Appendix (For online publication only)

Cooking fuel choice and child mortality in India

Arnab K. Basu, Tsenguunjav Byambasuren, Nancy H. Chau and Neha Khanna

April 9, 2024

This supplementary material presents the results from robustness checks (Appendix A), additional figures and tables that report additional results (Appendix B), and findings from heterogeneity analysis by household size (Appendix C).

A Robustness check tables

Table A.1: Effect of cooking fuel choice on infant mortality from IV probit regression

	1^{st} stage	2^{nd} stage				
	(1) Polluting fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
Polluting fuel for cooking		0.456**	-0.141	0.501**	0.362	0.502**
		(0.210)	(0.574)	(0.213)	(0.324)	(0.238)
Forest cover	0.065***					
	(0.009)					
Owns agricultural land	0.054***					
	(0.002)					
Observations	194254	194254	192079	194254	193859	194254
R^2	0.53					
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	272.04					
Critical value 2SLS ($\tau = 10\%$)	8.27					
Model Wald χ^2		3855.19	395.07	3669.80	1463.43	2499.14
Model degrees of freedom	20.00	53.00	46.00	53.00	52.00	53.00
Model Wald p-value	0.00	0.00	0.00	0.00	0.00	0.00
Exogeneity test Wald p-value		0.09	0.69	0.06	0.34	0.10
Wald χ^2 test of exogeneity		2.90	0.16	3.57	0.92	2.75

Notes: The first column reports the result from the first-stage OLS regression of IV probit using DHS-4 where the dependent variable is a binary variable for polluting fuel. The effective F-statistic on IVs—2013 district-wise forest cover calculated as a percentage of forested area in the total geographical area of the region using satellite-based data and an indicator variable for household's agricultural land ownership—confirms that the instruments create a significant variation in polluting fuel for cooking. Columns (2)–(6) report coefficient estimates from the estimation of equation (1) using IV probit regression with different dependent variables and a similar specification. All specifications contain an unreported vector of baseline demographic controls, state-by-year fixed effects, and a constant term. Some state FEs are excluded because they perfectly predict the outcome variable in child and post-neonatal mortality regressions. Heteroskedasticity-robust standard errors, clustered by PSUs, are in parentheses. The standard errors of the key regressors in the first-stage regression and the effective F-statistic on the IVs are different from those in Column (1) of Table 7 due to differences in cluster level. However, the statistical significance of the key regressors is the same for all IV probit regressions when the standard errors are clustered by districts. Significance: $^*p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table A.2: Effect of dirtiness level of cooking fuels on infant mortality

	1^{st} stage	2^{nd} stage				
	(1) Polluting fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
Dirtiness level of cooking fuels		0.007**	-0.000	0.007**	0.002	0.005**
		(0.003)	(0.001)	(0.003)	(0.002)	(0.002)
Forest cover	0.191***					
	(0.051)					
Owns agricultural land	0.343***					
	(0.014)					
Observations	194254	194254	194254	194254	194254	194254
R^2	0.52	0.02	0.00	0.02	0.01	0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	273.62					
Critical value 2SLS ($\tau = 10\%$)	6.10					
Hansen J statistic		0.89	0.23	1.20	0.22	0.99
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.34	0.63	0.27	0.64	0.32

Notes: All specifications contain an unreported vector of baseline demographic controls, state-by-year fixed effects, and a constant term. The Hansen's J-statistic suggests that the excluded IVs are exogenous. Parentheses contain heteroskedasticity-robust standard errors clustered by PSUs. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table A.3: Effect of cooking fuel choice on infant mortality
(IVs = 2011 satellite-based forest cover and agricultural land ownership)

	1^{st} stage	2^{nd} stage				
	(1)	(2)	(3)	(4)	(5)	(6)
	Polluting fuel use	Under-five	Child	Infant	Post-neonatal	Neonatal
Polluting fuel for cooking		0.038**	-0.000	0.038**	0.010	0.028**
		(0.018)	(0.005)	(0.017)	(0.010)	(0.014)
Forest cover (*satellite-based, 2011*)	0.054***					
	(0.009)					
Owns agricultural land	0.057***					
	(0.002)					
Observations	203737	203737	203737	203737	203737	203737
R^2	0.53	0.02	0.00	0.02	0.01	0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	292.31					
Critical value 2SLS ($\tau = 10\%$)	8.38					
Hansen J statistic		2.41	0.46	3.12	0.66	2.45
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.12	0.50	0.08	0.42	0.12

Notes: The first column reports the result from the first-stage regression of our 2SLS regression using DHS-4 data. The dependent variable is a binary variable of whether fuel choice. The effective F-statistic on IVs—2011 district-wise forest cover calculated as a percent of the total geographical area of the region using satellite-based data and an indicator variable for household's agricultural land ownership—verifies that the instruments generate a plausible variation in polluting fuel for cooking. Columns (2)–(6) report results from the second-stage regressions of 2SLS regression with different dependent variables and similar specifications. All specifications contain an unreported vector of baseline demographic controls, state-by-year fixed effects, and a constant term. The Hansen's J-statistic suggests that the excluded IVs are exogenous. Parentheses contain heteroskedasticity-robust standard errors clustered by PSUs. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table A.4: Effect of cooking fuel choice on infant mortality
(IVs = Census-based forest cover and agricultural land ownership)

	1^{st} stage	2^{nd} stage				
	(1)	(2)	(3)	(4)	(5)	(6)
	Polluting fuel use	Under-five	Child	Infant	Post-neonatal	Neonatal
Polluting fuel for cooking		0.028	-0.002	0.030*	0.000	0.029**
		(0.018)	(0.005)	(0.017)	(0.010)	(0.015)
Forest cover (*census-based, 2011*)	0.051***					
	(0.012)					
Owns agricultural land	0.057***					
	(0.002)					
Observations	204697	204697	204697	204697	204697	204697
R^2	0.53	0.02	0.00	0.02	0.01	0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	292.89					
Critical value 2SLS ($\tau = 10\%$)	8.06					
Hansen J statistic		4.88	0.21	5.65	2.06	3.80
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.03	0.64	0.02	0.15	0.05

Notes: The first column reports the result from the first-stage regression of 2SLS regression using DHS-4 where the dependent variable is a binary variable for polluting fuel. The effective F-statistic on IVs—district-wise forest cover calculated as a percent of the total geographical area of the region using the 2011 Indian Census and an indicator variable for household's agricultural land ownership—verifies that the instruments generate a plausible variation in polluting fuel for cooking. Columns (2)–(6) report results from the second-stage regressions of 2SLS regression with different dependent variables and similar specifications. All specifications contain an unreported vector of baseline demographic controls, state-by-year fixed effects, and a constant term. Hansen's J-statistic suggests that the excluded IVs are exogenous. Parentheses contain heteroskedasticity-robust standard errors clustered by PSUs. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table A.5: Probit: Marginal impact of cooking fuel choice on under-five mortality (Alternative definition of mortality outcomes)

	Dependent variable: Child mortality for various age-groups		
	(1)	(2)	(3)
	Panel A. Child		
Polluting fuel for cooking	0.0043***	0.0009**	0.0009**
	(0.0004)	(0.0004)	(0.0004)
Observations	326,996	211,211	211,211
Probit log-likelihood	-10,152	-3,880	-3,879
	Panel B. Infant		
Polluting fuel for cooking	0.020***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)
Observations	344,223	221,308	221,308
Probit log-likelihood	-65,115	-33,868	-33,859
	Panel C. Post-neonatal		
Polluting fuel for cooking	0.009***	0.001*	0.001*
	(0.0007)	(0.0008)	(0.0008)
Observations	333,553	214,989	214,989
Probit log-likelihood	-29,062	-13,691	-13,690
	Panel D. Neonatal		
Polluting fuel for cooking	0.012***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
Observations	338,236	218,548	218,548
Probit log-likelihood	-46,484	-25,262	-25,253
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
Demographic controls	No	Yes	Yes
State $\times$ Year FE	No	No	Yes

Notes: Each column reports AMEs for probit regression where the key explanatory variable is polluting fuel for cooking. The dependent variable is child mortality for different age groups: child, infant, post-neonatal, and neonatal in panels A-D, respectively, defined using an alternative definition. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The demographic controls include household characteristics: place of residence, household wealth, number of household members, place where food is cooked, and type of house; mother characteristics: age and educational attainment; infant characteristics: gender and breastfeeding status; and district characteristics: age-specific adult mortality rates. The unit of observation is the child. Standard errors of the probit regressions are clustered at the PSU level, and standard errors of the AMEs in parentheses are calculated by applying the Delta method. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table A.6: Effect of cooking fuel choice on infant mortality from IV regressions (Alternative definition of mortality outcomes)

	Dependent variable: Age-specific mortality dummy			
	(1) Child	(2) Infant	(3) Post-neonatal	(4) Neonatal
	Panel A. Indoor air pollution = Polluting fuel			
Polluting fuel for cooking	-0.0005	0.041**[**]	0.013	0.031**[**]
	(0.0053)	(0.019)	(0.011)	(0.015)
	[0.0048]	[0.017]	[0.010]	[0.014]
Observations	186542	193731	188431	191319
R^2	0.002	0.02	0.01	0.01
Hansen J statistic	0.27	2.12	0.47	2.33
Degree of overidentification	1.00	1.00	1.00	1.00
p-value of Hansen J statistic	0.60	0.15	0.49	0.13
	Panel B. Indoor air pollution = Biomass fuel			
Biomass fuels for cooking	-0.0002	0.038**[**]	0.012	0.028*[**]
	(0.0050)	(0.018)	(0.011)	(0.015)
	[0.0046]	[0.016]	[0.010]	[0.014]
Observations	181851	188870	183693	186514
R^2	0.002	0.02	0.01	0.01
Hansen J statistic	0.50	2.32	0.51	2.58
Degree of overidentification	1.00	1.00	1.00	1.00
p-value of Hansen J statistic	0.48	0.13	0.47	0.11

Notes: The table checks the robustness of our IV estimation results on the effect of polluting fuel use on child, infant, post-neonatal, and neonatal mortality to alternative definitions of these mortality outcomes. The key explanatory variable is the use of polluting fuels for cooking in the top panel, while it is the use of biomass fuels for cooking in the bottom panel. Baseline instruments are used. All specifications contain an unreported constant term, baseline demographic controls, and state-by-year fixed effects. Unit of observation: child. Heteroskedasticity-robust standard errors clustered by districts are in parentheses. Robust to multiple-LATEs and heteroscedasticity standard errors (Lee, 2018) of the key regressor and statistical significance based on them are in brackets. The statistical significance of the key regressors is the same for all regressions when the standard errors are clustered by PSUs. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table A.7: Effect of cooking fuel choice on infant mortality from IV regressions (Controlling for child's breastfeeding duration)

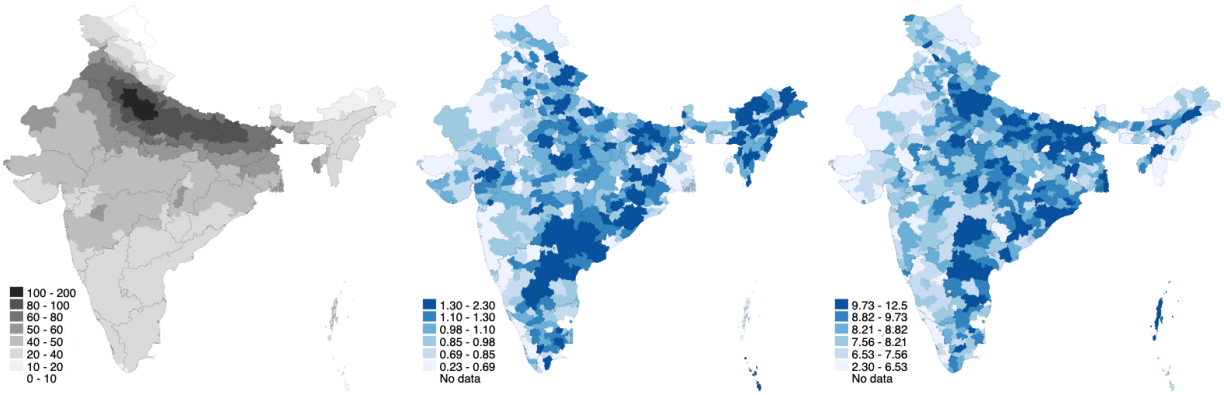
	1^{st} stage	2^{nd} stage				
	(1) Polluting fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
Polluting fuel for cooking		0.058***	0.002	0.056***	0.020**	0.036**
		(0.019)	(0.004)	(0.018)	(0.010)	(0.016)
Forest cover	0.064***					
	(0.023)					
Owns agricultural land	0.057***					
	(0.003)					
Observations	150201	150201	150201	150201	150201	150201
R^2	0.54	0.05	0.00	0.05	0.01	0.04
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	53.48					
Critical value 2SLS ($\tau = 10\%$)	17.75					
Hansen J statistic		7.64	0.24	8.70	3.39	7.52
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.01	0.63	0.00	0.07	0.01

Notes: The first column reports the result from the first-stage regression of our IV (2SLS) regression using DHS-4 where the dependent variable is a binary variable for polluting fuel. Columns (2)–(6) report results from the second-stage regressions of IV regression with different dependent variables and similar specifications. All specifications contain an unreported vector of demographic controls, state-by-year fixed effects, and a constant term. The demographic controls include the duration of the child's breastfeeding instead of breastfeeding status. Heteroskedasticity-robust standard errors clustered by districts are in parentheses. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

B Additional figures and tables

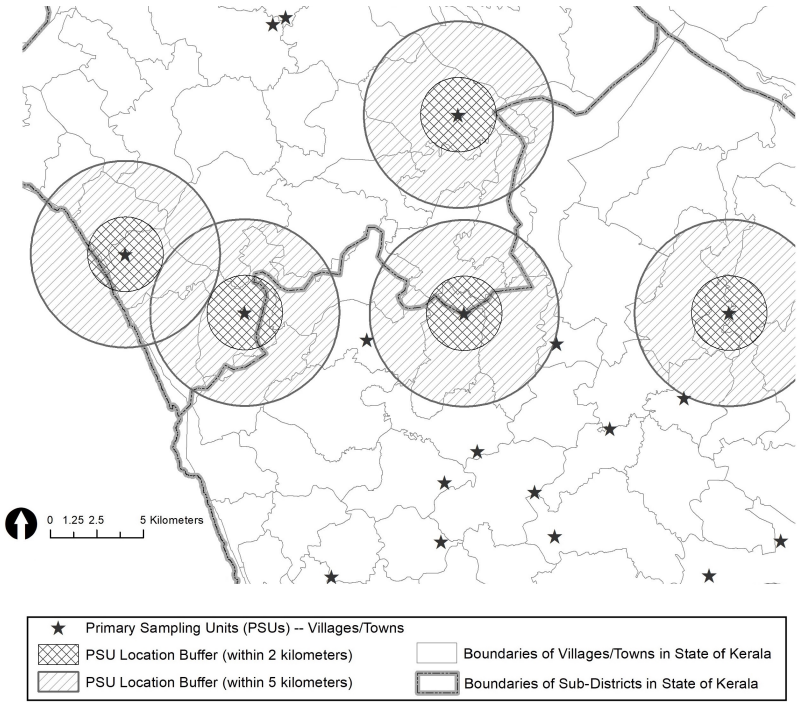
Figure B.1: India's district-wise outdoor air pollution and adult mortality rates

(a) $PM_{2.5}$ level (2014) | (b) Mortality rate (%, 15-49 years old) | (c) Mortality rate (%, 50+ years old)



Notes: Panel (a) depicts the 2014 district-wise ground-level ambient air pollution (measured by $PM_{2.5}$ level) obtained from satellite-based data. Panels (b) and (c) show district-level mortality measures for different age-groups of adults calculated using equation (2).

Figure B.2: Displacement of PSUs (villages/city blocks) in India's DHS-4 (2015–16)



Notes: The figure shows how the PSU points are displaced in the DHS-4 (2015-2016) survey based on a few PSU points in the Kerala district. To ensure that respondent confidentiality is maintained, the GPS (latitude/longitude positions) of respondent locations are randomly displaced according to the "random direction, random distance" method. The displacement is randomly carried out so that (i) urban clusters are displaced up to 2 kilometers, (ii) rural clusters are displaced up to 5 kilometers, with 1% of the rural clusters displaced up to 10 kilometers. According to the description of the DHS GPS data provided by the DHS Program, the displacement is restricted so that the points stay within the same country, state, and district areas as the undisplaced cluster. The buffer analysis on a few PSU points in Kerala district as an example suggests that identification of villages/towns and sub-districts (or *tehsils*) is questionable because 2-5-kilometer buffers intersect with boundaries of villages/towns and sub-districts.

Table B.1: Relationship between household's electricity access, use of electricity for cooking, and access to electric devices

	(1)	(2)	(3)	(4)
	Panel A. Use of electricity as a primary cooking fuel			
Access to electricity	0.0097***	0.0052***	0.0041***	-0.0004
	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Observations	164248	164248	164248	157561
R^2	0.00	0.06	0.15	0.16
	Panel B. Access to electric fan			
Access to electricity	0.7224***	0.7022***	0.6529***	0.5201***
	(0.0026)	(0.0031)	(0.0035)	(0.0037)
Observations	172059	172059	172059	164962
R^2	0.30	0.40	0.48	0.56
	Panel C. Access to TV			
Access to electricity	0.6364***	0.5166***	0.4607***	0.2466***
	(0.0024)	(0.0032)	(0.0034)	(0.0030)
Observations	172059	172059	172059	164962
R^2	0.21	0.29	0.33	0.52
State×Year FE	No	Yes	No	No
District×Year FE	No	No	Yes	Yes
Household characteristics	No	No	No	Yes

Notes: The table presents the OLS estimates on the relationship between a household's access to electricity, the use of electricity as a primary fuel for cooking, and access to electric devices, including an electric fan and television. The key explanatory variable is a dummy variable indicating whether a household has electricity. The outcome variable is a dummy variable indicating whether the household uses electricity as a primary cooking fuel (Panel A), whether a household has an electric fan (Panel B), and whether a household has a television (Panel C). Unit of observation: household. Household characteristics include all baseline household covariates, including place of residence, household wealth, number of household members, place where food is cooked, and type of house. Heteroskedasticity-robust standard errors clustered by PSUs are in parentheses. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table B.2: Summary statistics (common sample)

	Mean	SD	Min	Max	Observations
Infant mortality (% total live births)					
Under-five	0.041	0.198	0	1	221913
Child	0.003	0.051	0	1	221913
Infant	0.038	0.191	0	1	221913
Post-neonatal	0.011	0.107	0	1	221913
Neonatal	0.027	0.161	0	1	221913
Type of cooking fuel (clean)	0.339	0.473	0	1	221913
Place of residence (urban)	0.276	0.447	0	1	221913
Number of household members	6.424	2.908	1	41	221913
Child's gender (female)	0.478	0.500	0	1	221913
Child's breastfeeding status (ever)	0.633	0.482	0	1	221913
Household wealth (wealth index)					
High	0.146	0.353	0	1	221913
Middle	0.380	0.485	0	1	221913
Low	0.474	0.499	0	1	221913
Place where food is cooked					
In same room as they live in	0.362	0.481	0	1	221913
In separate kitchen inside the house	0.441	0.496	0	1	221913
In a separate building	0.100	0.300	0	1	221913
Outdoors	0.097	0.296	0	1	221913
Type of house					
Pucca	0.491	0.500	0	1	221913
Semi-pucca	0.434	0.496	0	1	221913
Kachha	0.075	0.264	0	1	221913
Mother's age (years)					
40-49	0.023	0.150	0	1	221913
$<$20	0.027	0.161	0	1	221913
20-29	0.706	0.456	0	1	221913
30-39	0.245	0.430	0	1	221913
Mother's education					
Secondary/Higher	0.554	0.497	0	1	221913
Primary	0.139	0.346	0	1	221913
No education	0.306	0.461	0	1	221913
Adult mortality rate (%, district-wise)					
15-49 years old	0.010	0.003	0.002	0.022	221913
50+ years old	0.085	0.016	0.023	0.125	221913

Notes: The table summarizes the household and individual characteristics of respondents in a common sample based on the DHS-4 survey. The unit of observation is the child, and sampling weights are applied. Neonatal = first 28 days after birth, Post-neonatal = period between the first 28 days after birth and end of the first year of life, Infant = first year of birth, and Child = period from age of one to five. Mean under-five mortality (the sum of mean child and infant mortality) exceeds any individual components that are subsets of a total number of incidents in the first five years of life. Similarly, infant mortality equals the sum of mean post-neonatal and neonatal mortality because these two preceding age groups make up the infant period. Units are % household unless otherwise specified.

Table B.3: Summary statistics of infant mortality and fuel choice (by states, common sample)

Panel A. Infant mortality (fraction)											
States	Under-five		Child		Infant		Post-neonatal		Neonatal		Obs.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Uttar Pradesh	0.063	0.243	0.004	0.066	0.058	0.235	0.018	0.133	0.041	0.197	33944
Chhattisgarh	0.054	0.226	0.005	0.068	0.049	0.216	0.012	0.108	0.038	0.190	7697
Madhya Pradesh	0.049	0.215	0.003	0.057	0.045	0.208	0.013	0.112	0.033	0.178	17542
Bihar	0.048	0.214	0.003	0.053	0.045	0.208	0.012	0.109	0.034	0.180	23155
Assam	0.047	0.211	0.003	0.051	0.044	0.205	0.015	0.121	0.029	0.168	9644
Jharkhand	0.045	0.208	0.002	0.048	0.043	0.203	0.014	0.116	0.029	0.168	11264
Rajasthan	0.043	0.203	0.003	0.054	0.040	0.197	0.013	0.113	0.027	0.163	14745
Mizoram	0.041	0.199	0.002	0.043	0.040	0.195	0.028	0.164	0.012	0.108	4565
Uttarakhand	0.039	0.195	0.003	0.05	0.037	0.189	0.012	0.109	0.025	0.156	5475
Dadra and Nagar Haveli	0.039	0.194	0.000	0.000	0.039	0.194	0.023	0.151	0.016	0.125	186
All States/UTs	0.041	0.198	0.003	0.051	0.038	0.191	0.011	0.107	0.027	0.161	221913

Panel B. Type of cooking fuel (fraction)				
States	Mean		SD	Obs.
	Polluting	Clean		
Bihar	0.883	0.117	0.322	23155
Jharkhand	0.874	0.126	0.332	11264
Meghalaya	0.861	0.139	0.346	4206
Odisha	0.846	0.154	0.361	9907
Assam	0.821	0.179	0.383	9644
Chhattisgarh	0.804	0.196	0.397	7697
West Bengal	0.788	0.212	0.409	4536
Tripura	0.759	0.241	0.428	1266
Nagaland	0.749	0.251	0.434	4108
Rajasthan	0.746	0.254	0.435	14745
All States/UTs	0.661	0.339	0.473	221913

Notes: The table summarizes the infant mortality of five different age groups (outcome variables, top panel) and the type of cooking fuel (key explanatory variable, bottom panel) by state using a common sample based on the DHS-4 survey. The sampling weights are applied. All 35 regions of India (29 states and six union territories–UTs) are considered, and we show 10 states/UTs with the highest incidence of child mortality and the highest share of households that use polluting fuel for cooking. Infant mortality and fuel choices significantly vary across regions throughout the country. In addition, five of these ten states/UTs (Chhattisgarh, Bihar, Assam, Jharkhand, and Rajasthan) are common in terms of the highest fraction of polluting fuel use and under-five mortality incidence proportion.

Table B.4: Probit: Marginal impact of cooking fuel choice on under-five mortality (common sample)

	Dependent variable: Child mortality for various age-groups		
	(1)	(2)	(3)
	Panel A. Under-five		
Polluting fuel for cooking	0.015***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)
Observations	221913	221913	221913
Probit log-likelihood	-37,989	-35,657	-35,647
	Panel B. Child		
Polluting fuel for cooking	0.0019***	0.0008**	0.0008**
	(0.0003)	(0.0004)	(0.0004)
Observations	219861	219861	219861
Probit log-likelihood	-4,100	-3,923	-3,923
	Panel C. Infant		
Polluting fuel for cooking	0.014***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)
Observations	221913	221913	221913
Probit log-likelihood	-36,117	-33,909	-33,901
	Panel D. Post-neonatal		
Polluting fuel for cooking	0.005***	0.001	0.001
	(0.001)	(0.001)	(0.001)
Observations	221518	221518	221518
Probit log-likelihood	-14,612	-13,840	-13,839
	Panel E. Neonatal		
Polluting fuel for cooking	0.009***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
Observations	221913	221913	221913
Probit log-likelihood	-26,877	-25,420	-25,411
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
Demographic controls	No	Yes	Yes
State $\times$ Year FE	No	No	Yes

Notes: Each column reports AMEs for probit regression where the key explanatory variable is polluting fuel for cooking. In each panel, the sample is based on the DHS-4 survey leveraged in Column (3) to evaluate the three different specifications on a common sample. The dependent variable is child mortality for different age groups: under-five, child, infant, post-neonatal, and neonatal in panels A through E, respectively. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The demographic controls include household characteristics: place of residence, household wealth, number of household members, place where food is cooked, and type of house; mother characteristics: age and educational attainment; infant characteristics: gender and breastfeeding status; and district characteristics: age-specific adult mortality rates. The unit of observation is the child. Standard errors of the probit regressions are clustered at the PSU level, and standard errors of the AMEs in parentheses are calculated by applying the Delta method. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table B.5: Probit: Marginal impact of cooking fuel choice on under-five mortality

	Dependent variable: Under-five mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.023***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)
Place of residence: Rural		0.004***	0.004***
		(0.001)	(0.001)
Household wealth: Middle		0.011***	0.011***
		(0.002)	(0.002)
Household wealth: Low		0.014***	0.014***
		(0.002)	(0.002)
Number of household members		-0.004***	-0.004***
		(0.000)	(0.000)
Mother's age: <20		0.019***	0.019***
		(0.003)	(0.003)
Mother's age: 20-29		-0.007***	-0.007***
		(0.002)	(0.002)
Mother's age: 30-39		-0.011***	-0.011***
		(0.002)	(0.002)
Mother's education: Primary		0.008***	0.008***
		(0.001)	(0.001)
Mother's education: No education		0.009***	0.009***
		(0.001)	(0.001)
Gender of child: Male		0.004***	0.004***
		(0.001)	(0.001)
Never breastfed		0.049***	0.049***
		(0.001)	(0.001)
Food cooked: In separate kitchen inside		-0.002	-0.002*
		(0.001)	(0.001)
Food cooked: In a separate building		-0.002	-0.002
		(0.002)	(0.002)
Food cooked: Outdoors		0.000	0.000
		(0.002)	(0.002)
House type: Semi-pucca		0.004***	0.004***
		(0.001)	(0.001)
House type: Kachha		0.005***	0.005**
		(0.002)	(0.002)
District-wise adult mortality rate (15-49 years)		0.007***	0.007***
		(0.002)	(0.002)
District-wise adult mortality rate (50+ years)		0.001***	0.001***
		(0.000)	(0.000)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State $\times$ Year FE	No	No	Yes
Observations	345,932	221,913	221,913
Probit log-likelihood	-69,858	-35,657	-35,647

Notes: Each column reports AMEs for a probit regression where the dependent variable is under-five mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors of the probit regressions are clustered at the PSU level, and standard errors of the AMEs in parentheses are calculated by applying the Delta method. Significance: $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table B.6: Probit: Marginal impact of cooking fuel choice on child mortality

	Dependent variable: Child mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.0040***	0.0008**	0.0008**
	(0.0004)	(0.0004)	(0.0004)
Place of residence: Rural		0.0004	0.0004
		(0.0004)	(0.0004)
Household wealth: Middle		0.0006	0.0006
		(0.0006)	(0.0006)
Household wealth: Low		0.0012*	0.0012*
		(0.0007)	(0.0007)
Number of household members		-0.0004***	-0.0004***
		(0.0001)	(0.0001)
Mother's age: <20		-0.0029***	-0.0029***
		(0.0011)	(0.0011)
Mother's age: 20-29		-0.0019***	-0.0019***
		(0.0005)	(0.0005)
Mother's age: 30-39		-0.0008	-0.0008
		(0.0005)	(0.0005)
Mother's education: Primary		0.0011***	0.0011***
		(0.0004)	(0.0004)
Mother's education: No education		0.0018***	0.0018***
		(0.0003)	(0.0003)
Gender of child: Male		-0.0004*	-0.0004*
		(0.0002)	(0.0002)
Never breastfed		0.0023***	0.0023***
		(0.0002)	(0.0002)
Food cooked: In separate kitchen inside		-0.0001	-0.0002
		(0.0003)	(0.0003)
Food cooked: In a separate building		-0.0003	-0.0003
		(0.0004)	(0.0004)
Food cooked: Outdoors		-0.0001	-0.0001
		(0.0004)	(0.0004)
House type: Semi-pucca		0.0002	0.0002
		(0.0003)	(0.0003)
House type: Kachha		0.0004	0.0004
		(0.0005)	(0.0005)
District-wise adult mortality rate (15-49 years)		0.0006	0.0006
		(0.0004)	(0.0004)
District-wise adult mortality rate (50+ years)		0.0000	0.0000
		(0.0001)	(0.0001)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State × Year FE	No	No	Yes
Observations	343,593	219,861	219,861
Probit log-likelihood	-10,264	-3,923	-3,923

Notes: Each column reports AMEs for probit regression where the dependent variable is child mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects (FEs) in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The number of observations is lower than that in Table B.5 because there exist five states for which state FEs perfectly explain child mortality, and thus those five state FEs are dropped because probit models cannot be estimated when the outcome variable is perfectly predicted by the regressor. The unit of observation is the child. Standard errors of the probit regressions are clustered at the PSU level, and standard errors of the AMEs in parentheses are calculated by applying the Delta method. Significance: $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table B.7: Probit: Marginal impact of cooking fuel choice on infant mortality

	Dependent variable: Infant mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.020***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)
Place of residence: Rural		0.003***	0.003**
		(0.001)	(0.001)
Household wealth: Middle		0.011***	0.010***
		(0.002)	(0.002)
Household wealth: Low		0.013***	0.013***
		(0.002)	(0.002)
Number of household members		-0.004***	-0.004***
		(0.000)	(0.000)
Mother's age: <20		0.021***	0.021***
		(0.003)	(0.003)
Mother's age: 20-29		-0.005**	-0.005**
		(0.002)	(0.002)
Mother's age: 30-39		-0.010***	-0.010***
		(0.002)	(0.002)
Mother's education: Primary		0.007***	0.007***
		(0.001)	(0.001)
Mother's education: No education		0.008***	0.008***
		(0.001)	(0.001)
Gender of child: Male		0.004***	0.004***
		(0.001)	(0.001)
Never breastfed		0.047***	0.047***
		(0.001)	(0.001)
Food cooked: In separate kitchen inside		-0.002	-0.002
		(0.001)	(0.001)
Food cooked: In a separate building		-0.002	-0.002
		(0.002)	(0.002)
Food cooked: Outdoors		0.001	0.001
		(0.002)	(0.002)
House type: Semi-pucca		0.004***	0.004***
		(0.001)	(0.001)
House type: Kachha		0.004**	0.004**
		(0.002)	(0.002)
District-wise adult mortality rate (15-49 years)		0.006***	0.007***
		(0.002)	(0.002)
District-wise adult mortality rate (50+ years)		0.001***	0.001***
		(0.000)	(0.000)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State × Year FE	No	No	Yes
Observations	345,932	221,913	221,913
Probit log-likelihood	-65,227	-33,909	-33,901

Notes: Each column reports AMEs for a probit regression where the dependent variable is infant mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects (FEs) in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors of the probit regressions are clustered at the PSU level, and standard errors of the AMEs in parentheses are calculated by applying the Delta method. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table B.8: Probit: Marginal impact of cooking fuel choice on post-neonatal mortality

	Dependent variable: Post-neonatal mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.0089***	0.0012	0.0012
	(0.0007)	(0.0008)	(0.0008)
Place of residence: Rural		0.0019***	0.0019***
		(0.0007)	(0.0007)
Household wealth: Middle		0.0053***	0.0053***
		(0.0011)	(0.0011)
Household wealth: Low		0.0071***	0.0071***
		(0.0013)	(0.0013)
Number of household members		-0.0013***	-0.0013***
		(0.0001)	(0.0001)
Mother's age: <20		0.0024	0.0025
		(0.0018)	(0.0018)
Mother's age: 20-29		-0.0041***	-0.0041***
		(0.0012)	(0.0012)
Mother's age: 30-39		-0.0043***	-0.0043***
		(0.0012)	(0.0012)
Mother's education: Primary		0.0031***	0.0031***
		(0.0007)	(0.0007)
Mother's education: No education		0.0051***	0.0051***
		(0.0006)	(0.0006)
Gender of child: Male		-0.0009*	-0.0009*
		(0.0005)	(0.0005)
Never breastfed		0.0152***	0.0152***
		(0.0006)	(0.0006)
Food cooked: In separate kitchen inside		-0.0001	-0.0002
		(0.0006)	(0.0006)
Food cooked: In a separate building		-0.0003	-0.0003
		(0.0009)	(0.0009)
Food cooked: Outdoors		0.0001	0.0001
		(0.0009)	(0.0009)
House type: Semi-pucca		0.0007	0.0007
		(0.0007)	(0.0007)
House type: Kachha		0.0007	0.0006
		(0.0010)	(0.0010)
District-wise adult mortality rate (15-49 years)		0.0018*	0.0019**
		(0.0009)	(0.0009)
District-wise adult mortality rate (50+ years)		0.0003	0.0002
		(0.0002)	(0.0002)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State × Year FE	No	No	Yes
Observations	345,932	221,518	221,518
Probit log-likelihood	-293,29	-13,840	-13,839

Notes: Each column reports AMEs for a probit regression where the dependent variable is post-neonatal mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects (FEs) in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The number of observations is slightly lower than that in Table B.5 because there exists one state for which state FE perfectly explains post-neonatal mortality, and thus that state FE is dropped because probit models cannot be estimated when the outcome variable is perfectly predicted by the regressor. The unit of observation is the child. Standard errors of the probit regressions are clustered at the PSU level, and standard errors of the AMEs in parentheses are calculated by applying the Delta method. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table B.9: Probit: Marginal impact of cooking fuel choice on neonatal mortality

	Dependent variable: Neonatal mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.011***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
Place of residence: Rural		0.002	0.001
		(0.001)	(0.001)
Household wealth: Middle		0.006***	0.005***
		(0.001)	(0.001)
Household wealth: Low		0.006***	0.006***
		(0.002)	(0.002)
Number of household members		-0.002***	-0.002***
		(0.000)	(0.000)
Mother's age: <20		0.019***	0.019***
		(0.003)	(0.003)
Mother's age: 20-29		-0.000	-0.000
		(0.002)	(0.002)
Mother's age: 30-39		-0.006***	-0.006***
		(0.002)	(0.002)
Mother's education: Primary		0.004***	0.004***
		(0.001)	(0.001)
Mother's education: No education		0.002***	0.003***
		(0.001)	(0.001)
Gender of child: Male		0.005***	0.005***
		(0.001)	(0.001)
Never breastfed		0.032***	0.032***
		(0.001)	(0.001)
Food cooked: In separate kitchen inside		-0.001	-0.001*
		(0.001)	(0.001)
Food cooked: In a separate building		-0.001	-0.001
		(0.001)	(0.001)
Food cooked: Outdoors		0.000	0.000
		(0.001)	(0.001)
House type: Semi-pucca		0.003***	0.003***
		(0.001)	(0.001)
House type: Kachha		0.004**	0.004**
		(0.002)	(0.002)
District-wise adult mortality rate (15-49 years)		0.004***	0.005***
		(0.001)	(0.001)
District-wise adult mortality rate (50+ years)		0.001***	0.001***
		(0.000)	(0.000)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State × Year FE	No	No	Yes
Observations	345,932	221,913	221,913
Probit log-likelihood	-46,777	-25,420	-25,411

Notes: Each column reports AMEs for a probit regression where the dependent variable is neonatal mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors of the probit regressions are clustered at the PSU level, and standard errors of the AMEs in parentheses are calculated by applying the Delta method. Significance: $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table B.10: OLS: Relationship between cooking fuel choice and under-five mortality

	Dependent variable: Under-five mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.019***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)
Place of residence: Rural		0.004***	0.004***
		(0.001)	(0.001)
Household wealth: Middle		0.007***	0.007***
		(0.001)	(0.001)
Household wealth: Low		0.011***	0.011***
		(0.002)	(0.002)
Number of household members		-0.004***	-0.004***
		(0.000)	(0.000)
Mother's age: <20		0.020***	0.021***
		(0.005)	(0.005)
Mother's age: 20-29		-0.011***	-0.011***
		(0.003)	(0.003)
Mother's age: 30-39		-0.016***	-0.016***
		(0.003)	(0.003)
Mother's education: Primary		0.008***	0.008***
		(0.001)	(0.001)
Mother's education: No education		0.010***	0.011***
		(0.001)	(0.001)
Gender of child: Male		0.004***	0.004***
		(0.001)	(0.001)
Never breastfed		0.055***	0.055***
		(0.001)	(0.001)
Food cooked: In separate kitchen inside		-0.002**	-0.003**
		(0.001)	(0.001)
Food cooked: In a separate building		-0.002	-0.003*
		(0.002)	(0.002)
Food cooked: Outdoors		-0.000	-0.000
		(0.002)	(0.002)
House type: Semi-pucca		0.004***	0.004***
		(0.001)	(0.001)
House type: Kachha		0.006***	0.006***
		(0.002)	(0.002)
District-wise adult mortality rate (15-49 years)		0.006***	0.007***
		(0.002)	(0.002)
District-wise adult mortality rate (50+ years)		0.001***	0.001***
		(0.000)	(0.000)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State $\times$ Year FE	No	No	Yes
Observations	345,932	221,913	221,913
R^2	0.01	0.03	0.03

Notes: Each column reports coefficient estimates for an OLS regression where the dependent variable is under-five mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors are clustered at the PSU level. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table B.11: OLS: Relationship between cooking fuel choice and child mortality

	Dependent variable: Child mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.0025***	0.0007**	0.0007**
	(0.0002)	(0.0003)	(0.0003)
Place of residence: Rural		0.0003	0.0003
		(0.0003)	(0.0003)
Household wealth: Middle		0.0001	0.0001
		(0.0003)	(0.0003)
Household wealth: Low		0.0008*	0.0008
		(0.0005)	(0.0005)
Number of household members		-0.0003***	-0.0003***
		(0.0000)	(0.0000)
Mother's age: <20		-0.0034***	-0.0034***
		(0.0011)	(0.0011)
Mother's age: 20-29		-0.0030***	-0.0030***
		(0.0010)	(0.0010)
Mother's age: 30-39		-0.0019*	-0.0019*
		(0.0011)	(0.0011)
Mother's education: Primary		0.0008**	0.0008**
		(0.0003)	(0.0003)
Mother's education: No education		0.0020***	0.0020***
		(0.0003)	(0.0003)
Gender of child: Male		-0.0004*	-0.0004*
		(0.0002)	(0.0002)
Never breastfed		0.0025***	0.0025***
		(0.0003)	(0.0003)
Food cooked: In separate kitchen inside		-0.0002	-0.0002
		(0.0003)	(0.0003)
Food cooked: In a separate building		-0.0004	-0.0004
		(0.0004)	(0.0004)
Food cooked: Outdoors		-0.0001	-0.0001
		(0.0005)	(0.0005)
House type: Semi-pucca		0.0002	0.0002
		(0.0003)	(0.0003)
House type: Kachha		0.0007	0.0007
		(0.0006)	(0.0006)
District-wise adult mortality rate (15-49 years)		0.0007	0.0007
		(0.0005)	(0.0005)
District-wise adult mortality rate (50+ years)		0.0000	0.0000
		(0.0001)	(0.0001)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State × Year FE	No	No	Yes
Observations	345,932	221,913	221,913
R^2	0.003	0.002	0.002

Notes: Each column reports coefficient estimates for an OLS regression where the dependent variable is child mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors are clustered at the PSU level. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table B.12: OLS: Relationship between cooking fuel choice and infant mortality

	Dependent variable: Infant mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.016***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)
Place of residence: Rural		0.003***	0.003***
		(0.001)	(0.001)
Household wealth: Middle		0.007***	0.007***
		(0.001)	(0.001)
Household wealth: Low		0.010***	0.010***
		(0.002)	(0.002)
Number of household members		-0.003***	-0.003***
		(0.000)	(0.000)
Mother's age: <20		0.024***	0.024***
		(0.005)	(0.005)
Mother's age: 20-29		-0.008**	-0.008**
		(0.003)	(0.003)
Mother's age: 30-39		-0.014***	-0.014***
		(0.003)	(0.003)
Mother's education: Primary		0.007***	0.007***
		(0.001)	(0.001)
Mother's education: No education		0.008***	0.009***
		(0.001)	(0.001)
Gender of child: Male		0.004***	0.004***
		(0.001)	(0.001)
Never breastfed		0.052***	0.052***
		(0.001)	(0.001)
Food cooked: In separate kitchen inside		-0.002**	-0.002**
		(0.001)	(0.001)
Food cooked: In a separate building		-0.002	-0.002
		(0.002)	(0.002)
Food cooked: Outdoors		-0.000	-0.000
		(0.002)	(0.002)
House type: Semi-pucca		0.004***	0.004***
		(0.001)	(0.001)
House type: Kachha		0.005**	0.005**
		(0.002)	(0.002)
District-wise adult mortality rate (15-49 years)		0.006***	0.006***
		(0.002)	(0.002)
District-wise adult mortality rate (50+ years)		0.001***	0.001***
		(0.000)	(0.000)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State × Year FE	No	No	Yes
Observations	345,932	221,913	221,913
R^2	0.01	0.03	0.03

Notes: Each column reports coefficient estimates for an OLS regression where the dependent variable is infant mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors are clustered at the PSU level. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table B.13: OLS: Relationship between cooking fuel choice and post-neonatal mortality

	Dependent variable: Post-neonatal mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.0068***	0.0012*	0.0012*
	(0.0005)	(0.0007)	(0.0007)
Place of residence: Rural		0.0019***	0.0019***
		(0.0006)	(0.0006)
Household wealth: Middle		0.0034***	0.0033***
		(0.0007)	(0.0007)
Household wealth: Low		0.0054***	0.0054***
		(0.0011)	(0.0011)
Number of household members		-0.0011***	-0.0011***
		(0.0001)	(0.0001)
Mother's age: <20		0.0005	0.0006
		(0.0027)	(0.0027)
Mother's age: 20-29		-0.0071***	-0.0071***
		(0.0021)	(0.0021)
Mother's age: 30-39		-0.0073***	-0.0073***
		(0.0021)	(0.0021)
Mother's education: Primary		0.0029***	0.0029***
		(0.0008)	(0.0008)
Mother's education: No education		0.0056***	0.0056***
		(0.0007)	(0.0007)
Gender of child: Male		-0.0010**	-0.0010**
		(0.0005)	(0.0005)
Never breastfed		0.0168***	0.0168***
		(0.0006)	(0.0006)
Food cooked: In separate kitchen inside		-0.0003	-0.0004
		(0.0006)	(0.0006)
Food cooked: In a separate building		-0.0004	-0.0004
		(0.0009)	(0.0009)
Food cooked: Outdoors		-0.0002	-0.0002
		(0.0010)	(0.0010)
House type: Semi-pucca		0.0007	0.0007
		(0.0007)	(0.0007)
House type: Kachha		0.0010	0.0010
		(0.0012)	(0.0012)
District-wise adult mortality rate (15-49 years)		0.0018*	0.0018*
		(0.0010)	(0.0010)
District-wise adult mortality rate (50+ years)		0.0003	0.0003
		(0.0002)	(0.0002)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State $\times$ Year FE	No	No	Yes
Observations	345,932	221,913	221,913
R^2	0.01	0.01	0.01

Notes: Each column reports coefficient estimates for an OLS regression where the dependent variable is post-neonatal mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors are clustered at the PSU level. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table B.14: OLS: Relationship between cooking fuel choice and neonatal mortality

	Dependent variable: Neonatal mortality		
	(1)	(2)	(3)
Polluting fuel for cooking	0.009***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)
Place of residence: Rural		0.002	0.001
		(0.001)	(0.001)
Household wealth: Middle		0.004***	0.004***
		(0.001)	(0.001)
Household wealth: Low		0.005***	0.005***
		(0.002)	(0.002)
Number of household members		-0.002***	-0.002***
		(0.000)	(0.000)
Mother's age: <20		0.023***	0.024***
		(0.004)	(0.004)
Mother's age: 20-29		-0.001	-0.001
		(0.003)	(0.003)
Mother's age: 30-39		-0.007***	-0.007***
		(0.003)	(0.003)
Mother's education: Primary		0.004***	0.004***
		(0.001)	(0.001)
Mother's education: No education		0.003***	0.003***
		(0.001)	(0.001)
Gender of child: Male		0.005***	0.005***
		(0.001)	(0.001)
Never breastfed		0.035***	0.035***
		(0.001)	(0.001)
Food cooked: In separate kitchen inside		-0.002**	-0.002**
		(0.001)	(0.001)
Food cooked: In a separate building		-0.002	-0.002
		(0.001)	(0.001)
Food cooked: Outdoors		0.000	0.000
		(0.002)	(0.002)
House type: Semi-pucca		0.003***	0.003***
		(0.001)	(0.001)
House type: Kachha		0.004**	0.004**
		(0.002)	(0.002)
District-wise adult mortality rate (15-49 years)		0.004***	0.004***
		(0.001)	(0.001)
District-wise adult mortality rate (50+ years)		0.001***	0.001**
		(0.000)	(0.000)
Year FE	Yes	Yes	No
State FE	Yes	Yes	No
State $\times$ Year FE	No	No	Yes
Observations	345,932	221,913	221,913
R^2	0.00	0.02	0.02

Notes: Each column reports coefficient estimates for an OLS regression where the dependent variable is neonatal mortality and the key explanatory variable is polluting fuel for cooking. The year fixed effects in Columns (2) and (3) include dummies for two years of interviews (2015 and 2016). The state fixed effects include dummies for 36 states. The unit of observation is the child. Standard errors are clustered at the PSU level. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

Table B.15: Effect of IAP on infant mortality from IV regressions
(IVs = Forest cover and agricultural land ownership)

	1^{st} stage	2^{nd} stage				
	(1) Polluting/Biomass fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
	Panel A. Indoor air pollution = Polluting fuel					
Polluting fuel for cooking		0.040**	-0.001	0.041**	0.011	0.030**
		(0.020)	(0.005)	(0.019)	(0.011)	(0.015)
Place of residence: Rural	0.225***	-0.004	0.001	-0.005	0.000	-0.005
	(0.008)	(0.005)	(0.001)	(0.005)	(0.003)	(0.004)
Household wealth: Middle	0.350***	-0.005	0.001	-0.005	-0.000	-0.005
	(0.008)	(0.007)	(0.002)	(0.007)	(0.004)	(0.005)
Household wealth: Low	0.643***	-0.011	0.002	-0.012	-0.002	-0.010
	(0.009)	(0.013)	(0.003)	(0.012)	(0.007)	(0.010)
Number of household members	0.014***	-0.004***	-0.000***	-0.004***	-0.001***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's age: <20	0.051***	0.021***	-0.003**	0.024***	-0.001	0.025***
	(0.007)	(0.005)	(0.001)	(0.005)	(0.003)	(0.004)
Mother's age: 20-29	0.032***	-0.012***	-0.003**	-0.009**	-0.007***	-0.002
	(0.005)	(0.004)	(0.001)	(0.004)	(0.003)	(0.003)
Mother's age: 30-39	0.002	-0.017***	-0.002*	-0.015***	-0.007***	-0.007***
	(0.005)	(0.004)	(0.001)	(0.004)	(0.002)	(0.003)
Mother's education: Primary	0.040***	0.007***	0.001	0.006***	0.002***	0.004***
	(0.004)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
Mother's education: No education	0.045***	0.009***	0.002***	0.007***	0.005***	0.002
	(0.003)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
Gender of child: Male	0.000	0.003***	-0.000*	0.004***	-0.001*	0.005***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Never breastfed	-0.015***	0.056***	0.003***	0.053***	0.017***	0.037***
	(0.002)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
Food cooked: In separate kitchen inside	-0.031***	-0.001	-0.000	-0.001	-0.000	-0.001
	(0.004)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Food cooked: In a separate building	0.062***	-0.005**	-0.000	-0.005**	-0.002	-0.004**
	(0.005)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Food cooked: Outdoors	0.166***	-0.006	0.000	-0.006	-0.002	-0.004
	(0.008)	(0.004)	(0.001)	(0.004)	(0.002)	(0.003)
House type: Semi-pucca	0.041***	0.003**	0.000	0.003*	0.000	0.002**
	(0.005)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
House type: Kachha	0.045***	0.006**	0.001*	0.005**	0.001	0.004**
	(0.006)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)
District-wise adult mortality rate (15-49 years)	0.003	0.007***	0.001	0.007***	0.002	0.005***
	(0.011)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
District-wise adult mortality rate (50+ years)	0.004	0.001	0.000	0.001	0.000	0.001**
	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Forest cover	0.065***					
	(0.022)					
Owns agricultural land	0.054***					
	(0.003)					
Observations	194254	194254	194254	194254	194254	194254
R^2	0.53	0.02	0.00	0.02	0.01	0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	54.12					
Critical value 2SLS ($\tau = 10\%$)	17.41					
Hansen J statistic		1.70	0.33	2.18	0.40	2.29
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.19	0.57	0.14	0.53	0.13

Table B.15: (Continued)

	1^{st} stage	2^{nd} stage				
	(1) Polluting/Biomass fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
	Panel B. Indoor air pollution = Biomass fuel					
Biomass fuels for cooking		0.037*	-0.001	0.037**	0.011	0.026*
		(0.019)	(0.005)	(0.018)	(0.010)	(0.014)
Place of residence: Rural	0.245***	-0.004	0.001	-0.004	0.000	-0.004
	(0.008)	(0.005)	(0.001)	(0.005)	(0.003)	(0.004)
Household wealth: Middle	0.330***	-0.003	0.001	-0.004	-0.000	-0.003
	(0.008)	(0.006)	(0.002)	(0.006)	(0.003)	(0.005)
Household wealth: Low	0.632***	-0.009	0.002	-0.010	-0.002	-0.009
	(0.009)	(0.012)	(0.003)	(0.011)	(0.007)	(0.009)
Number of household members	0.015***	-0.004***	-0.000***	-0.004***	-0.001***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's age: <20	0.049***	0.021***	-0.003**	0.024***	-0.001	0.025***
	(0.007)	(0.005)	(0.001)	(0.005)	(0.003)	(0.004)
Mother's age: 20-29	0.031***	-0.013***	-0.003**	-0.010**	-0.007***	-0.002
	(0.005)	(0.004)	(0.001)	(0.004)	(0.003)	(0.003)
Mother's age: 30-39	0.002	-0.017***	-0.002	-0.015***	-0.007***	-0.008***
	(0.005)	(0.004)	(0.001)	(0.004)	(0.003)	(0.003)
Mother's education: Primary	0.039***	0.007***	0.001	0.006***	0.002***	0.004***
	(0.004)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
Mother's education: No education	0.045***	0.009***	0.002***	0.007***	0.005***	0.002*
	(0.003)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
Gender of child: Male	-0.000	0.003***	-0.000*	0.004***	-0.001*	0.005***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Never breastfed	-0.015***	0.056***	0.002***	0.053***	0.017***	0.036***
	(0.002)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
Food cooked: In separate kitchen inside	-0.027***	-0.002	-0.000	-0.002	-0.000	-0.001
	(0.004)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Food cooked: In a separate building	0.069***	-0.005**	-0.000	-0.005**	-0.002	-0.003*
	(0.006)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Food cooked: Outdoors	0.172***	-0.006	-0.000	-0.006	-0.002	-0.004
	(0.008)	(0.004)	(0.001)	(0.004)	(0.002)	(0.003)
House type: Semi-pucca	0.047***	0.003**	0.000	0.003*	0.000	0.003**
	(0.005)	(0.002)	(0.000)	(0.002)	(0.001)	(0.001)
House type: Kachha	0.051***	0.007**	0.001*	0.005**	0.001	0.004*
	(0.007)	(0.003)	(0.001)	(0.003)	(0.001)	(0.002)
District-wise adult mortality rate (15-49 years)	0.005	0.007***	0.001	0.007***	0.002	0.005***
	(0.011)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
District-wise adult mortality rate (50+ years)	0.003	0.001	0.000	0.001	0.000	0.001**
	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Forest cover	0.067***					
	(0.022)					
Owns agricultural land	0.058***					
	(0.003)					
Observations	189384	189384	189384	189384	189384	189384
R^2	0.55	0.02	0.00	0.02	0.01	0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	57.82					
Critical value 2SLS ($\tau = 10\%$)	17.74					
Hansen J statistic		1.74	0.58	2.40	0.44	2.54
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.19	0.45	0.12	0.51	0.11

Notes: The first column reports the result from the first-stage regression of our IV (2SLS) regression using DHS-4 data. The dependent variable is a binary variable of whether fuel choice: polluting fuel (top panel) and biomass fuel (bottom panel). The effective F-statistic on IVs—2013 district-wise forest cover calculated as a percentage of forested area in the total geographical area of the region using satellite-based data and an indicator variable for household's agricultural land ownership—verifies that the instruments generate a plausible variation in polluting (top panel) and biomass (bottom panel) fuels for cooking. Columns (2)–(6) report results from the estimation of equation (1) using IV regression with different dependent variables and similar specifications. The number of observations decreases in the bottom panel as we drop non-biomass polluting fuels from the sample to contrast the effect of using polluting biomass fuels to that of using polluting fuels. The Hansen's J-statistic suggests that the excluded IVs are exogenous. All specifications contain an unreported vector of baseline demographic controls, state-by-year fixed effects, and a constant term. Heteroskedasticity-robust standard errors clustered by districts are in parentheses. Robust to multiple-LATEs and heteroscedasticity standard errors (Lee, 2018) of the key regressor and statistical significance based on them are in brackets. The statistical significance of the key regressors is the same for all regressions when the standard errors are clustered by PSUs. Significance: $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table B.16: Relevance of instruments with variables other than cooking fuel choice

	Child outcomes	Maternal health	Mother's health
	(1)	(2)	(3)
	Ever breastfed=1	# Antenatal visits	# Injection
Forest cover	-0.0076	-0.5561*	-0.2536
	(0.0193)	(0.2915)	(0.2481)
Owns agricultural land	-0.0003	-0.0421	0.0570
	(0.0030)	(0.0305)	(0.0385)
Observations	205680	203827	204477
R^2	0.06	0.27	0.03
Montiel Olea-Pflueger weak IV test			
Effective F-statistic ($\alpha = 5\%$)	0.13	1.58	1.34
Critical value 2SLS ($\tau = 10\%$)	15.54	20.83	15.22

Notes: The table presents the results from testing the relevance of our baseline instruments (forest cover and agricultural land ownership) with variables other than our key explanatory variable of cooking fuel choice. The dependent variables are stated at the top. All specifications contain unreported constant term and baseline demographic controls. The state-by-year fixed effects are also controlled. Unit of observation: child. Heteroskedasticity-robust standard errors clustered by districts are in parentheses. Significance: $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table B.17: IV: Effect of other household and mother characteristics on under-five mortality

	Panel A		Panel B	
	1^{st} stage	2^{nd} stage	1^{st} stage	2^{nd} stage
	(1) Household size	(2) Under-five	(3) Mother's education	(4) Under-five
Household size		-0.0009		
		(0.0012)		
Mother's educational attainment				0.0133
				(0.0086)
Forest cover	-0.3727**		0.0593	
	(0.1498)		(0.0627)	
Owns agricultural land	0.9461***		0.1325***	
	(0.0325)		(0.0090)	
Observations	205680	205680	205680	205680
R^2	0.10	0.02	0.32	0.02
Montiel Olea-Pflueger weak IV test				
Effective F-statistic ($\alpha = 5\%$)	96.88		22.94	
Critical value 2SLS ($\tau = 10\%$)	18.85		18.57	
Hansen J statistic		0.00		0.32
Degree of overidentification		1.00		1.00
p-value of Hansen J statistic		0.97		0.57

Notes: The table presents the effects of household size and mother's education on under-five mortality instrumented by our baseline instruments including forest cover and agricultural land ownership status as a "placebo" test of exclusion restriction of our instruments. All specifications contain unreported constant term and baseline demographic controls. The state-by-year fixed effects are also controlled. Unit of observation: child. Heteroskedasticity-robust standard errors clustered by districts are in parentheses. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

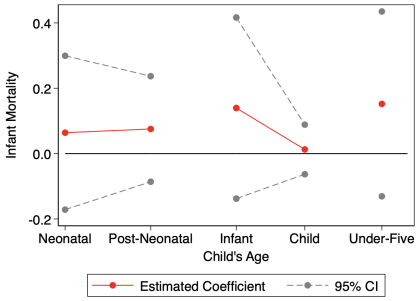
C Additional household size heterogeneity results

In this appendix, we check whether household size classification affects our results of heterogeneous LATE by household size. We do so by choosing other sets of subpopulations of households with fewer than twelve members. Figure C.1 and Table C.1 show the heterogeneous treatment effects of IAP on infant mortality when the sample of households with twelve or fewer members is classified into four groups (3-by-3) in terms of household size. The result reveals that the effect on neonatal mortality rate is 0.098 for households with 4-6 members (Column (5) in panel B). The effects on under-five and post-neonatal mortality are positive and statistically significant at least at the 5% level; however, the excluded IVs are not exogenous in these two particular regressions according to the Hansen test. The effects of IAP on infant mortality for all age groups are not significantly different from zero for households with members other than 4-6.

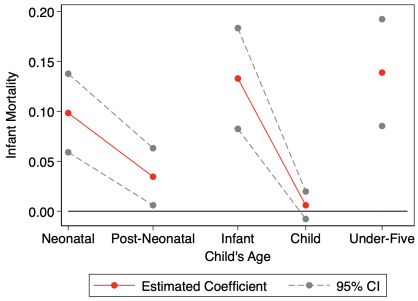
We then further reclassify the sample of households with twelve or fewer members into six groups (2-by-2) in terms of household size (Figure C.2 and Table C.2). The results suggest the effect on neonatal mortality rate is 0.085 for households with 5-6 members (Column (5) of panel B). For these households with 5-6 members, the impact of solid fuel use on under-five mortality is statistically significant at the 1% level, but the excluded IVs are not exogenous in this regression. The effects of IAP on infant mortality for all age groups are essentially zero for households with 1-2 and those with more than 7 members. Child mortality is not affected by IAP in any of the households with different pairs of family members. The local average effects of IAP on under-five, post-neonatal, and neonatal mortality are positive and significantly different from zero for households with three and four members at least at the 5% level. However, the Hansen test values imply that the instruments are not valid in three regressions for households with three and four members since the null hypothesis is rejected at the 5%, 1%, and 10% levels, respectively.

Figure C.1: Heterogeneous mortality effects of IAP by child's age and household size (3-by-3 groups)

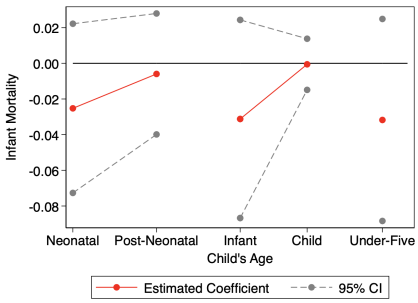
(a) Household size = [1-3], N = 15,403



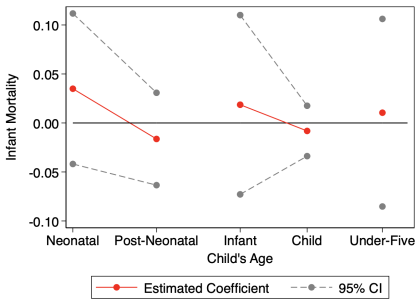
(b) Household size = [4-6], N = 103,517



(c) Household size = [7-9], N = 51,110



(d) Household size = [10-12], N = 16,772



Notes: The figure presents heterogeneous treatment effects of polluting fuel use on infant mortality by both child's age and household size using four distinct subpopulations of households with fewer than twelve members covered in DHS-4 data. Panels (a)–(d) consider each of the four subsamples in an order of 1-3 to 10-12 members. Each panel reports results from the estimation of equation (1) using IV regression with different dependent variables and similar specifications. The outcome variable in each IV regression is a binary variable of infant mortality for each of the five different age groups, and the endogenous regressor (i.e., the outcome variable in the first-stage regression) is whether fuel choice: polluting fuel. The forest cover and agricultural land ownership status are used as instruments, and first-stage coefficient estimates on the IVs are both positive and statistically significant at least at the 5% level except for the forest cover in the first-stage regression for households with 10-12 members. The effective F-statistic on IVs verifies that the instruments generate a plausible variation in polluting fuel for cooking in all regressions. The Hansen's J-statistic suggests that the excluded IVs are not exogenous in under-five, infant, and post-neonatal mortality regressions for households with 4-6 members since a rejection of the null hypothesis of the Hansen's J test is encountered at the 5% level. All specifications contain a vector of baseline demographic controls, state-by-year fixed effects, and a constant term. Heteroskedasticity-robust standard errors are clustered by districts.

Table C.1: Heterogeneous mortality effects of IAP by child's age and household size (3-by-3 groups)

	1^{st} stage	2^{nd} stage				
	(1) Polluting fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
	Panel A. Number of household members $= [1\text{-}3]$					
Polluting fuel for cooking		0.152	0.013	0.139	0.075	0.064
		(0.144)	(0.039)	(0.141)	(0.082)	(0.120)
Forest cover	0.094***					
	(0.031)					
Owns agricultural land	0.027***					
	(0.006)					
Observations	15403	15403	15403	15403	15403	15403
R^2	0.59	0.11	0.00	0.11	0.01	0.09
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	12.03					
Critical value 2SLS ($\tau = 10\%$)	12.95					
Critical value 2SLS ($\tau = 20\%$)	8.86					
Hansen J statistic		2.53	0.51	3.42	1.20	2.04
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.11	0.47	0.06	0.27	0.15
	Panel B. Number of household members $= [4\text{-}6]$					
Polluting fuel for cooking		0.139***	0.006	0.133***	0.035**	0.098***
		(0.027)	(0.007)	(0.026)	(0.015)	(0.020)
Forest cover	0.076***					
	(0.022)					
Owns agricultural land	0.050***					
	(0.003)					
Observations	103517	103517	103517	103517	103517	103517
R^2	0.55	-0.03	0.00	-0.03	-0.00	-0.02
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	43.38					
Critical value 2SLS ($\tau = 10\%$)	16.25					
Hansen J statistic		5.47	0.18	5.53	5.06	2.60
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.02	0.67	0.02	0.02	0.11
	Panel C. Number of household members $= [7\text{-}9]$					
Polluting fuel for cooking		-0.032	-0.001	-0.031	-0.006	-0.025
		(0.029)	(0.007)	(0.028)	(0.017)	(0.024)
Forest cover	0.056**					
	(0.025)					
Owns agricultural land	0.056***					
	(0.004)					
Observations	51110	51110	51110	51110	51110	51110
R^2	0.52	0.01	0.00	0.01	0.01	0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	44.73					
Critical value 2SLS ($\tau = 10\%$)	13.92					
Hansen J statistic		0.16	0.23	0.08	2.56	0.70
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.69	0.63	0.78	0.11	0.40

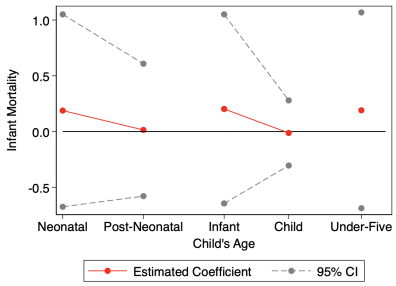
Table C.1: (Continued)

	1^{st} stage	2^{nd} stage				
	(1) Polluting fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
	Panel D. Number of household members $= [10\text{-}12]$					
Polluting fuel for cooking		0.010	-0.008	0.019	-0.016	0.035
		(0.049)	(0.013)	(0.047)	(0.024)	(0.039)
Forest cover	0.005					
	(0.039)					
Owns agricultural land	0.071***					
	(0.008)					
Observations	16772	16772	16772	16772	16772	16772
R^2	0.52	0.01	-0.00	0.01	0.00	0.00
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	25.64					
Critical value 2SLS ($\tau = 10\%$)	9.23					
Hansen J statistic		0.60	0.42	0.36	0.00	0.50
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.44	0.52	0.55	0.98	0.48

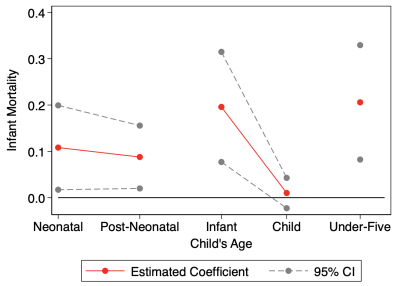
Notes: The table presents heterogeneous treatment effects of IAP on infant mortality by both child's age and household size using four subpopulations of households with fewer than twelve members covered in DHS-4 data. Panel A–D considers each of the four subsamples in an order of 1-3 to 10-12 members. The first column provides results from the first-stage regressions of the IV (2SLS) regressions, where the dependent variable is a binary variable of whether fuel choice: polluting fuel. Columns (2)–(6) report results from the estimation of equation (1) using IV regression with different dependent variables and similar specifications. The outcome variable in an IV regression is a binary variable of infant mortality for each of the five different age groups. All specifications contain an unreported vector of baseline demographic controls, state-by-year fixed effects, and a constant term. Heteroskedasticity-robust standard errors clustered by districts are in parentheses. Significance: $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Figure C.2: Heterogeneous mortality effects of IAP by child's age and household size (2-by-2 groups)

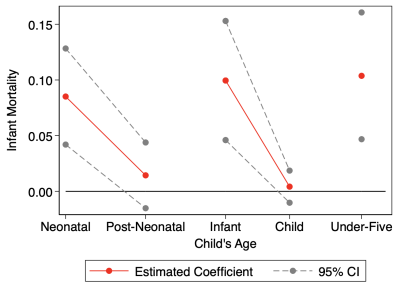
(a) Household size = [1-2], N = 1,217



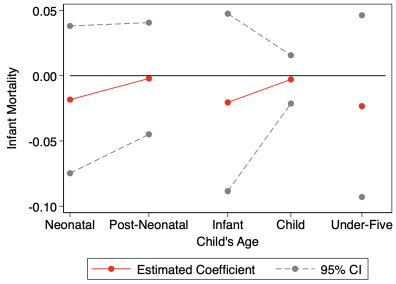
(b) Household size = [3-4], N = 47,559



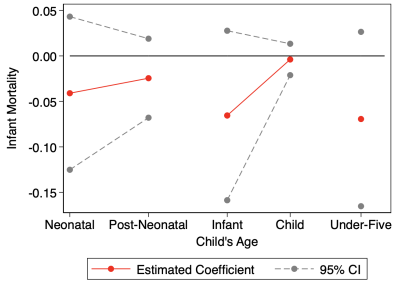
(c) Household size = [5-6], N = 70,143



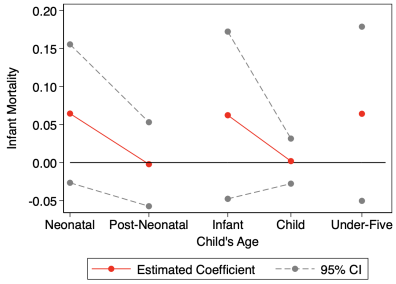
(d) Household size = [7-8], N = 39,604



(e) Household size = [9-10], N = 19,645



(f) Household size = [11-12], N = 8,633



Notes: The figure presents heterogeneous treatment effects of polluting fuel use on infant mortality by both child's age and household size using six distinct subpopulations of households with fewer than twelve members covered in DHS-4 data. Panels (a)–(f) consider each of the six subsamples in an order of 1-2 to 11-12 members. Each panel reports results from the estimation of equation (1) using IV regression with different dependent variables and similar specifications. The outcome variable in each IV regression is a binary variable of infant mortality for each of the five different age groups, and the endogenous regressor or outcome variable in the first-stage regression is whether fuel choice: polluting fuel. The forest cover and agricultural land ownership status are used as instruments, and first-stage coefficient estimates on the IVs are both positive and statistically significant at least at the 5% level except for the forest cover in the first-stage regression for households with 9-10 and 11-12 members. The effective F-statistic on IVs verifies that the instruments generate a plausible variation in polluting fuel for cooking, except for a subpopulation of households with 1-2 members. The Hansen's J-statistic suggests that the excluded IVs are not exogenous in under-five, infant, and post-neonatal mortality regressions with the statistically significant coefficient on polluting fuel use since a rejection of the null hypothesis of the Hansen's J test is encountered at least at the 10% level. All specifications contain a vector of baseline demographic controls, state-by-year fixed effects, and a constant term. Heteroskedasticity-robust standard errors are clustered by districts.

Table C.2: Heterogeneous mortality effects of IAP by child's age and household size (2-by-2 groups)

	1^{st} stage	2^{nd} stage				
	(1) Polluting fuel use	(2) Under-five	(3) Child	(4) Infant	(5) Post-neonatal	(6) Neonatal
	Panel A. Number of household members = [1-2]					
Polluting fuel for cooking		0.191	-0.012	0.203	0.015	0.188
		(0.446)	(0.149)	(0.431)	(0.302)	(0.439)
Forest cover	0.195**					
	(0.078)					
Owns agricultural land	0.048**					
	(0.022)					
Observations	1217	1217	1217	1217	1217	1217
R^2	0.52	0.25	0.02	0.23	0.04	0.16
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	5.34					
Critical value 2SLS ($\tau = 10\%$)	5.91					
Critical value 2SLS ($\tau = 20\%$)	4.61					
Hansen J statistic		0.09	0.02	0.06	0.80	0.13
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.77	0.89	0.81	0.37	0.72
	Panel B. Number of household members = [3-4]					
Polluting fuel for cooking		0.206***	0.010	0.196***	0.088**	0.108**
		(0.063)	(0.017)	(0.061)	(0.035)	(0.046)
Forest cover	0.079***					
	(0.024)					
Owns agricultural land	0.035***					
	(0.004)					
Observations	47559	47559	47559	47559	47559	47559
R^2	0.58	-0.03	0.00	-0.03	-0.03	-0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	22.24					
Critical value 2SLS ($\tau = 10\%$)	14.38					
Hansen J statistic		5.24	0.95	6.93	6.89	2.92
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.02	0.33	0.01	0.01	0.09
	Panel C. Number of household members = [5-6]					
Polluting fuel for cooking		0.104***	0.004	0.100***	0.014	0.085***
		(0.029)	(0.007)	(0.027)	(0.015)	(0.022)
Forest cover	0.076***					
	(0.022)					
Owns agricultural land	0.055***					
	(0.004)					
Observations	70143	70143	70143	70143	70143	70143
R^2	0.54	-0.01	0.00	-0.01	0.00	-0.02
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	48.66					
Critical value 2SLS ($\tau = 10\%$)	14.96					
Hansen J statistic		3.10	0.49	2.88	1.96	1.45
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.08	0.48	0.09	0.16	0.23

Table C.2: (Continued)

	1^{st} stage	2^{nd} stage				
	(1)	(2)	(3)	(4)	(5)	(6)
	Polluting fuel use	Under-five	Child	Infant	Post-neonatal	Neonatal
	Panel D. Number of household members = [7-8]					
Polluting fuel for cooking		-0.023	-0.003	-0.020	-0.002	-0.018
		(0.035)	(0.009)	(0.035)	(0.022)	(0.029)
Forest cover	0.059**					
	(0.026)					
Owns agricultural land	0.053***					
	(0.005)					
Observations	39604	39604	39604	39604	39604	39604
R^2	0.51	0.01	0.00	0.01	0.01	0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	35.72					
Critical value 2SLS ($\tau = 10\%$)	13.68					
Hansen J statistic		0.00	0.32	0.03	0.95	0.85
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		1.00	0.57	0.87	0.33	0.36
	Panel E. Number of household members = [9-10]					
Polluting fuel for cooking		-0.069	-0.004	-0.065	-0.024	-0.041
		(0.049)	(0.009)	(0.047)	(0.022)	(0.043)
Forest cover	0.039					
	(0.035)					
Owns agricultural land	0.064***					
	(0.008)					
Observations	19645	19645	19645	19645	19645	19645
R^2	0.53	0.00	-0.00	0.00	0.00	0.00
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	26.53					
Critical value 2SLS ($\tau = 10\%$)	9.30					
Hansen J statistic		0.00	0.17	0.00	1.58	0.62
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.96	0.68	0.97	0.21	0.43
	Panel F. Number of household members = [11-12]					
Polluting fuel for cooking		0.064	0.002	0.062	-0.002	0.064
		(0.058)	(0.015)	(0.056)	(0.028)	(0.046)
Forest cover	-0.023					
	(0.046)					
Owns agricultural land	0.084***					
	(0.012)					
Observations	8633	8633	8633	8633	8633	8633
R^2	0.51	-0.00	0.00	-0.00	0.00	-0.01
Montiel Olea-Pflueger weak IV test						
Effective F-statistic ($\alpha = 5\%$)	21.63					
Critical value 2SLS ($\tau = 10\%$)	5.74					
Hansen J statistic		0.06	0.19	0.02	0.27	0.01
Degree of overidentification		1.00	1.00	1.00	1.00	1.00
p-value of Hansen J statistic		0.81	0.66	0.90	0.60	0.91

Notes: The table presents heterogeneous treatment effects of IAP on infant mortality by both child's age and household size using six subpopulations of households with fewer than twelve members covered in DHS-4 data. Panel A–F considers each of the six subsamples in an order of 1-2 to 11-12 members. The first column provides results from the first-stage regressions of the IV (2SLS) regressions, where the dependent variable is a binary variable of whether fuel choice: polluting fuel. Columns (2)–(6) report results from the estimation of equation (1) using IV regression with different dependent variables and similar specifications. The outcome variable in an IV regression is a binary variable of infant mortality for each of the five different age groups. All specifications contain an unreported vector of baseline demographic controls, state-by-year fixed effects, and a constant term. Heteroskedasticity-robust standard errors clustered by districts are in parentheses. Significance: *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.