV (conditioned on IV existing) is less than that for OLS. The final columns of the table give some percentiles of the CDF's of $n^{1/2}(b_{OLS}-\beta)$ and $n^{1/2}(b_{IV}-\beta)$. One expects that this expression for OLS will drift due to the effect of bias, whereas the corresponding expression for b_{IV} will be approximately stationary. The results demonstrate the relatively thick tails of the expression for b_{IV}.

Mild Contamination, Moderately Good Instrument: $\rho = 0.2$, $\lambda = 0.5$													
Sample Size	Bias		IV Pct. Finite	MSE		IV Pct. Better	Probability (pct. less than)					
	OLS	IV		OLS	IV		OLS			IV		
							-2.5	0.0	2.5	-2.5	0.0	2.5
10	0.65	0.10	93.9	1.32	6.81	26.4	0	36	96	17	59	86
20	0.90	-0.10	98.9	1.79	8.50	34.1	0	24	94	14	58	88
30	1.11	-0.17	99.8	2.16	8.24	42.0	0	17	90	13	56	90
40	1.30	-0.11	100	2.65	6.42	45.5	0	13	86	12	55	89
60	1.55	-0.12	100	3.35	5.22	54.9	0	7	82	12	52	92
80	1.79	-0.13	100	4.15	7.45	61.3	0	5	77	13	54	90
100	1.98	-0.12	100	4.91	5.11	64.4	0	2	70	13	54	90
150	2.45	-0.08	100	6.92	4.33	76.2	0	1	53	10	52	91
200	2.84	-0.07	100	9.03	4.12	81.6	0	0	36	12	54	90
250	3.16	-0.08	100	11.0	4.01	85.9	0	0	24	11	53	91
300	3.47	-0.06	100	13.0	4.19	88.2	0	0	16	13	51	90

Existence of the IV estimator is a problem only for sample sizes under 40. IV is better a majority of the time for sample sizes above 40. Because IV has large deviations, its MSE is large even when one conditions on the existence of the IV estimator, so that in terms of this criterion, IV is better only for sample sizes over 100. The distribution of the OLS estimator is strongly shifted to the right, and increasingly so with sample size, due to the bias. The distribution of the IV estimator is roughly symmetric, with thick tails.

Severe Contamination, Moderately Good Instrument: $\rho = 0.5$, $\lambda = 0.4$												
Sample Size	Bias		IV Pct. Finite	MSE		IV Pct. Better	Probability (pct. less than)					
	OLS	IV		OLS	IV		OLS			IV		
							-2.5	0.0	2.5	-2.5	0.0	2.5
10	1.57	0.18	91.0	3.19	9.68	43.1	0	8	77	24	57	85
20	2.21	-0.49	97.4	5.60	17.4	59.8	0	1	61	21	56	88
30	2.71	-0.87	98.6	8.08	23.8	66.8	0	0	34	22	56	88
40	3.15	-1.01	99.6	10.6	29.2	73.2	0	0	19	23	56	86
60	3.85	-0.81	99.8	15.6	22.4	81.8	0	0	6	21	56	87
80	4.46	-0.63	100	20.6	11.3	86.1	0	0	2	21	55	89
100	4.99	-0.50	100	25.7	9.63	91.1	0	0	0	22	54	88
150	6.12	-0.39	100	38.1	8.04	94.9	0	0	0	20	55	87
200	7.07	-0.31	100	50.7	7.43	96.9	0	0	0	19	54	86
250	7.91	-0.28	100	63.3	7.37	98.2	0	0	0	18	53	85
300	8.68	-0.17	100	76.1	7.06	99.1	0	0	0	18	52	85

Existence of the IV estimator is an issue for sample sizes below 40. The IV estimator is better a majority of the time for sample sizes of 20 and higher. In terms of MSE, IV is better for sample sizes over 60.

Mild Contamination, Weak Instrument: $\rho = 0.2$, $\lambda = 0.2$													
Sample Size	Bias		IV Pct. Finite	MSE		IV Pct. Better	Probability (pct. less than)					
	OLS	IV		OLS	IV		OLS			IV		
							-2.5	0.0	2.5	-2.5	0.0	2.5
10	0.59	0.27	78.3	1.33	12.7	18.5	0	38	95	30	61	73
20	0.88	0.42	86.9	1.71	31.1	18.5	0	25	95	28	55	73
30	1.09	-0.07	91.7	2.11	52.9	21.8	0	19	91	32	56	71
40	1.25	-0.04	94.3	2.54	80.9	20.2	0	14	86	33	57	70
60	1.52	-0.08	96.9	3.31	101	23.9	0	8	83	32	56	68
80	1.74	-0.33	98.0	4.01	105	26.7	0	5	79	32	54	68
100	1.95	-0.48	98.3	4.76	121	28.8	0	3	71	34	55	71
150	2.40	-0.77	99.9	6.72	123	36.4	0	1	54	32	54	71
200	2.82	-0.82	100	8.90	71	40.5	0	0	38	32	54	71
250	3.16	-0.77	100	11.0	112	45.1	0	0	25	33	54	70
300	3.46	-0.43	100	13.0	36.7	49.0	0	0	16	33	54	69

Existence of the IV estimator is a substantial problem for sample sizes below 80. IV is never better in a majority of cases, even at a sample size of 300.

Severe Contamination, Weak Instrument: $\rho = 0.5$, $\lambda = 0.2$													
Sample Size	Bias		IV Pct. Finite	MSE		IV Pct. Better	Probability (pct. less than)					
	OLS	IV		OLS	IV		OLS			IV		
							-2.5	0.0	2.5	-2.5	0.0	2.5
10	1.57	1.16	81.6	3.21	13.4	29.9	0	7	77	25	50	67
20	2.26	0.94	89.6	5.85	29.7	36.6	0	1	59	23	48	66
30	2.79	0.79	93.1	8.50	48.0	41.6	0	0	30	26	47	64
40	3.21	0.28	95.9	11.0	65.5	47.3	0	0	16	27	48	66
60	3.91	-0.37	96.8	16.0	101	54.5	0	0	5	30	50	67
80	4.49	-0.70	98.5	20.8	141	61.4	0	0	1	30	50	69
100	5.03	-0.82	99.2	26.0	101	69.1	0	0	0	30	50	71
150	6.13	-1.37	99.7	38.3	105	76.5	0	0	0	30	53	71
200	7.06	-1.29	99.9	50.6	80.0	81.5	0	0	0	31	52	70
250	7.90	-1.19	100	63.1	72.4	84.6	0	0	0	31	53	70
300	8.66	-0.93	100	75.7	47.5	88.6	0	0	0	31	53	71

Existence of the IV estimator is a substantial problem for sample sizes below 80. The IV estimator is better in a majority of cases for sample sizes above 40. In terms of MSE, IV is better for sample sizes above 250.

Econ. 240B, Fall 2002, Dan McFadden