

Kernel Density Estimation on Grouped Data:

The Case of Poverty Assessment

SUPPLEMENTARY ONLINE APPENDIX

Camelia Minoiu and Sanjay G. Reddy

February 22, 2012

This document accompanies the manuscript available on

<http://www.camelia-minoiu.com/kde.pdf>

Table A1. Summary statistics (alternate bandwidths)

Bandwidths: S1, S3, over-smoothed, DPI-2

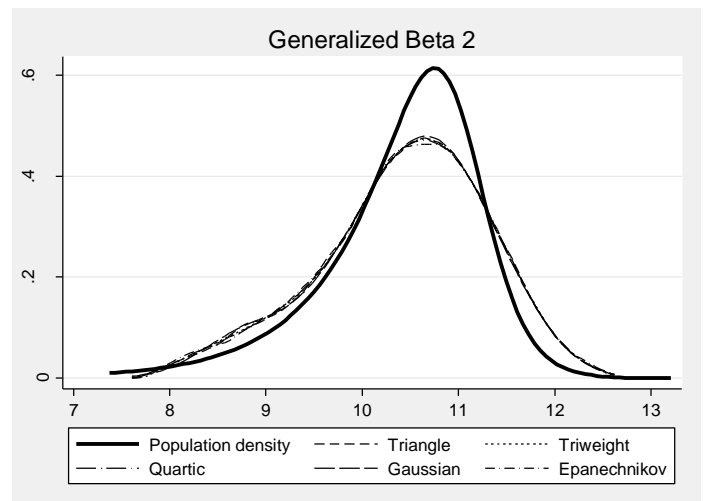
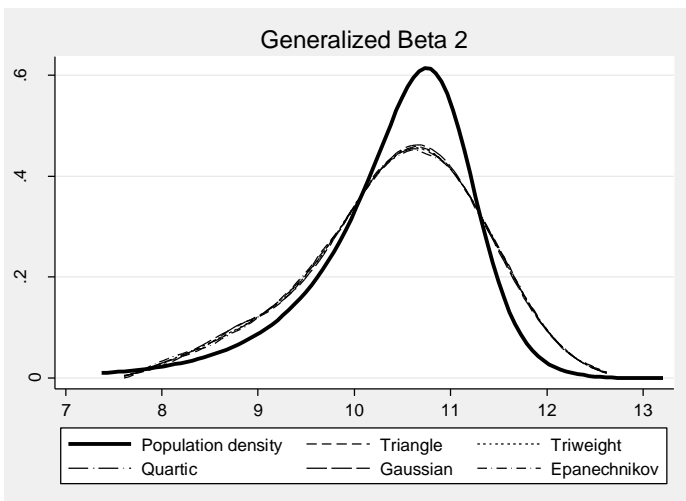
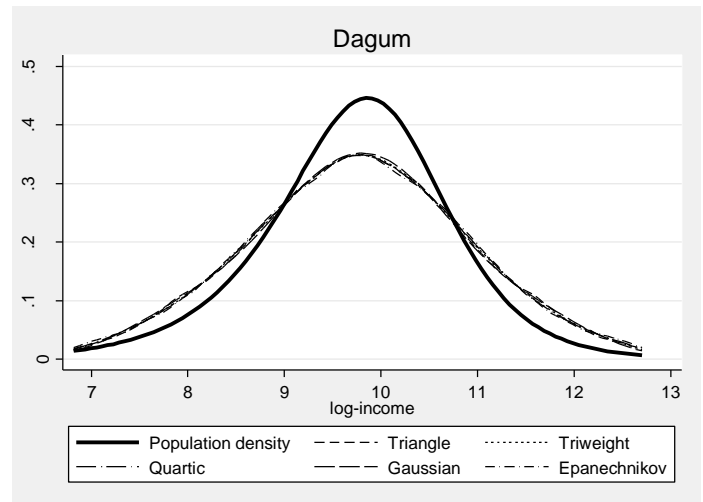
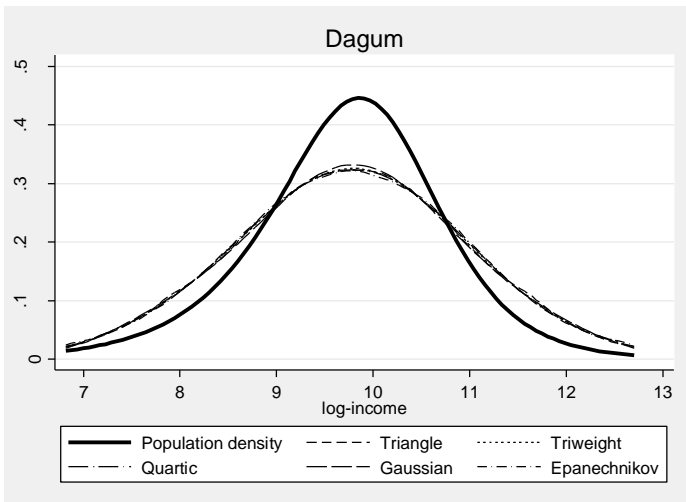
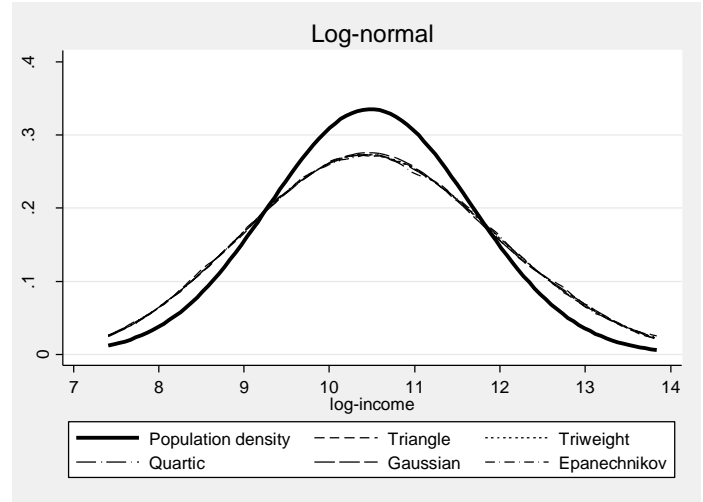
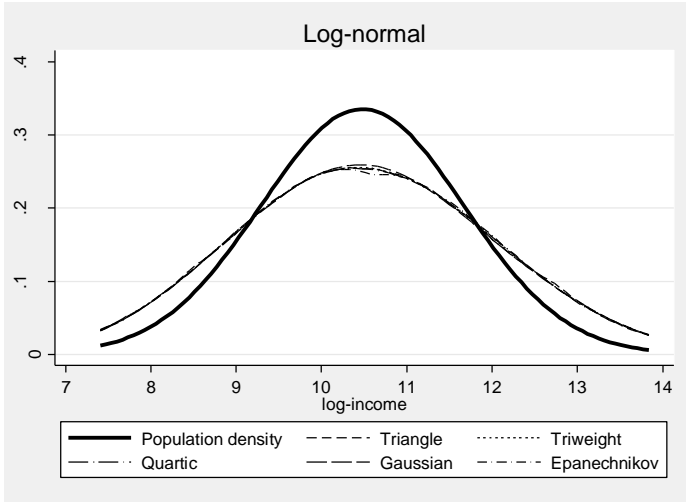
Kernel	Bandwidth	Underlying distribution	Summary statistics												
			Mean	Median	St. Dev.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Epanechnikov	S1	Log-normal	1.38	1.01	1.46	0.72	0.78	0.85	0.92	0.99	1.07	1.17	0.70	1.49	1.69
		Dagum	1.22	0.99	1.33	0.83	0.79	0.83	0.89	0.96	1.05	1.15	1.28	1.45	1.41
		GB2	1.11	0.95	1.31	0.98	0.86	0.88	0.91	0.94	0.99	1.05	1.13	1.23	1.38
	S3	Log-normal	1.25	1.00	1.21	0.82	0.84	0.89	0.94	1.00	1.05	1.12	0.66	1.37	1.45
		Dagum	1.11	0.99	1.08	0.94	0.86	0.89	0.94	0.98	1.03	1.09	1.18	1.31	1.19
		GB2	1.08	0.96	1.19	1.01	0.89	0.92	0.93	0.96	0.99	1.04	1.10	1.17	1.27
	over-smoothed	Log-normal	1.45	1.00	1.61	0.68	0.75	0.82	0.90	0.98	1.08	1.19	0.72	1.54	1.83
		Dagum	1.28	0.98	1.45	0.80	0.76	0.81	0.87	0.95	1.05	1.17	1.31	1.51	1.51
		GB2	1.17	0.94	1.51	0.92	0.82	0.83	0.87	0.92	0.99	1.07	1.19	1.34	1.57
	DPI-2	Log-normal	1.59	1.01	1.86	0.60	0.69	0.78	0.88	0.98	1.10	1.24	0.76	1.68	2.08
		Dagum	1.28	0.98	1.47	0.80	0.76	0.81	0.87	0.95	1.05	1.17	1.31	1.51	1.52
		GB2	1.14	0.94	1.43	0.95	0.84	0.85	0.88	0.93	0.99	1.06	1.16	1.29	1.49
Gaussian	S1	Log-normal	1.15	1.01	0.95	0.89	0.87	0.93	0.99	1.06	1.14	1.25	0.76	1.72	1.44
		Dagum	1.08	0.99	0.95	0.99	0.86	0.89	0.95	1.02	1.10	1.20	1.35	1.62	1.24
		GB2	1.07	0.94	1.16	1.10	0.90	0.90	0.93	0.97	1.03	1.10	1.20	1.36	1.36
	S3	Log-normal	1.08	1.01	0.85	0.98	0.92	0.96	1.01	1.06	1.12	1.19	0.71	1.56	1.27
		Dagum	1.01	0.99	0.84	1.09	0.93	0.95	0.98	1.02	1.07	1.14	1.24	1.45	1.08
		GB2	1.03	0.96	1.02	1.16	0.97	0.96	0.97	0.99	1.02	1.07	1.13	1.23	1.19
	over-smoothed	Log-normal	1.19	1.01	1.01	0.86	0.85	0.91	0.98	1.06	1.16	1.28	0.78	1.80	1.53
		Dagum	1.11	0.98	1.00	0.96	0.84	0.88	0.94	1.01	1.10	1.22	1.39	1.69	1.31
		GB2	1.08	0.94	1.20	1.08	0.88	0.88	0.91	0.96	1.03	1.11	1.22	1.40	1.42
	DPI-2	Log-normal	1.27	1.01	1.11	0.78	0.80	0.88	0.97	1.07	1.19	1.34	0.84	1.99	1.73
		Dagum	1.11	0.98	1.01	0.96	0.83	0.87	0.94	1.01	1.10	1.22	1.39	1.68	1.31
		GB2	1.07	0.94	1.15	1.10	0.90	0.90	0.93	0.97	1.02	1.10	1.20	1.35	1.36

Note: The figures represent the ratio between the KDE-based estimate from decile means (averaged over 2,000 replications) and the population value. Q1-Q10 are the ten decile means, from lowest (Q1) to highest (Q10). Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Figure A1. Visual illustrations: Impact of kernel (alternate bandwidths)

Bandwidth: DPI-1

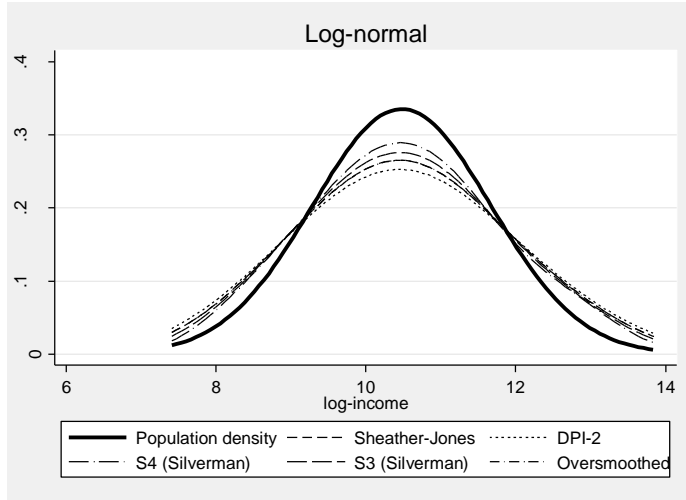
Bandwidth: S4



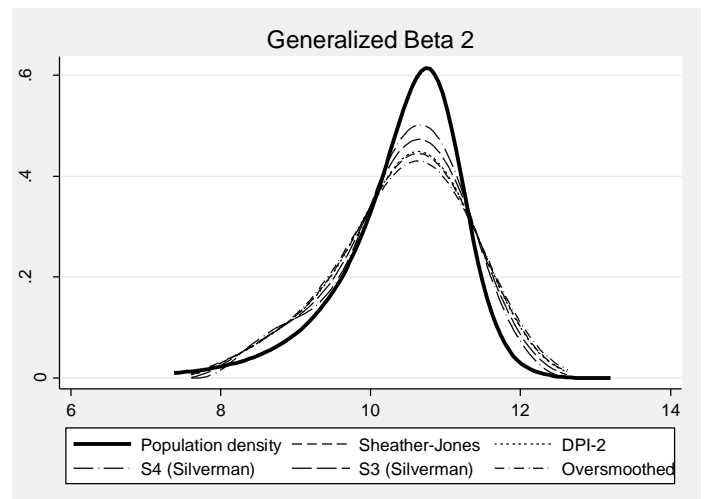
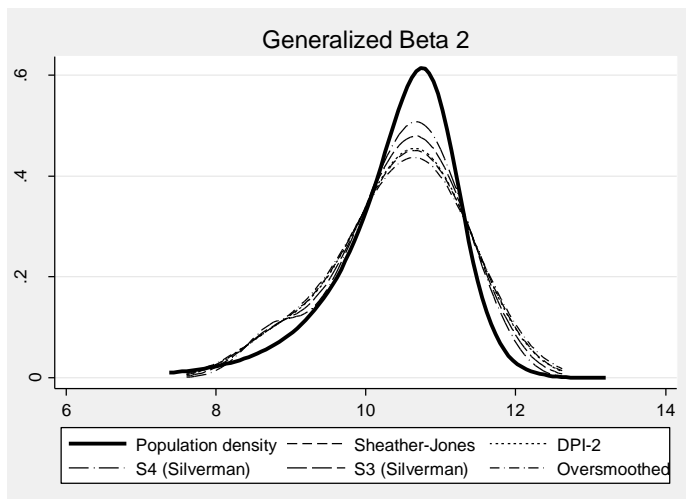
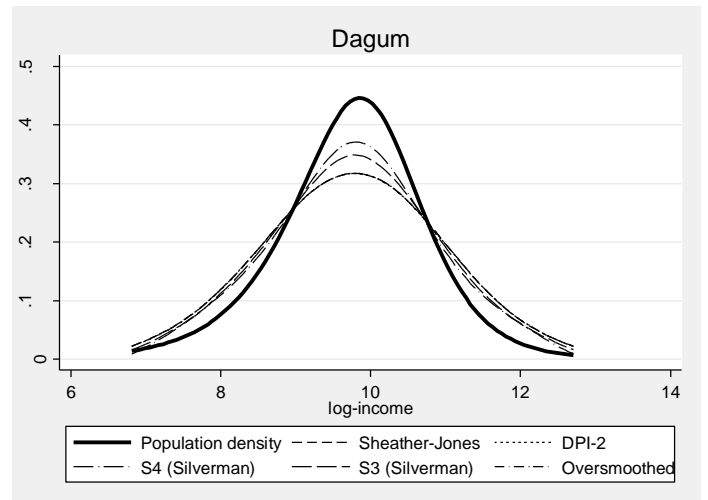
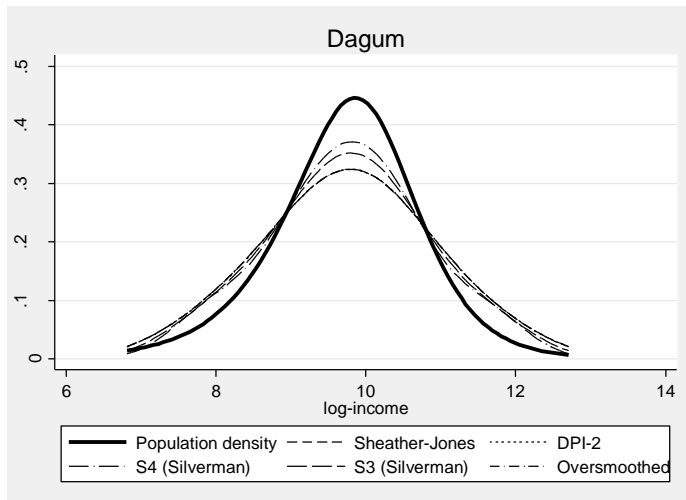
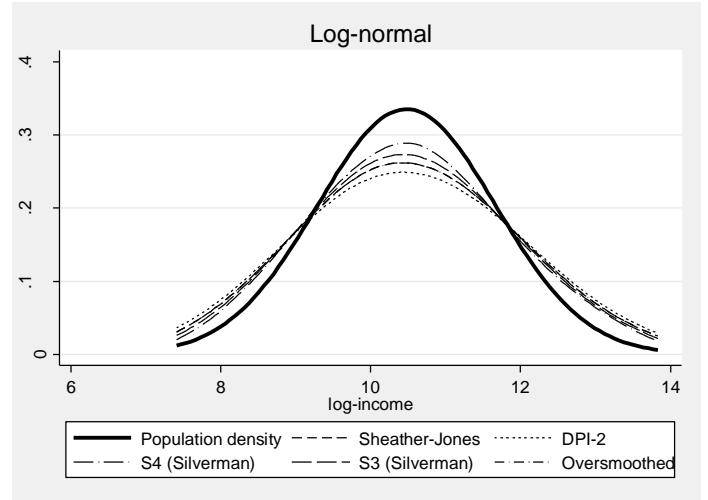
Note: The estimates are based on decile means.

Figure A2. Visual illustrations: Impact of bandwidth (alternate kernels)

Kernel: Gaussian



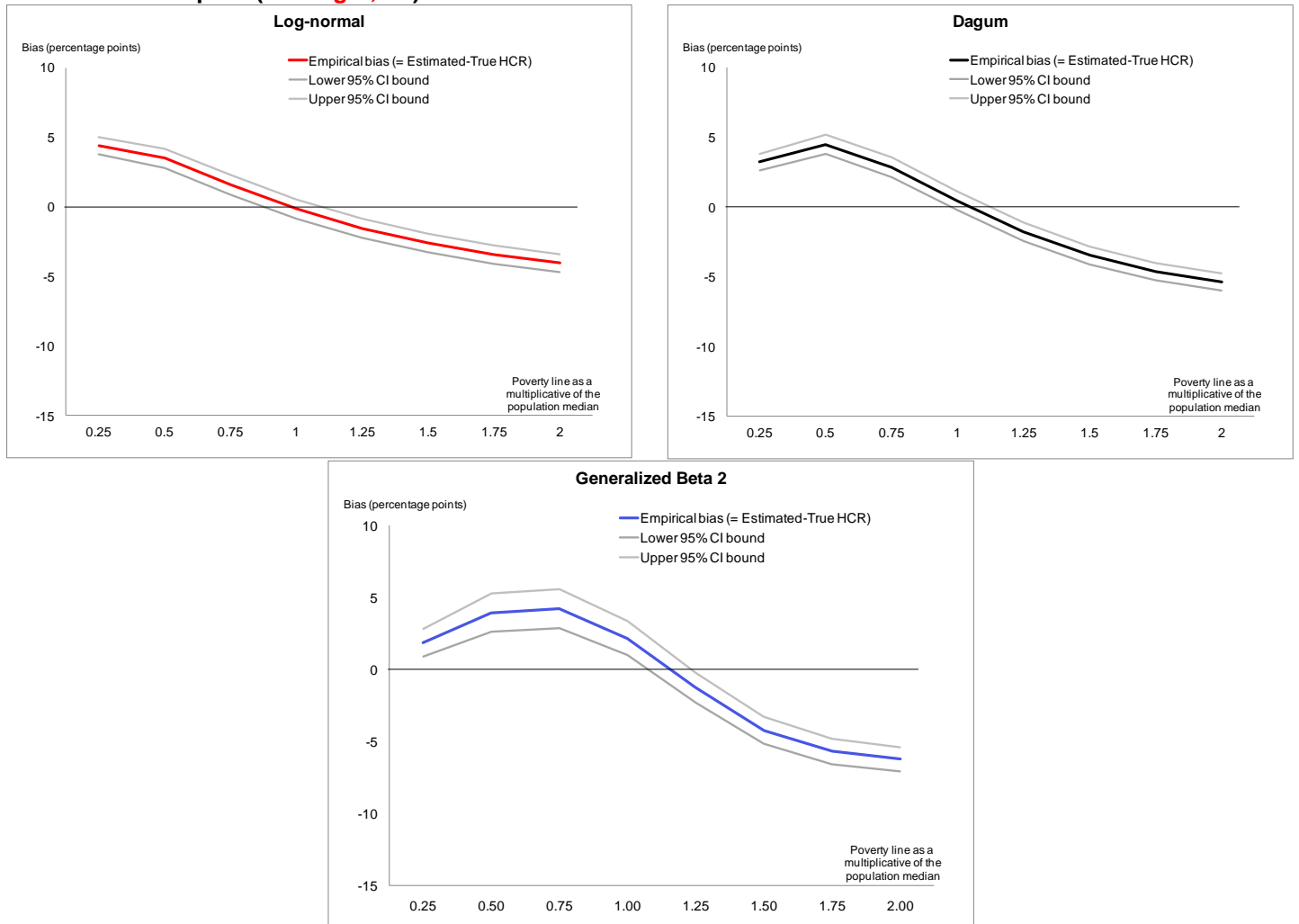
Kernel: Triweight



Note: The estimates are based on decile means.

Figure A3. How does the bias vary across poverty lines? (alternate kernel-bandwidth pairs)

Kernel-bandwidth pair: (Triweight, S2)



Note: The figures show the empirical bias in the poverty headcount ratio (that is, the difference between the estimated and the true headcount ratio, averaged over 2,000 replications), in percentage points, for different poverty lines. The poverty lines are expressed as multiples of the population median, ranging from a quarter of the median (0.25) to twice (2) the median. The estimates are based on decile means.

Table A2. How does the bias vary across poverty lines? (alternate kernel-bandwidth pairs)**Kernel: Gaussian, All bandwidths**

Bandwidth	Underlying distribution	Poverty line (as multiple of population median)							
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
S4	Log-normal	2.2	1.9	0.8	-0.2	-1.0	-1.7	-2.1	-2.5
	Dagum	1.3	2.6	1.9	0.4	-1.0	-2.1	-2.8	-3.3
	Generalized Beta 2	0.4	2.1	3.1	2.0	-0.3	-2.4	-3.3	-3.5
S3	Log-normal	2.2	1.9	0.8	-0.2	-1.0	-1.6	-2.1	-2.5
	Dagum	1.4	2.6	1.8	0.3	-1.1	-2.2	-2.9	-3.3
	Generalized Beta 2	0.5	2.1	3.0	2.0	-0.3	-2.4	-3.3	-3.6
S2	Log-normal	3.1	2.7	1.1	-0.3	-1.4	-2.2	-2.9	-3.3
	Dagum	2.1	3.6	2.4	0.4	-1.5	-2.9	-3.8	-4.4
	Generalized Beta 2	0.8	3.2	3.9	2.4	-0.5	-3.1	-4.3	-4.7
Sheather-Jones	Log-normal	3.3	2.8	1.2	-0.3	-1.4	-2.3	-3.0	-3.5
	Dagum	2.8	4.4	2.9	0.4	-1.8	-3.5	-4.6	-5.3
	Generalized Beta 2	1.2	4.2	4.8	2.7	-0.8	-3.8	-5.3	-5.8
DPI-1	Log-normal	3.9	3.2	1.4	-0.3	-1.6	-2.6	-3.4	-4.0
	Dagum	3.2	5.0	3.3	0.6	-1.9	-3.8	-5.0	-5.8
	Generalized Beta 2	1.2	4.4	5.0	2.9	-0.7	-3.8	-5.3	-5.9
DPI-2	Log-normal	3.9	3.2	1.4	-0.3	-1.6	-2.6	-3.4	-4.0
	Dagum	3.2	5.0	3.3	0.6	-1.9	-3.8	-5.0	-5.8
	Generalized Beta 2	1.5	5.0	5.5	3.0	-0.9	-4.2	-5.8	-6.5
S1	Log-normal	4.4	3.5	1.5	-0.3	-1.8	-2.9	-3.8	-4.4
	Dagum	2.9	4.6	3.0	0.5	-1.8	-3.5	-4.6	-5.3
	Generalized Beta 2	1.0	3.9	4.7	2.7	-0.6	-3.5	-4.9	-5.4
Over-smoothed	Log-normal	4.9	3.9	1.6	-0.4	-2.0	-3.2	-4.1	-4.8
	Dagum	3.3	5.0	3.2	0.5	-1.9	-3.8	-5.0	-5.8
	Generalized Beta 2	1.1	4.2	4.9	2.8	-0.7	-3.7	-5.2	-5.7

Note: The figures represent the percentage point difference between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means.

**Table A2. How does the bias vary across poverty lines? (alternate kernel-bandwidth pairs)
cont'd**

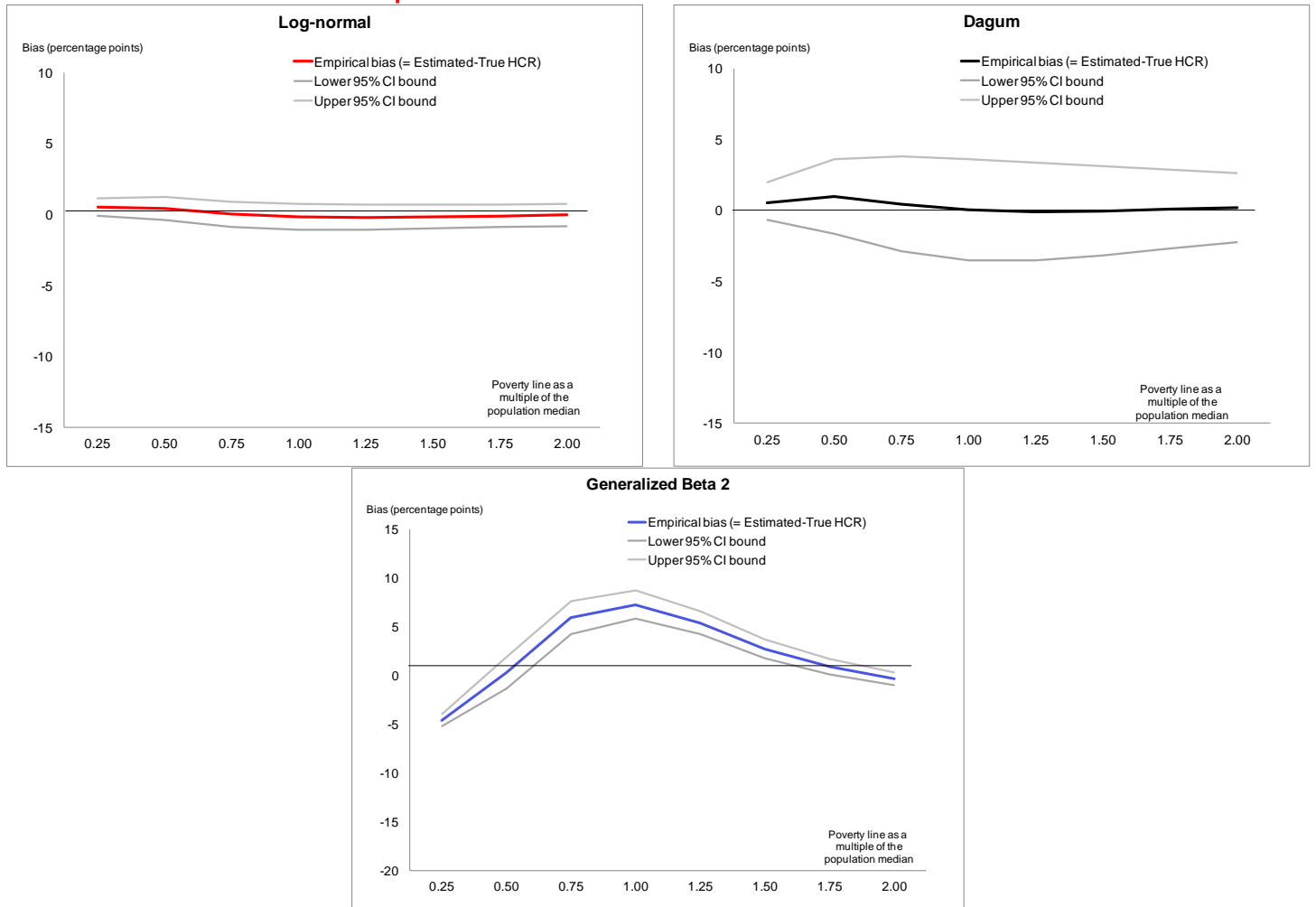
Kernel: Triweight, All bandwidths

Bandwidth	Underlying distribution	Poverty line (as multiple of population median)							
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
S4	Log-normal	3.4	2.7	1.2	-0.1	-1.2	-2.0	-2.7	-3.2
	Dagum	2.4	3.4	2.2	0.4	-1.3	-2.6	-3.5	-4.1
	Generalized Beta 2	1.5	2.9	3.5	2.0	-0.9	-3.4	-4.5	-4.9
S3	Log-normal	3.4	2.7	1.2	-0.1	-1.2	-2.0	-2.6	-3.1
	Dagum	2.4	3.4	2.2	0.4	-1.3	-2.7	-3.6	-4.2
	Generalized Beta 2	1.6	2.8	3.3	1.9	-0.9	-3.4	-4.5	-4.9
S2	Log-normal	4.4	3.5	1.6	-0.1	-1.5	-2.6	-3.4	-4.1
	Dagum	3.2	4.5	2.9	0.4	-1.8	-3.5	-4.6	-5.4
	Generalized Beta 2	1.9	3.9	4.2	2.2	-1.3	-4.2	-5.7	-6.2
Sheather-Jones	Log-normal	4.7	3.7	1.7	-0.1	-1.6	-2.8	-3.6	-4.2
	Dagum	4.0	5.4	3.4	0.5	-2.2	-4.1	-5.4	-6.3
	Generalized Beta 2	2.3	5.1	5.1	2.4	-1.6	-5.0	-6.8	-7.5
DPI-1	Log-normal	1.0	0.5	0.2	-0.1	-0.3	-0.5	-0.7	-0.9
	Dagum	1.3	0.9	0.6	0.1	-0.4	-0.8	-1.1	-1.2
	Generalized Beta 2	1.8	0.7	0.9	0.6	-0.1	-0.6	-1.0	-1.7
DPI-2	Log-normal	5.2	4.1	1.9	-0.2	-1.8	-3.0	-4.0	-4.7
	Dagum	4.5	6.0	3.8	0.6	-2.3	-4.4	-5.9	-6.8
	Generalized Beta 2	2.5	5.8	5.8	2.6	-1.8	-5.5	-7.5	-8.3
S1	Log-normal	4.8	3.8	1.7	-0.1	-1.6	-2.8	-3.7	-4.3
	Dagum	3.3	4.6	3.0	0.5	-1.7	-3.5	-4.7	-5.4
	Generalized Beta 2	1.7	3.7	4.2	2.2	-1.1	-4.0	-5.4	-5.9
Over-smoothed	Log-normal	4.7	3.7	1.7	-0.1	-1.6	-2.7	-3.6	-4.2
	Dagum	3.2	4.5	2.9	0.5	-1.7	-3.4	-4.5	-5.2
	Generalized Beta 2	1.6	3.4	3.9	2.1	-1.0	-3.8	-5.1	-5.6

Note: The figures represent the percentage point difference between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means.

Figure A4. How does the bias vary across poverty lines? (alternate estimation method)

Estimation method: **Lorenz Curve parametric estimatio: GQ**



Note: The figures show the empirical bias in the poverty headcount ratio (that is, the difference between the estimated and the true headcount ratio, averaged over 2,000 replications), in percentage points, for different poverty lines. The poverty lines are expressed as multiples of the population median, ranging from a quarter of the median (0.25) to twice (2) the median. The estimates are based on decile means.

Table A3. How does the bias vary across bandwidths? (alternate kernels)**Kernel: Triweight**

Underlying distribution	Poverty line (as multiple of population median)	Bandwidth							
		S4	S3	S2	Sheather-Jones	DPI-1	DPI-2	S1	Over-smoothed
Log-normal	0.25	1.28	1.28	1.36	1.38	1.09	1.43	1.39	1.38
	0.50	1.10	1.28	1.36	1.38	1.09	1.43	1.39	1.38
	0.75	1.03	1.03	1.04	1.04	1.01	1.05	1.04	1.04
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	1.25	0.98	0.98	0.97	0.97	0.99	0.97	0.97	0.97
	1.50	0.97	0.97	0.96	0.96	0.99	0.95	0.96	0.96
	1.75	0.96	0.96	0.95	0.95	0.99	0.94	0.95	0.95
2.00	0.96	0.96	0.94	0.94	0.99	0.93	0.94	0.94	
Dagum	0.25	1.26	1.26	1.34	1.42	1.13	1.48	1.35	1.33
	0.50	1.14	1.26	1.34	1.42	1.13	1.48	1.35	1.33
	0.75	1.06	1.06	1.08	1.09	1.02	1.10	1.08	1.08
	1.00	1.01	1.01	1.01	1.01	1.00	1.01	1.01	1.01
	1.25	0.98	0.98	0.97	0.96	0.99	0.96	0.97	0.97
	1.50	0.96	0.96	0.95	0.94	0.99	0.93	0.95	0.95
	1.75	0.95	0.95	0.94	0.93	0.99	0.92	0.94	0.94
2.00	0.95	0.95	0.93	0.92	0.98	0.91	0.93	0.93	
Generalized Beta 2	0.25	1.19	1.20	1.23	1.28	1.22	1.31	1.21	1.20
	0.50	1.14	1.20	1.23	1.28	1.22	1.31	1.21	1.20
	0.75	1.10	1.10	1.12	1.15	1.03	1.16	1.12	1.11
	1.00	1.04	1.04	1.04	1.05	1.01	1.05	1.04	1.04
	1.25	0.99	0.99	0.98	0.98	1.00	0.97	0.98	0.98
	1.50	0.96	0.96	0.94	0.93	0.99	0.93	0.95	0.95
	1.75	0.95	0.95	0.93	0.92	0.99	0.91	0.93	0.94
2.00	0.94	0.94	0.93	0.92	0.98	0.91	0.93	0.94	

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Table A3. How does the bias vary across bandwidths? (alternate kernels) cont'd**Kernel: Gaussian**

Underlying distribution	Poverty line (as multiple of population median)	Bandwidth							
		S4	S3	S2	Sheather-Jones	DPI-1	DPI-2	S1	Over-smoothed
Log-normal	0.25	1.18	1.18	1.25	1.32	1.36	1.40	1.27	1.32
	0.50	1.07	1.07	1.09	1.11	1.13	1.14	1.10	1.11
	0.75	1.02	1.02	1.03	1.03	1.04	1.04	1.03	1.03
	1.00	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99
	1.25	0.98	0.98	0.98	0.97	0.97	0.97	0.97	0.97
	1.50	0.97	0.97	0.96	0.96	0.95	0.95	0.96	0.96
	1.75	0.97	0.97	0.96	0.95	0.94	0.94	0.96	0.95
2.00	0.97	0.97	0.95	0.94	0.94	0.93	0.95	0.94	
Dagum	0.25	1.14	1.14	1.22	1.34	1.31	1.35	1.30	1.34
	0.50	1.11	1.11	1.15	1.21	1.19	1.21	1.19	1.21
	0.75	1.05	1.05	1.06	1.09	1.08	1.09	1.08	1.09
	1.00	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
	1.25	0.98	0.98	0.98	0.97	0.97	0.97	0.97	0.97
	1.50	0.97	0.97	0.96	0.94	0.95	0.94	0.95	0.94
	1.75	0.96	0.96	0.95	0.93	0.94	0.93	0.94	0.93
2.00	0.96	0.96	0.94	0.93	0.93	0.93	0.93	0.93	
Generalized Beta 2	0.25	1.05	1.07	1.10	1.15	1.12	1.14	1.15	1.19
	0.50	1.10	1.10	1.15	1.21	1.19	1.21	1.21	1.25
	0.75	1.09	1.09	1.11	1.14	1.13	1.14	1.14	1.16
	1.00	1.04	1.04	1.05	1.06	1.05	1.06	1.05	1.06
	1.25	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99
	1.50	0.97	0.97	0.96	0.95	0.95	0.95	0.95	0.94
	1.75	0.96	0.96	0.95	0.94	0.94	0.94	0.94	0.93
2.00	0.96	0.96	0.95	0.93	0.94	0.94	0.93	0.93	

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Table A4. How does the bias vary across kernels? (alternate bandwidths)**Bandwidth: S1**

Underlying distribution	Kernel	Poverty line (as multiple of population median)							
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Log-normal	Epanechnikov	1.38	1.13	1.04	1.00	0.97	0.96	0.95	0.94
	Gaussian	1.27	1.10	1.03	0.99	0.97	0.96	0.96	0.95
	Quartic	1.38	1.13	1.04	1.00	0.97	0.96	0.95	0.94
	Triweight	1.38	1.13	1.04	1.00	0.97	0.96	0.95	0.94
Dagum	Epanechnikov	1.41	1.24	1.09	1.01	0.96	0.94	0.92	0.92
	Gaussian	1.30	1.19	1.08	1.01	0.97	0.95	0.94	0.93
	Quartic	1.42	1.23	1.09	1.01	0.96	0.94	0.93	0.92
	Triweight	1.42	1.23	1.09	1.01	0.96	0.94	0.93	0.92
Generalized Beta 2	Epanechnikov	1.26	1.25	1.15	1.05	0.97	0.93	0.92	0.91
	Gaussian	1.15	1.21	1.14	1.05	0.99	0.95	0.94	0.93
	Quartic	1.28	1.25	1.15	1.05	0.98	0.93	0.92	0.91
	Triweight	1.28	1.25	1.15	1.05	0.98	0.93	0.92	0.92

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Bandwidth: DPI-2

Underlying distribution	Kernel	Poverty line (as multiple of population median)							
		0.25	0.5	0.75	1	1.25	1.5	1.75	2
Log-normal	Epanechnikov	1.54	1.18	1.06	1.00	0.96	0.94	0.93	0.92
	Gaussian	1.40	1.14	1.04	0.99	0.97	0.95	0.94	0.93
	Quartic	1.54	1.18	1.06	1.00	0.96	0.94	0.93	0.92
	Triweight	1.38	1.13	1.04	1.00	0.97	0.96	0.95	0.94
Dagum	Epanechnikov	1.47	1.26	1.10	1.01	0.96	0.93	0.92	0.91
	Gaussian	1.35	1.21	1.09	1.01	0.97	0.94	0.93	0.93
	Quartic	1.48	1.26	1.10	1.01	0.96	0.93	0.92	0.91
	Triweight	1.33	1.19	1.08	1.01	0.97	0.95	0.94	0.93
Generalized Beta 2	Epanechnikov	1.25	1.25	1.15	1.05	0.98	0.93	0.92	0.91
	Gaussian	1.14	1.21	1.14	1.06	0.99	0.95	0.94	0.94
	Quartic	1.26	1.25	1.15	1.05	0.98	0.93	0.92	0.92
	Triweight	1.20	1.17	1.11	1.04	0.98	0.95	0.94	0.94

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Table A5. How does the bias vary across poverty indicators? (alternate kernel-bandwidth pairs)

Kernel-bandwidth pair: (Gaussian, S4)

Underlying distribution	Poverty indicator	Poverty line (as multiple of population median)							
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Log-normal	Poverty headcount ratio (FGT0)	1.18	1.07	1.02	1.00	0.98	0.97	0.97	0.97
	Poverty gap ratio (FGT1)	1.12	1.11	1.07	1.04	1.02	1.01	1.00	1.00
	Squared poverty gap (FGT2)	1.01	1.12	1.10	1.07	1.05	1.04	1.03	1.02
	FGT3	0.89	1.10	1.10	1.09	1.07	1.06	1.05	1.04
	FGT4	0.76	1.07	1.10	1.10	1.08	1.07	1.06	1.05
Dagum	Poverty headcount ratio (FGT0)	1.14	1.11	1.05	1.01	0.98	0.97	0.96	0.96
	Poverty gap ratio (FGT1)	0.88	1.08	1.08	1.05	1.03	1.02	1.00	0.99
	Squared poverty gap (FGT2)	0.64	1.01	1.06	1.06	1.05	1.04	1.03	1.02
	FGT3	0.45	0.91	1.03	1.05	1.06	1.05	1.04	1.03
	FGT4	0.31	0.81	0.98	1.03	1.05	1.05	1.05	1.04
Generalized Beta 2	Poverty headcount ratio (FGT0)	1.05	1.10	1.09	1.04	1.00	0.97	0.96	0.96
	Poverty gap ratio (FGT1)	0.71	1.01	1.07	1.07	1.05	1.02	1.01	1.00
	Squared poverty gap (FGT2)	0.44	0.89	1.01	1.05	1.05	1.04	1.03	1.02
	FGT3	0.26	0.77	0.94	1.01	1.03	1.04	1.04	1.03
	FGT4	0.15	0.65	0.86	0.96	1.00	1.03	1.03	1.04

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means. The poverty lines are expressed as multiples of the population median, ranging from a quarter of the median (0.25) to twice (2) the median. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Kernel-bandwidth pair: (Gaussian, DPI-2)

Underlying distribution	Poverty indicator	Poverty line (as multiple of population median)							
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Log-normal	Poverty headcount ratio (FGT0)	1.40	1.14	1.04	0.99	0.97	0.95	0.94	0.93
	Poverty gap ratio (FGT1)	1.53	1.29	1.17	1.10	1.06	1.03	1.01	1.00
	Squared poverty gap (FGT2)	1.56	1.38	1.26	1.18	1.13	1.09	1.07	1.05
	FGT3	1.53	1.44	1.32	1.24	1.19	1.15	1.12	1.09
	FGT4	1.47	1.48	1.37	1.29	1.24	1.19	1.16	1.13
Dagum	Poverty headcount ratio (FGT0)	1.35	1.21	1.09	1.01	0.97	0.94	0.93	0.93
	Poverty gap ratio (FGT1)	1.14	1.25	1.18	1.12	1.07	1.04	1.01	0.99
	Squared poverty gap (FGT2)	1.32	1.39	1.32	1.25	1.18	1.13	1.10	1.07
	FGT3	0.71	1.14	1.21	1.20	1.17	1.14	1.11	1.09
	FGT4	0.55	1.05	1.18	1.20	1.19	1.17	1.14	1.12
Generalized Beta 2	Poverty headcount ratio (FGT0)	1.14	1.21	1.14	1.06	0.99	0.95	0.94	0.94
	Poverty gap ratio (FGT1)	0.81	1.12	1.15	1.13	1.08	1.05	1.02	1.00
	Squared poverty gap (FGT2)	0.54	1.00	1.11	1.13	1.12	1.09	1.07	1.05
	FGT3	0.35	0.87	1.05	1.10	1.12	1.11	1.09	1.08
	FGT4	0.23	0.75	0.97	1.06	1.10	1.11	1.10	1.09

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means. The poverty lines are expressed as multiples of the population median, ranging from a quarter of the median (0.25) to twice (2) the median. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Table A5*. How does the bias vary across poverty indicators? (alternate estimation method)

Estimation method: Lorenz Curve parametric estimation: GQ

Underlying distribution	Poverty indicator	Poverty line (as multiple of population median)							
		0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Log-normal	Poverty headcount ratio (FGT0)	1.04	1.01	1.00	1.00	1.00	1.00	1.00	1.00
	Poverty gap ratio (FGT1)	0.91	1.01	1.01	1.00	1.00	1.00	1.00	1.00
	Squared poverty gap (FGT2)	0.75	0.97	1.00	1.00	1.00	1.00	1.00	1.00
	FGT3	0.59	0.92	0.98	0.99	1.00	1.00	1.00	1.00
	FGT4	0.46	0.86	0.95	0.98	0.99	1.00	1.00	1.00
Dagum	Poverty headcount ratio (FGT0)	1.06	1.04	1.01	1.00	1.00	1.00	1.00	1.00
	Poverty gap ratio (FGT1)	1.06	1.04	1.01	1.00	1.00	1.00	1.00	1.00
	Squared poverty gap (FGT2)	0.68	0.96	1.01	1.01	1.01	1.01	1.01	1.01
	FGT3	0.52	0.89	0.98	1.00	1.01	1.01	1.01	1.01
	FGT4	0.39	0.81	0.94	0.98	1.00	1.01	1.01	1.01
Generalized Beta 2	Poverty headcount ratio (FGT0)	0.43	1.01	1.17	1.15	1.08	1.04	1.01	1.00
	Poverty gap ratio (FGT1)	0.22	0.69	0.96	1.06	1.08	1.07	1.06	1.05
	Squared poverty gap (FGT2)	0.13	0.49	0.77	0.92	1.00	1.04	1.05	1.05
	FGT3	0.08	0.35	0.61	0.79	0.90	0.97	1.00	1.02
	FGT4	0.05	0.26	0.50	0.68	0.81	0.89	0.95	0.98

Note: The figures represent the ratio between the grouped data, parametric estimate of the poverty headcount ratio (averaged over 2,000 replications) and the population value. The estimates are based on decile means. The poverty lines are expressed as multiples of the population median, ranging from a quarter of the median (0.25) to twice (2) the median. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Table A6. How does the bias vary with the number of quantile means? (alternate bandwidth-kernel pairs)

Kernel-bandwidth pairs: (Epanechnikov, S3); (Gaussian, S3)

Kernel	Underlying distribution	Number of data points	Poverty line (as multiple of population median)							
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Epanechnikov	Log-normal	Quintiles	1.16	1.08	1.02	0.99	0.97	0.96	0.95	0.94
		Deciles	1.27	1.10	1.03	1.00	0.98	0.97	0.96	0.96
		Ventiles	1.23	1.08	1.02	1.00	0.99	0.98	0.97	0.97
	Dagum	Quintiles	1.07	1.14	1.05	1.00	0.97	0.95	0.93	0.92
		Deciles	1.25	1.15	1.06	1.01	0.98	0.96	0.95	0.94
		Ventiles	1.20	1.11	1.05	1.01	0.98	0.97	0.96	0.96
	Generalized Beta 2	Quintiles	0.95	1.17	1.08	1.02	0.98	0.95	0.95	0.95
		Deciles	1.47	1.32	1.14	1.01	0.92	0.86	0.84	0.83
		Ventiles	1.12	1.10	1.08	1.03	0.99	0.97	0.96	0.96
Gaussian	Log-normal	Quintiles	0.98	1.03	1.00	0.99	0.98	0.97	0.97	0.96
		Deciles	1.18	1.07	1.02	1.00	0.98	0.97	0.97	0.97
		Ventiles	1.17	1.06	1.02	1.00	0.99	0.98	0.98	0.97
	Dagum	Quintiles	0.84	1.07	1.03	1.00	0.98	0.96	0.96	0.95
		Deciles	1.14	1.11	1.05	1.01	0.98	0.97	0.96	0.96
		Ventiles	1.13	1.09	1.04	1.01	0.99	0.98	0.97	0.97
	Generalized Beta 2	Quintiles	0.58	1.08	1.05	1.03	0.99	0.97	0.97	0.97
		Deciles	1.07	1.10	1.09	1.04	1.00	0.97	0.96	0.96
		Ventiles	1.04	1.09	1.07	1.03	1.00	0.97	0.97	0.97

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The poverty lines are expressed as multiples of the population median, ranging from a quarter of the median (0.25) to twice (2) the median. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Kernel-bandwidth pairs: (Quartic, S4); (Triweight, S4)

Kernel	Underlying distribution	Number of data points	Poverty line (as multiple of population median)							
			0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Quartic	Log-normal	Quintiles	1.16	1.08	1.02	0.99	0.97	0.96	0.95	0.94
		Deciles	1.27	1.10	1.03	1.00	0.98	0.97	0.96	0.96
		Ventiles	1.22	1.07	1.02	1.00	0.99	0.98	0.97	0.97
	Dagum	Quintiles	1.08	1.13	1.05	1.00	0.97	0.95	0.93	0.93
		Deciles	1.25	1.14	1.06	1.01	0.98	0.96	0.95	0.95
		Ventiles	1.19	1.11	1.05	1.01	0.98	0.97	0.97	0.96
	Generalized Beta 2	Quintiles	0.93	1.15	1.07	1.03	0.98	0.96	0.95	0.95
		Deciles	1.19	1.14	1.10	1.04	0.99	0.95	0.95	0.94
		Ventiles	1.11	1.11	1.08	1.03	0.99	0.97	0.96	0.96
Triweight	Log-normal	Quintiles	1.16	1.08	1.02	0.99	0.97	0.96	0.95	0.94
		Deciles	1.28	1.10	1.03	1.00	0.98	0.97	0.96	0.96
		Ventiles	1.22	1.07	1.02	1.00	0.99	0.98	0.97	0.97
	Dagum	Quintiles	1.09	1.13	1.05	1.00	0.97	0.95	0.94	0.93
		Deciles	1.26	1.14	1.06	1.01	0.98	0.96	0.95	0.95
		Ventiles	1.19	1.11	1.05	1.01	0.98	0.97	0.97	0.96
	Generalized Beta 2	Quintiles	0.92	1.15	1.07	1.03	0.98	0.96	0.95	0.95
		Deciles	1.19	1.14	1.10	1.04	0.99	0.96	0.95	0.94
		Ventiles	1.11	1.11	1.08	1.03	0.99	0.97	0.96	0.96

Note: The figures represent the ratio between the grouped data KDE-based poverty headcount ratio (averaged over 2,000 replications) and the population value. The poverty lines are expressed as multiples of the population median, ranging from a quarter of the median (0.25) to twice (2) the median. Figures in boldface represent cases in which the population value lies inside the 95 percent confidence interval around the average estimate.

Table A7. List of countries and availability of distributional data

Country	Initial year	Final year	Country	Initial year	Final year
Albania	1997	2005	Kyrgyz Republic	1993	2004
Argentina	1996	2005	Latvia	1995	2004
Armenia	1996	2003	Lithuania	1996	2004
Azerbaijan	1995	2005	Madagascar	1997	2005
Bangladesh	1995	2005	Malawi	1997	2004
Belarus	1995	2005	Malaysia	1995	2004
Bolivia	1997	2005	Mali	1994	2006
Brazil	1995	2005	Mexico	1995	2006
Bulgaria	1995	2003	Moldova, Republic	1997	2004
Burkina Faso	1994	2003	Mongolia	1995	2005
Cambodia	1994	2004	Nepal	1995	2003
Central African Republic	1993	2003	Nicaragua	1993	2005
Chile	1996	2006	Niger	1994	2005
China ¹	1995	2005	Nigeria	1996	2004
Colombia	1995	2006	Pakistan	1996	2004
Costa Rica	1996	2005	Panama	1995	2006
Dominican Republic	1996	2005	Paraguay	1995	2005
Ecuador	1994	2005	Peru	1996	2005
Egypt	1996	2004	Philippines	1997	2006
El Salvador	1995	2005	Poland	1996	2005
Estonia	1995	2004	Romania	1994	2005
Ethiopia	1995	2005	Russian Federation	1996	2005
Georgia	1996	2005	Senegal	1994	2005
Guinea	1994	2003	Slovenia	1993	2004
Honduras	1997	2005	Thailand	1996	2004
Hungary	1993	2004	Turkey	1994	2005
India ¹	1999	2005	Uganda	1996	2005
Indonesia ¹	1996	2005	Ukraine	1996	2005
Iran	1994	2005	Uruguay	1996	2005
Jamaica	1996	2004	Venezuela, RB	1995	2005
Jordan	1997	2006	Vietnam	1993	2006
Kazakhstan	1996	2003	Zambia	1996	2004
Kenya	1997	2005			

Note: For China, India, and Indonesia we use national distributional data from the WDI (World Development Indicators) and the UNU-WIDER World Income Inequality Database instead of the World Bank's Povcalnet, which reports tabulations separately for rural/urban areas.¹

¹ The databases are respectively available on <http://data.worldbank.org/data-catalog/world-development-indicators> and http://www.wider.unu.edu/research/Database/en_GB/database/. Povcalnet is available on <http://go.worldbank.org/7X6J3S7K90> (Accessed: August 15, 2011).

**Table A8. Sensitivity analysis of global poverty to alternative bandwidths (1995–2005)
(alternate kernels)**

Kernel: Epanechnikov

Bandwidth	Hybrid	S3	S4	Sheather- Jones	DPI-1	DPI-2	Over- smoothed	Ratio between highest and lowest value in columns 1-7	Pp diff. between highest and lowest estimate in columns 1-7
Poverty line	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
<i>Panel A. Poverty headcount ratio (% poor)</i>									
1995									
\$1.00	5.8	6.2	6.9	7.9	7.9	8.4	8.5	1.5	2.7
\$1.25	9.6	10.1	11.2	11.9	11.9	12.5	12.6	1.3	3.0
\$1.45	13.3	13.4	14.8	15.1	15.5	15.6	16.3	1.2	3.1
\$2.00	22.6	23.0	24.2	24.7	25.2	25.1	25.1	1.1	2.6
\$2.50	31.7	31.5	32.4	32.9	32.8	32.8	32.9	1.0	1.4
2005									
\$1.00	2.0	2.3	2.8	3.3	3.2	3.5	3.6	1.8	1.7
\$1.25	3.7	4.2	4.9	5.8	5.5	5.8	6.1	1.7	2.4
\$1.45	5.7	6.1	6.8	7.8	7.9	8.1	8.1	1.4	2.4
\$2.00	12.5	12.7	14.0	15.1	14.7	15.2	15.1	1.2	2.7
\$2.50	19.3	19.5	20.6	21.3	21.1	22.0	21.9	1.1	2.7
<i>Panel B. Poverty headcount (millions)</i>									
1995									
\$1.00	242	259	292	330	333	354	357	1.5	(millions) 115
\$1.25	403	423	472	498	501	526	529	1.3	126
\$1.45	558	563	622	636	651	654	687	1.2	129
\$2.00	948	965	1017	1035	1057	1056	1055	1.1	109
\$2.50	1333	1325	1363	1382	1379	1380	1380	1.0	58
2005									
\$1.00	92	108	131	153	150	162	170	1.8	(millions) 78
\$1.25	174	197	228	274	258	271	288	1.7	114
\$1.45	268	285	321	366	369	379	379	1.4	111
\$2.00	585	597	658	708	689	711	709	1.2	126
\$2.50	904	916	967	1000	988	1030	1027	1.1	127
<i>Panel C. Reduction in the poverty headcount ratio (pp)</i>									
1995-2005									
\$1.00	3.8	3.9	4.2	4.6	4.7	5.0	4.9	1.3	1.2
\$1.25	5.9	5.9	6.4	6.0	6.4	6.8	6.5	1.2	0.9
\$1.45	7.6	7.3	8.0	7.4	7.6	7.5	8.3	1.1	0.9
\$2.00	10.1	10.2	10.2	9.6	10.5	10.0	10.0	1.1	0.9
\$2.50	12.5	12.0	11.8	11.6	11.8	10.9	11.0	1.1	1.6
<i>Panel D. Reduction in the poverty headcount ratio (%)</i>									
1995-2005									
\$1.00	66	63	60	59	60	59	57	1.1	8.5
\$1.25	61	58	57	51	54	54	51	1.2	10.5
\$1.45	57	55	54	49	49	48	51	1.2	8.9
\$2.00	45	45	42	39	42	40	40	1.2	6.0
\$2.50	39	38	36	35	36	33	33	1.2	6.1
<i>Panel E. Reduction in the poverty headcount (millions)</i>									
1995-2005									
\$1.00	-150	-151	-161	-177	-183	-192	-187	1.3	(millions) 41.5
\$1.25	-229	-226	-244	-224	-243	-256	-242	1.1	31.2
\$1.45	-289	-278	-301	-271	-282	-274	-308	1.1	37.3
\$2.00	-363	-368	-359	-328	-368	-345	-346	1.1	40.1
\$2.50	-429	-408	-395	-383	-391	-349	-353	1.2	79.6

Note: The estimates are based on decile means.

**Table A9. Sensitivity analysis of global poverty to alternative bandwidths (1995-2005)
(alternate kernels)**

Kernel: Gaussian

Bandwidth	Hybrid	S3	S4	Sheather-Jones	DPI-1	DPI-2	Over-smoothed	Ratio between highest and lowest value in columns 1-7	Pp diff. between highest and lowest estimate in columns 1-7
Poverty line	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
<i>Panel A. Poverty headcount ratio (% poor)</i>									
1995									
\$1.00	5.8	6.4	7.3	7.7	7.9	8.5	8.5	1.5	2.7
\$1.25	9.9	10.4	10.9	11.9	12.1	12.2	12.8	1.3	2.8
\$1.45	13.2	13.8	14.5	15.7	15.3	16.0	16.1	1.2	2.9
\$2.00	22.8	23.7	23.8	24.9	24.7	25.4	25.3	1.1	2.6
\$2.50	31.4	31.9	32.2	33.3	32.7	33.2	33.2	1.1	1.9
2005									
\$1.00	2.1	2.4	3.0	3.5	3.5	3.7	3.9	1.9	1.8
\$1.25	3.9	4.2	5.0	6.0	5.8	6.0	6.3	1.6	2.3
\$1.45	5.7	6.1	7.0	8.0	7.8	8.3	8.3	1.5	2.6
\$2.00	12.5	12.7	13.8	14.8	14.5	14.9	15.2	1.2	2.8
\$2.50	18.8	19.6	20.6	21.4	21.4	21.3	21.9	1.2	3.1
<i>Panel B. Poverty headcount (millions)</i>									
1995									
\$1.00	243	269	308	323	333	355	357	1.5	114
\$1.25	416	436	458	502	507	513	536	1.3	120
\$1.45	555	578	611	659	642	673	676	1.2	121
\$2.00	956	994	999	1044	1036	1066	1064	1.1	111
\$2.50	1320	1342	1352	1399	1372	1393	1396	1.1	79
2005									
\$1.00	98	113	139	166	165	175	184	1.9	86
\$1.25	184	199	236	280	272	282	293	1.6	110
\$1.45	268	287	330	374	365	391	391	1.5	123
\$2.00	584	594	649	695	678	697	714	1.2	130
\$2.50	883	920	967	1003	1004	998	1029	1.2	146
<i>Panel C. Reduction in the poverty headcount ratio (pp)</i>									
1995-2005									
\$1.00	3.7	4.0	4.4	4.1	4.4	4.7	4.6	1.3	1.0
\$1.25	6.0	6.1	5.9	6.0	6.3	6.2	6.5	1.1	0.6
\$1.45	7.5	7.7	7.5	7.7	7.5	7.7	7.7	1.0	0.3
\$2.00	10.3	11.0	10.0	10.0	10.2	10.5	10.1	1.1	1.0
\$2.50	12.6	12.3	11.6	11.9	11.3	11.9	11.3	1.1	1.3
<i>Panel D. Reduction in the poverty headcount ratio (%)</i>									
1995-2005									
\$1.00	64	63	60	54	56	56	54	1.2	10.2
\$1.25	60	59	54	50	52	51	51	1.2	10.4
\$1.45	57	56	52	49	49	48	48	1.2	8.8
\$2.00	45	46	42	40	41	41	40	1.2	6.6
\$2.50	40	39	36	36	34	36	34	1.2	6.1
<i>Panel E. Reduction in the poverty headcount (millions)</i>									
1995-2005									
\$1.00	-145	-157	-169	-157	-168	-180	-173	1.2	34.5
\$1.25	-232	-237	-221	-222	-235	-231	-242	1.1	20.8
\$1.45	-287	-292	-281	-285	-278	-282	-285	1.0	13.9
\$2.00	-371	-400	-350	-349	-357	-369	-350	1.1	50.6
\$2.50	-437	-422	-385	-396	-368	-395	-367	1.2	70.4

Note: The estimates are based on decile means.