Table 1: Variables and their distribution (n=7668, unless otherwise stated)

Variables	N	%
Dependent variables		
Illness episodes (n=7705)*	667	8.66
Treatment sought (n=667)	555	83.21
Choice of providers (n=555)		
Self	378	68.11
Traditional healers	40	7.21
Medical care	137	24.68
Independent variables		
Age		
0-5 years	1,275	16.63
6-15 years	2,258	29.45
16-60 years	3,694	48.17
60+ years	441	5.75
Male	3,975	51.84
Schooling		
Never attended	6,249	81.49
Primary	1,181	15.4
Secondary or higher	238	3.1
Ethnicity		
Bwaba	1,763	22.99
Marka	3,027	39.48
Mossi	1,326	17.29
Others	1,552	20.24
Agriculture as first occupation	3,771	49.18
Married	2,852	37.19
Resident of Nouna	2,567	33.48
	Mean	**Std. Dev.**
Household six-month expenditure**	319,622	848,321
Medical care costs (n=555)**	3957	4701

* 37 individuals reported illness more than once.
** In CFA: 1 Euro = 655 CFA.

Table 2: Estimated coefficients in models with and without sample selection

Independent variables	Sample selection Probit			Ordinary probit			% change in Coefficient[+]
	Coeff	SE[$]	Sig.	Coeff	SE[$]	Sig.	
Intercept	1.151	0.467	***	-0.069	0.507		
Age (Ref group: 0-5 years)							
6-15 years	0.293	0.181		0.152	0.238		-93
16-60 years	0.410	0.231	*	0.585	0.259	**	30
60+ years	0.125	0.285		0.543	0.268	**	77
Male	0.179	0.111		0.233	0.133	*	23
Expenditure quintile (Ref group: Quintile 1)							
Quintile 2	0.192	0.197		0.138	0.256		-39
Quintile 3	0.034	0.173		0.045	0.222		24
Quintile 4	0.356	0.177	**	0.371	0.224	*	4
Quintile 5	0.501	0.181	***	0.608	0.201	***	18
Log of medical care costs	-0.167	0.068	**	-0.235	0.067	***	29
Schooling (Ref group: Never attended)							
Primary	-0.285	0.173		-0.388	0.203	**	27
Secondary or higher	0.086	0.245		0.141	0.328		39
Ethnicity (Ref group: Others)							
Bwaba	0.395	0.187	**	0.553	0.195	***	29
Marka	0.186	0.138		0.147	0.183		-27
Mossi	0.169	0.163		0.195	0.209		13
Married	0.082	0.131		0.101	0.179		19
Agriculture as the first occupation	-0.051	0.137		-0.152	0.167		66
Resident of Nouna	0.298	0.148	**	0.391	0.170	**	24
Number of observation	7578 (uncensored=540)[++]			540			
Log likelihood	-2135			-264			
Wald chi2(17)	29.68**			41.20***			
Rho	-0.73	(0.192)					
LR test of independent equations (Rho=0)	Chi2(1) = 5.06**			Model goodness of fit: Hosmer-Lemeshow chi2(8) = 5.06 (p=0.75)			

[$]Standard Errors adjusted for intra-individual correlation.
[+] Base: ordinary probit
[++] 15 observations excluded as observed medical care costs was zero, leading to undefined log of medical care costs.
***significant at 1%; **significant at 5%; *significant at 10%.
Haussman specification test (that difference in coefficients in two models not systematic): Chi2(17)=3.20; p=0.99
Likelihood Ratio test (that separate probit models are nested in sample selection model): Chi2(1)=105.04; p=0.00

Appendix A: The hedonic equation used to impute medical care costs

$$COST = 1{,}606.96 + 52.599*AGE + 2{,}076.26*SEX - 1{,}133.91*DIS2 + 97.093*DIS3$$

$$(0.69) \qquad (-1.5) \qquad (2.00)** \qquad (-0.97) \qquad (0.07)$$

$$+ 7{,}951.84*DIS4 + 4{,}060.82*DIS5 - 25555.046*NOUNA$$

$$(1.55) \qquad (1.99)** \qquad (-1.69)*$$

$$+ 1.443*POPULATION - 3{,}778.04*LAMBDA$$

$$(1.90)* \qquad (-3.23)***$$

Total observations = 137; R-squared = 0.24; Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Codes: DIS2= Infectious diseases other than malaria; DIS3=Non-communicable diseases;
DIS4=Injury; DIS5=Others; NOUNA= the town of Nouna; LAMBDA= the selection term generated
from a logit model of provider choice (the estimation details are available from the authors).

Appendix B: Estimated coefficients of selection equation (illness reporting)

Variables	Coefficient	Std. Err.[+]	Sig.
Age (Ref group: 0-5 years)			
6-15 years	-0.314	0.075	***
16-60 years	0.005	0.077	
60+ years	0.423	0.092	***
Male	-0.009	0.045	
Ethnicity (Ref group: Others)			
Bwaba	0.000	0.069	
Marka	-0.075	0.062	
Mossi	0.055	0.073	
Schooling (Ref group: Never attended)			
Primary	-0.027	0.072	
Secondary or higher	-0.002	0.143	
Agriculture as the first occupation	-0.091	0.060	
Household size (Ref group: With less than 5 members)			
With 5-10 members	-0.156	0.062	***
With more than 10 members	-0.422	0.068	***
Expenditure quintile (Ref group: Quintile 1, the poorest)			
Quintile 2	-0.131	0.088	
Quintile 3	0.071	0.082	
Quintile 4	0.004	0.085	
Quintile 5 (richest)	0.175	0.085	**
Housing and sanitation (instrumental variables)			
Metal roof	0.010	0.078	
Cemented soil	-0.069	0.076	
Latrine available	0.015	0.058	
Running water available	-0.142	0.057	***
Water disposed at courtyard	-0.032	0.048	
Water pots always covered	-0.098	0.049	**
Intercept	-1.033	0.114	***
Correlation coefficient (Rho)	-0.73**	0.192	

Number of observations = 7578 (Censored 7038)
Log likelihood = -2135**
Wald test of independent equations (rho = 0): Chi2(1) =5.06**
[+]Std. Err. adjusted for intra-individual correlation
*** significant at 1% level; **significant at 5% level; *significant at 10% level