APPENDIX A: ATTRITION IN THE KHDS SURVEY

In comparison with most longitudinal surveys, attrition in the KHDS is low; in 93% of baseline households, at least one household member was reinterviewed in 2004. Among all surviving members of the baseline households, about 82% were reinterviewed in 2004. Among the target group in this article, non-orphaned children aged 6–15 at baseline, attrition is slightly higher; 79% of the nondeceased children were reinterviewed in 2004. About half these were found through the extensive efforts to locate people who had physically moved to a new village. Any area (town, district, or even neighboring country) to which children were reported to have moved was visited by a specific search and survey team to minimize nonrandom attrition. Despite these efforts and the low attrition rate, it is still possible that attrition may cause problems for our analysis.[1] Attrition linked to unobservables cannot be addressed in this data set; as in Fitzgerald et al. (1998), we focus on attrition linked to observables.

These results suggest some potential problem of nonrandom attrition. While it is plausible that the descriptive statistics in this article are affected to some extent, our focus is on the regression analysis aiming to establish causal links between orphanhood and outcomes at adulthood. First, any attrition linked to particular clusters is not likely to be problematic because all the regression analyses control for cluster fixed effects so that all effects are identified using within-cluster variation, while for every cluster in the sample, at least 55% of the cluster sample was reinterviewed. Second, we use the approach suggested by Fitzgerald et al. (1998) and use the inverse of the probabilities of an individual to remain in the sample by 2004, as predicted by baseline characteristics and with the probabilities normalized to keep the sample size at 718 observations. These probabilities are based on the same specification as in Appendix Table A1 but using the logit model with cluster dummy variables rather than the conditional fixed-effects model; otherwise, some sampled individuals could not be assigned a probability of remaining in the sample. The inverse probabilities are then used as weights. The results of the specification equivalent to Table 5 in the main text are presented in Appendix Table A2, and (not surprisingly given the discussion in Deaton 1997) the results are virtually identical to Table 5 in the main text. Third, it could be that there are behavioral differences for those groups we specifically had difficulty tracking. In line with Becketti et al. (1998), a simple approach to explore this is to interact the variables related to orphanhood with those observable traits correlated with attrition. Interaction terms related to consumption per capita and whether the child was living with the father at baseline were found to be insignificant, lending further credence to the conclusion that attrition is not affecting our results.

1. As reported, 4% of the children died before the 2004 round, but the small absolute number involved makes it hardly sensible to conduct a detailed analysis.

Appendix Table A1. Correlates of Attrition (dependent variable = if child dropped out of sample): Conditional Fixed-Effects Logit Model

Variable	Coefficient
Male	0.063 (0.175)
Mother Resides in Household	0.007 (0.223)
Father Resides in Household	0.705** (0.249)
Mother's Years of Schooling	–0.001 (0.035)
Father's Years of Schooling	–0.014 (0.037)
Height in Centimeters	–0.010 (0.012)
Attending School Now?	–0.031 (0.219)
Household Head Male	0.179 (0.305)
Household Head Age	0.011 (0.007)
Household Head Years of Schooling	–0.040 (0.043)
Ln per Capita Consumption (TZ shillings)	–0.533* (0.215)
Dwelling Has Good Flooring	0.184 (0.267)
Number of Observations	987

Notes: Cluster and age fixed-effects logit regressions, with standard errors indicated in parentheses. Some observations were dropped because the fixed-effects logit is identified only for the cluster in which there was some attrition.

$^{*}p < .05$; $^{**}p < .01$

Appendix Table A2. Determinants of Height and Years of Schooling in 2004

	Dependent Variable	
Variable	Ln Height (1)	Years of Schooling (2)
Mother Died	–0.013† (0.007)	–1.240** (0.423)
Father Died	–0.001 (0.005)	–0.324 (0.306)
Number of Observations	718	718

Notes: Weighted 2SLS estimates with community and age fixed effects. Weights are the inverse of the predicted probability of inclusion in sample. Logit model predictions with specification as in Table A1. Standard errors are shown in parentheses. All other independent variables are as in Table 5 (coefficients not reported).

$^{\dagger}p < .10$; $^{**}p < .01$

APPENDIX B: TABLE 5 FULL REGRESSION RESULTS

Appendix Table B1. First-Stage Regression Results (Table 5)

	Dependent Variable	
Variable	Ln Height at Baseline (1)	Years of Schooling at Baseline (2)
Share of Farm Harvest Lost Prior to Baseline	–0.025*	
	(0.010)	
Deviation of Long Rains When Child Was 2–3 Years Old		0.001*
		(0.000)
7 Years Old	0.025**	–0.056
	(0.013)	(0.240)
8 Years Old	0.082†	0.122
	(0.013)	(0.240)
9 Years Old	0.125†	0.144
	(0.014)	(0.257)
10 Years Old	0.144†	0.514*
	(0.014)	(0.246)
11 Years Old	0.178†	1.168†
	(0.014)	(0.245)
12 Years Old	0.225†	2.172†
	(0.013)	(0.233)
13 Years Old	0.253†	2.807†
	(0.014)	(0.253)
14 Years Old	0.278†	3.884†
	(0.014)	(0.244)
15 Years Old	0.309†	4.549†
	(0.014)	(0.256)
Male	–0.007	–0.106
	(0.005)	(0.094)
Household Head Years of Schooling	–0.001	0.015
	(0.002)	(0.028)
Household Head Male	0.011	–0.219
	(0.011)	(0.202)
Household Head Age	0.000	0.008**
	(0.000)	(0.004)
Ln per Capita Consumption (TZ shillings)	0.007	0.348†
	(0.007)	(0.117)
Ln per Capita Consumption Missing	0.076	4.183†
	(0.080)	(1.418)
Dwelling Has Good Flooring	0.015**	0.248
	(0.009)	(0.163)
Height of Mother	0.002†	0.009
	(0.001)	(0.010)
Height of Mother Missing	0.248†	1.147
	(0.089)	(1.586)
Mother Resides in Household	0.007	0.118
	(0.015)	(0.260)

(continued)

(Appendix Table B1, continued)

Variable	Dependent Variable: Ln Height at Baseline (1)	Dependent Variable: Years of Schooling at Baseline (2)
Father Resides in Household	–0.012	0.181
	(0.009)	(0.161)
Mother's Years of Schooling	0.000	0.013
	(0.001)	(0.019)
Father's Years of Schooling	0.002	0.034
	(0.001)	(0.025)
Mother Died Before Child Was 15 Years Old	–0.012	0.197
	(0.011)	(0.191)
Father Died Before Child Was 15 Years Old	–0.006	–0.095
	(0.008)	(0.146)
Number of Observations	718	718
Test of Excluded Instruments *F* Value	5.58	4.95
Test of Excluded Instruments *p* Value	.019	.026

$^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$

Appendix Table B2. Second-Stage Regression Results (Table 5)

	Dependent Variable	
Variable	Ln Height in 2004 (1)	Years of Schooling in 2004 (2)
Ln Height at Baseline	–0.025*	
	(0.011)	
Years of School at Baseline		0.001*
		(0.000)
Mother Died Before Child Was 15 Years Old	–0.012	0.197
	(0.011)	(0.191)
Father Died Before Child Was 15 Years Old	–0.006	–0.095
	(0.008)	(0.146)
7 Years Old	0.025**	–0.056
	(0.013)	(0.240)
8 Years Old	0.082†	0.122
	(0.013)	(0.240)
9 Years Old	0.125†	0.144
	(0.014)	(0.257)
10 Years Old	0.144†	0.514*
	(0.014)	(0.246)
11 Years Old	0.178†	1.168†
	(0.014)	(0.245)
12 Years Old	0.225†	2.172†
	(0.013)	(0.233)
13 Years Old	0.253†	2.807†
	(0.014)	(0.253)
14 Years Old	0.278†	3.884†
	(0.014)	(0.244)
15 Years Old	0.309†	4.549†
	(0.014)	(0.256)
Male	–0.007	–0.106
	(0.005)	(0.094)
Household Head Years of Schooling	–0.001	0.015
	(0.002)	(0.028)
Household Head Male	0.011	–0.219
	(0.011)	(0.202)
Household Head Age	0.000	0.008**
	(0.000)	(0.004)
Ln per Capita Consumption (TZ shillings)	0.007	0.348†
	(0.007)	(0.117)
Ln per Capita Consumption Missing	0.076	4.183†
	(0.080)	(1.418)
Dwelling Has Good Flooring	0.015**	0.248
	(0.009)	(0.163)
Height of Mother	0.002†	0.009
	(0.001)	(0.010)
Height of Mother Missing	0.248†	1.147
	(0.089)	(1.586)

(continued)

(Appendix Table B2, continued)

	Dependent Variable	
Variable	Ln Height in 2004 (1)	Years of Schooling in 2004 (2)
Mother Resides in Household	0.007	0.118
	(0.015)	(0.260)
Father Resides in Household	–0.012	0.181
	(0.009)	(0.161)
Mother's Years of Schooling	0.000	0.013
	(0.001)	(0.019)
Father's Years of Schooling	0.002	0.034
	(0.001)	(0.025)
Number of Observations	718	718

$^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$