

LETTER • OPEN ACCESS

## Quantifying the impacts of clean cooking transitions on future health-age trajectories in South Africa

To cite this article: Arda Aktas *et al* 2022 *Environ. Res. Lett.* **17** 055001

View the [article online](#) for updates and enhancements.

### You may also like

- [Jointly modeling the adoption and use of clean cooking fuels in rural India](#)  
Carlos F Gould, Xiaoxue Hou, Jennifer Richmond *et al.*
- [Household, community, sub-national and country-level predictors of primary cooking fuel switching in nine countries from the PURE study](#)  
Matthew Shupler, Perry Hystad, Paul Gustafson *et al.*
- [The cost of cooking a meal. The case of Nyeri County, Kenya](#)  
Francesco Fuso Nerini, Charlotte Ray and Youssef Boulkaid

ENVIRONMENTAL RESEARCH  
LETTERS

## LETTER

## Quantifying the impacts of clean cooking transitions on future health-age trajectories in South Africa

## OPEN ACCESS

RECEIVED  
6 April 2021REVISED  
10 March 2022ACCEPTED FOR PUBLICATION  
30 March 2022PUBLISHED  
19 April 2022

Original Content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.

Arda Aktas<sup>1,\*</sup> , Miguel Poblete-Cazenave<sup>2,3</sup> and Shonali Pachauri<sup>3</sup> <sup>1</sup> Population and Just Societies Program, International Institute of Applied Systems Analysis (IIASA), Laxenburg, Austria<sup>2</sup> Yale School of the Environment, Yale University, New Haven, CT, United States of America<sup>3</sup> Energy, Climate and Environment Program, International Institute of Applied Systems Analysis (IIASA), Laxenburg, Austria

\* Author to whom any correspondence should be addressed.

E-mail: [aktas@iiasa.ac.at](mailto:aktas@iiasa.ac.at)**Keywords:** clean cooking, microsimulation, scenario analysis, health impactsSupplementary material for this article is available [online](#)**Abstract**

Reliance on highly polluting cooking technologies poses a significant risk for human health. This study quantifies and compares the impact of different clean cooking access scenarios on future health-age trajectories among population subgroups in South Africa. Using microdata from five waves of the South African National Income Dynamics Study, we develop a dynamic microsimulation model and a composite metric of individual health status that is used to explore how health status changes under alternative access scenarios for the period 2010–2030. We find that there are clear gains of using clean cooking technologies for population health, and that electrification alone does not improve health status, if it is not accompanied by an increase in the use of clean cooking technologies in homes. Our results imply that achieving universal access to clean cooking in South Africa can by itself improve average population health by almost 4% by 2030 compared to a scenario without clean cooking technologies, with the health of individuals of genders and races with the poorest health and well-being endowments improving the most. Thus, clean cooking can contribute to narrowing existing inequalities by improving health for the most vulnerable population groups that disproportionately depend on polluting cooking technologies.

**1. Introduction**

Exposure to pollution in homes caused by the incomplete combustion of solid fuels and kerosene in inefficient cooking stoves and devices is a major cause of premature death and acute illness due to respiratory, cardiovascular, and circulatory diseases (Smith and Pillarisetti 2017, Shupler *et al* 2018, Hystad *et al* 2019, Arku *et al* 2020). Globally, household air pollution from cooking with solid fuels has been estimated to cause between 1.6 and 3.8 million deaths annually (GBD 2017 Risk Factor Collaborators 2018, Landrigan *et al* 2018, WHO 2018). In South Africa alone, an estimated 13 642 [8218–19 762] premature deaths were estimated to be attributable to household air pollution in 2016 (WHO 2020). These adverse health impacts are disproportionately borne by women and children, as they are most often exposed to most of the indoor air pollution due to

them spending longer periods at home, and particularly by women, as they are mainly responsible for cooking (Edwards and Langpap 2012, Lin *et al* 2013, Dutta and Banerjee 2014, Daset *et al* 2018, Bede-Ojimadu and Orisakwe 2020).

Research on estimating health impacts of cooking with polluting fuels and stoves principally use comparative risk assessment (CRA) methods employed by the global burden of disease (GBD 2017 Risk Factor Collaborators 2018). These methods rely on findings from epidemiological research on relative risk estimates for specific health outcomes as a function of exposure level with impacts measured in terms of deaths, years of life lost, years lived with disability, or disability-adjusted life-years (for example see Abtahi *et al* 2017, Owili *et al* 2017, Arku *et al* 2018, Yu *et al* 2018). Relatively few studies have investigated how cooking choices affect broader health status using individual level health indicators such as self-assessed

health status (Liu *et al* 2018) and activities of daily living (i.e. ADL and IADL) (Liu *et al* 2020). Also, though microsimulation modeling has been used to quantify the impacts of environmental exposures on health (Schofield *et al* 2018, Symonds *et al* 2019), the application of this method to estimate how cooking with polluting fuels and stoves affects health has not been explored thus far.

In this study, we quantify the impact of different clean cooking access scenarios on future health-age trajectories across population subgroups in South Africa to examine the extent that existing health disparities among population subgroups in South Africa can be attenuated through the adoption of clean cooking technologies. Specifically, we develop a dynamic microsimulation model, using data from the South African National Income Dynamics Study (NIDS), to project health trajectories of individuals in South Africa over the period 2010–2030 under different clean cooking access scenarios.

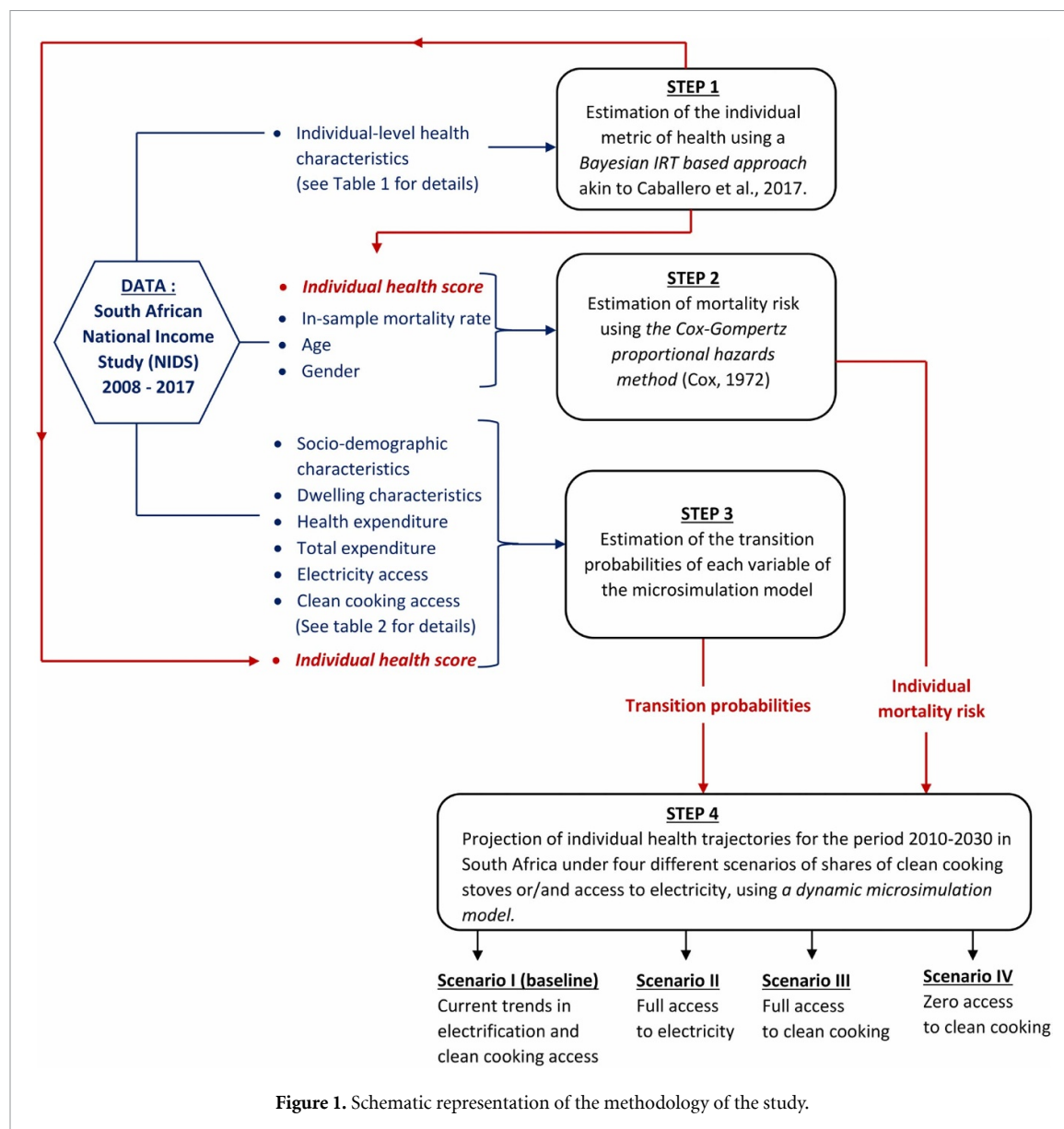
As the use of polluting fuels and stoves is a well-established risk factor for multiple diseases and mortality, we focus our analysis on the question of how much the latent health level of South Africans could improve if everyone used clean cooking technologies. To address this question, we follow a procedure similar to Marois and Aktas (2021), which allows us to create health-age trajectories over time for people exhibiting different behaviors. Specifically, we use a microsimulation model to analyze the impacts of different cooking technologies, as well as other relevant factors, on the health trajectories of individuals over time. This microsimulation model requires three preparatory stages. First, we estimate a composite measure of individual health status, following a Bayesian Multi-Level Item Response Theory methodology developed and validated in Caballero *et al* (2017) and de la Fuente *et al* (2018) using a large set of health characteristics reported in the NIDS. Second, to model how the survival probabilities of individuals vary with their level of health, we estimate the hazard rates for individuals in our sample as a function of the estimated health metric. Third, we use the empirical data to estimate the transition probabilities of the relevant factors affecting both health and the adoption of different cooking technologies, which we represent as variables of our microsimulation model. Finally, we use these inputs in our microsimulation model to project future health-age trajectories for the period 2010–2030 under different scenarios of levels of access to clean cooking. We consider four alternative scenarios: a baseline scenario, which assumes the continuation of current trends and no new access policies; two different policy scenarios, one that assumes universal access to electricity and another that assumes universal access to clean cooking; and a final extreme counterfactual scenario that assumes no access to clean cooking.

Our results show that there are clear gains to using clean cooking technologies in terms of population health. In fact, our results show a more than 7% increase in the number of females in good health in 2030 that can be attributed to the adoption of clean cooking technologies, with almost half of these being females of African descent in urban areas with some level of schooling. However, we also find that electrification alone does not improve health if it is not accompanied by an increase in the uptake of clean cooking technologies. Our findings imply that achieving universal access to clean cooking fuels and technologies in South Africa can bring significant improvements in population health and may also contribute to narrowing existing inequalities in health observed by gender and education by improving the health of the most vulnerable population groups who disproportionately rely on polluting cooking technologies.

We contribute to the literature in two significant aspects. First, the novel health metric we estimate captures various aspects of individual health in a single composite measure, in contrast to other objective health measures often used in the literature, such as the prevalence of a particular disease or adverse health conditions. Another advantage of this health metric, especially over indicators which are based on people's subjective assessment of their health status (i.e. self-rated health), is that it is not time and culture dependent, therefore it allows for a comparison of health status across countries and sub-populations, and over time. Moreover, the health metric is constructed as a continuous variable, which allows us to quantify the marginal and cumulative effect of cooking with polluting stoves on individual health. Second, we develop a novel dynamic microsimulation model to explore how future health-age trajectories of the South African population change under alternative scenarios of a transition to clean cooking fuels and technologies. While microsimulation models have been developed to assess the health impacts of environmental pollution in previous literature (e.g. Pimpin *et al* 2018, Symonds *et al* 2019), we are unaware of any studies that employ this method to evaluate the health impacts of clean cooking policies through scenario modeling.

## 2. Data and estimation of parameters of the microsimulation model

For this study, we used longitudinal data from the five available waves of the South African NIDS corresponding to the years between 2008 to 2017. NIDS is a nationally representative panel survey that contains detailed information on households demographic and socio-economic characteristics (Woolard *et al* 2010). The survey is particularly well suited for our purposes as it also includes information on household cooking energy sources and health-related variables



needed to compute the health metric. We use this data for three different purposes related to the preparation of the variables and parameters of the microsimulation model:

- Use data on health characteristics of individuals from the NDIS sample to estimate a metric of individual health status.
- Use mortality rates from the NDIS sample to estimate the survival probabilities of individuals depending on their health.
- Use observed changes over the five waves of the NDIS sample to estimate transition probabilities of a set of factors relevant to the modeling of the impact of cooking technologies on individual health over time.

For these estimations, we only retain observations from the NDIS sample for individuals that have non-missing information on all the variables needed,

namely, the variables presented in tables 1 and 2. All the variables and parameters estimated are then used in the construction of the microsimulation model, that is further described in section 3. A schematic overview of the methodological framework, along with the data and indicators used in each step is depicted in figure 1. We present the specific data and estimation procedures for these three purposes independently in the subsections below.

### 2.1. Estimation of the measure of individual health status

The health status of individuals in each sample is estimated following a procedure akin to Caballero *et al* (2017). The methodology for developing a composite health metric and the estimation method has already been presented and validated, first in (Caballero *et al* 2017) and also in other studies (de la Fuente *et al* 2018, Daskalopoulou *et al* 2019, Marois and Aktas 2021) using different datasets. The

suitability of the metric for causal inference has also been tested (Kollia *et al* 2018).

Briefly, this approach assumes that there exists a latent health variable that can be inferred from a set of observed health-related characteristics. The distribution of a health score is estimated in a way that it reflects the distribution of the observed health status in a particular sample. Formally, an individual  $i$  belonging to a group  $j$  that has a health score  $\theta_{ij}$  has a probability of having a health characteristic  $k$  such that:

$$P(H_{kji} = 1|\theta_{ij}) = \phi(a_{kj}\theta_{ij} + b_{kj}),$$

where  $H_{kji}$  is a dummy reflecting whether or not individual  $i$  in group  $j$  has health characteristic  $k$ ,  $\phi(\cdot)$  is the c.d.f. of the standard normal distribution, and where

$$a_{kj} \sim \mathcal{N}(1, \omega_a^2)$$

$$b_{kj} \sim \mathcal{N}(\mu_b, \omega_b^2)$$

are group-specific ‘discrimination’ and ‘difficulty’ parameters, respectively. The discrimination parameter is related to how the likelihood of having a particular health condition  $k$  decreases with the health score  $\theta_{ij}$ , while the difficulty parameter represents how likely (or hard) it is to have a particular health characteristic  $k$ .

Finally, the health score  $\theta_{ij}$  has a random group and individual-group components:

$$\theta_{ij} = u_j + e_{ij}$$

$$u_j \sim \mathcal{N}(0, \sigma_{L1}^2), e_{ij} \sim \mathcal{N}(0, \sigma_{L2}^2).$$

The values of  $\omega_a^2$ ,  $\mu_b$ ,  $\omega_b^2$ ,  $\sigma_{L1}^2$  and  $\sigma_{L2}^2$  are estimated using a Bayesian multilevel item-response theory (MLIRT) approach on a set of health characteristics including self-reported health questions and measured tests obtained from the longitudinal household surveys.

We adapt the Bayesian MLIRT approach described in de la Fuente *et al* (2018) to our context, to estimate health metric scores using a consistent set of 14 health characteristics available in all five waves of the NIDS (see table 1). In our application, the group variable  $j$  represents the particular wave where the observation is taken from, hence creating a longitudinally consistent version of the health metric (Verhagen and Fox 2013). Following de la Fuente *et al* (2018), the Markov Chain Monte Carlo estimation is conducted using 5000 iterations and 100 burn-in iterations. The latent health score is then created by normalizing the Expected *A Posteriori* estimates on a scale from 0 to 1, with higher values indicating better health.

In contrast to other objective or subjective health measures commonly used in related literature, such as the prevalence of a particular disease (e.g. chronic illnesses, respiratory diseases, cardiovascular diseases, cerebrovascular diseases), or self-rated health status, our health metric is able to simultaneously capture

different aspects of individual health in a single measure. Additionally, given that it is estimated as a continuous variable defined on 0–1 interval, it does not suffer from epistemic problems due to the discretization of health into different categories.

Figure 2 presents the kernel density estimates of the health metric for different population subgroups: urban/rural place of residence, gender and race<sup>4</sup>. It is important here to reiterate that the calculation of the health metric is at an individual level and does not involve additional factors besides the health characteristics presented in table 1. Hence, all the results presented here cannot be traced to specific coefficients associated to any of the particular demographic characteristics used for the population groupings. As an additional validation mechanism, we compare this distribution with the distribution of self-rated health, another widely used metric of health. We find that our estimated health metric is able to appropriately capture the trends that are observed in self-rated health, but in a continuous manner. In particular, the distribution of health seems heavily skewed to the right, indicating more people are in good health. As it is normally observed, men have better health than women, in general. However, interestingly, men of African descent in our sample (figure 2(a)) seem to be the specific population subgroup with better health.

## 2.2. Estimation of the individual survival probabilities

In order to make our model as self-contained as possible, we use in-sample mortality data to estimate hazard rates. To do so, we employ the Cox regression method (Cox 1972), to estimate a Gompertz proportional hazards models of mortality risk, using our health index as the only additional explanatory covariate, namely:

$$h_t = \frac{\eta}{\nu} * \exp(t\nu) * \exp(\beta * health_t) + \epsilon_t \quad (1)$$

where  $h_t$  is hazard rate at age  $t$  (i.e. the probability of dying at a given age  $t$ ),  $\eta$  and  $\nu$  are shape and scale parameters, and  $\beta$  is the effect of health on the overall hazard rate. We deliberately decided not to include additional explanatory variables that affect survival probabilities (e.g. a dummy for clean cooking), as we want to investigate all such additional effects indirectly through health.

The estimated parameters of equation (1) can be found in table S2. Visually, the estimated survival curves for our sample are presented in figure 3. We find that the gender differences observed in health

<sup>4</sup> The original sample presents four different racial groups: individuals of African descent, Colored, Asian/Indian and White. Of these, only individuals of African descent exhibit significant differences in health characteristics compared to the other groups. Therefore, we focus in our analysis on differences between African descendents and others, accordingly.

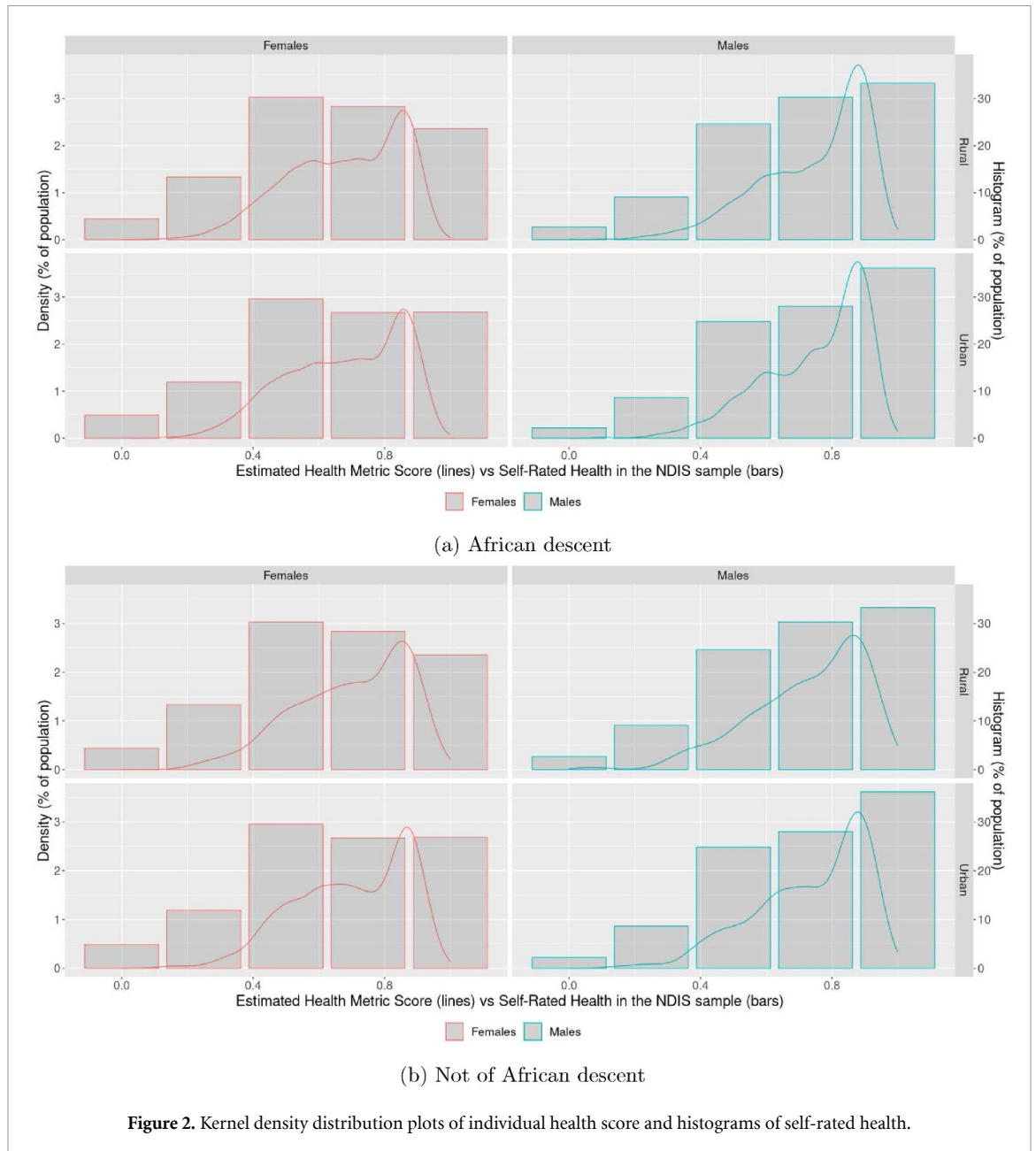
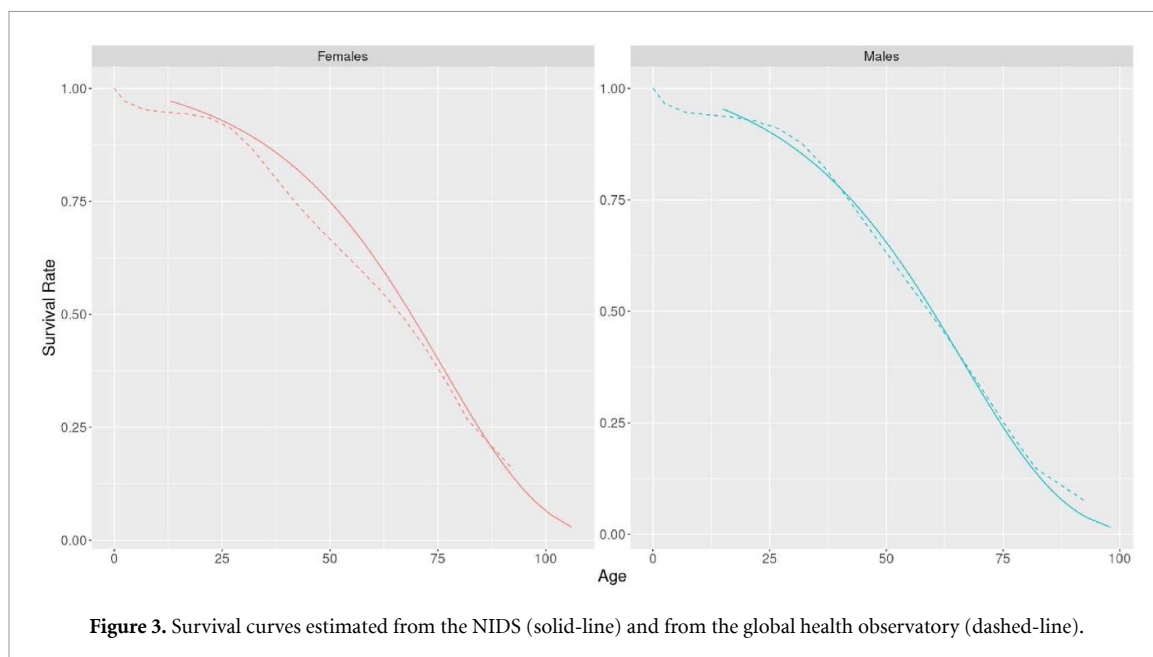


Figure 2. Kernel density distribution plots of individual health score and histograms of self-rated health.

Table 1. Descriptive statistics for the health characteristics in each survey sample.

	Fever <sup>a</sup>	Persistent cough <sup>a</sup>	Cough with blood <sup>a</sup>	Chest pain <sup>a</sup>	Body ache <sup>a</sup>	Headache <sup>a</sup>	Back ache <sup>a</sup>
2008	14.2%	11.4%	2.0%	10.7%	17.5%	32.6%	17.0%
2010	12.9%	10.7%	1.1%	7.1%	11.8%	21.3%	11.2%
2012	18.8%	17.4%	1.2%	9.9%	17.9%	30.9%	16.3%
2015	21.1%	12.5%	1.1%	9.2%	19.8%	35.3%	16.0%
2017	22.1%	11.4%	1.0%	7.3%	17.0%	31.3%	14.6%
	Joint pain <sup>a</sup>	Diarrhea <sup>a</sup>	Painful urination <sup>a</sup>	Swelling of ankles <sup>a</sup>	Weight loss <sup>a</sup>	Good eyesight	Good hearing
2008	13.5%	3.6%	2.9%	5.6%	2.7%	79.9%	91.3%
2010	9.1%	1.6%	1.4%	3.2%	2.6%	87.5%	90.1%
2012	11.3%	2.5%	1.8%	4.4%	2.8%	90.8%	95.6%
2015	12.3%	3.5%	1.8%	5.1%	3.3%	90.1%	96.1%
2017	10.8%	4.3%	1.2%	4.1%	3.2%	89.2%	95.2%

<sup>a</sup> Has been experienced in the past 30 days.



status are also reflected in the survival rates but in opposite direction, i.e. mortality risk of men is significantly higher than mortality risk of women, whereas men are healthier than women. This finding is in line with the literature (Verbrugge *et al* 1987, Case and Paxson 2005, Alberts *et al* 2014). We validate our estimates by comparing our survival curves with the survival projections by gender from the WHO Global Health Observatory, which can also be seen in figure 3 represented by dashed lines. Our estimates match these quite closely, and are only slightly higher for females around 50 years old.

### 2.3. Estimation of the transition probabilities

In our microsimulation model, we focus on a limited set of variables that can be related to either the adoption of clean fuels, or the damages to health associated with the use of polluting stoves<sup>5</sup>. In this regard, it is important to acknowledge that South Africa is a very special case among Sub-Saharan countries, due to ambitious policies in place aimed at increasing electrification. A side effect of these policies is that we do not find significant evidence of fuel stacking, nor the major use of other sources of clean energy for satisfying the cooking needs of the population besides electricity (see table S1 in

<sup>5</sup> As a simplifying assumption, we assign to individuals the level of education indicated by them in the last wave of the survey. We assume this to avoid the computational burden of adding an additional layer of simulation which, for most individuals, would not make any difference given their age. We define four education categories: no schooling, less than high school, high school graduate, college degree. Table 2 presents the average level of education of the sample, assuming values of 0 for no schooling, 1 for less than high school degree, 2 for high school degree, and 3 for college degree, respectively. However, in the estimation, independent dummies are used for each of the 4 education categories.

the supplementary information available online at [stacks.iop.org/ERL/17/055001/mmedia](https://stacks.iop.org/ERL/17/055001/mmedia)). Additionally, Kolmogorov-Smirnov tests on the distribution of health over the different cooking fuels does not show statistical differences between the health of those using electricity or gas, nor differences between the distribution of health for those using kerosene, coal or biomass (table S2). Hence, we constrain the microsimulation model to only two cooking technology choices: clean (i.e. electricity or gas) or other, non-clean fuels.

In table 2, which presents descriptive statistics for selected variables, we can see the rapid improvements in the living conditions of individuals in our sample during the observation period. For example, electrification increased 10% during the period, while the adoption of clean cooking stoves improved almost 20%. This probably attenuated the gradual decline in the health of individuals as the sample ages. A noticeable impact of the 2008 financial crisis is also observable, with some estimates for the 2010 wave presenting great deviations from the overall trends.

There are three types of variables in the sample: dichotomous variables, that is, those taking two discrete values (e.g. gender, urbanization); left-bounded continuous variables, taking values from zero upwards (e.g. total expenditure); and fully bounded continuous variables, e.g. health, that take values between zero and one. Given these particularities of the data, we use different parameterizations to empirically estimate transition pathways for each variable.

In general, the estimation of the transition rates through the simulation period for these variables, with the exception of health and health expenditures, follow the general form:

**Table 2.** Descriptive statistics of the microsimulation estimation sample.

	Female	African descent	Level of education <sup>a</sup>	Urban	Total expenditure <sup>b</sup>	Health expenditure <sup>b,c</sup>
2008	67.7%	83.7%	—	45.9%	1628.0	29.6
2010	67.7%	83.7%	—	46.7%	1456.9	18.5
2012	67.7%	83.7%	—	49.6%	1863.1	37.7
2015	67.7%	83.7%	—	51.5%	2150.4	32.9
2017	67.7%	83.7%	1.187	51.6%	2151.2	31.0
	Has electricity	Single room	Light material	Clean cooking <sup>d</sup>	Works	Average health score
2008	79.8%	07.8%	86.4%	66.4%	30.7%	0.713
2010	77.8%	05.9%	81.7%	75.5%	31.0%	0.742
2012	85.0%	08.5%	83.8%	77.6%	35.9%	0.702
2015	88.2%	10.9%	82.7%	80.2%	40.6%	0.685
2017	87.2%	10.1%	70.5%	83.1%	40.7%	0.688

<sup>a</sup> Mean level, see footnote 5.

<sup>b</sup> Monthly mean values in 2008 South African Rands.

<sup>c</sup> Out-of-pocket.

<sup>d</sup> Electricity or gas.

$$\begin{aligned}
 Y_{w+1} \sim & \text{age}_w + \text{hh\_size}_w + \log(\text{hh\_expenditure}_w) + \log(\text{hh\_health\_expenditure}_w) \\
 & + \text{urban}_w + \text{has\_electricity}_w + \text{single\_room}_w + \text{light\_material}_w + \text{clean\_cooking}_w \\
 & + \text{health}_w + \text{female} + \text{african\_descent} + \text{educ\_level} + \text{working}_w + \epsilon_w.
 \end{aligned} \tag{2}$$

This form is equivalent to a panel autoregressive model with covariates, where  $Y$  is the variable of interest and  $w$  is the wave. For the case of dichotomous variables, we carry out a logit regression, whereas for household expenditure, we estimate a log-linear regression. Additionally, in some cases, and in order to overcome some of the anomalies in the data, dummies for specific years are added to the estimation (e.g. a 2008-to-2010 dummy to control for the effect of the financial crisis).

Two of the variables presented in table 2, health and out-of-pocket health expenditures, are estimated using a form different than (2). For the case of health, we estimate different transition pathways for each member of the simulated population depending on whether the individual lives in a household that uses clean cooking or not. In general, similar to Marois and Aktas (2021), the form used to estimate health transitions is:

$$\begin{aligned}
 \text{logit\_health}_{w+1} - \text{logit\_health}_w \sim & \text{age}_w + \text{hh\_size}_w + \log(\text{hh\_expenditure}_w) \\
 & + \log(\text{hh\_health\_expenditure}_w) \\
 & + \text{urban}_w + \text{has\_electricity}_w + \text{single\_room}_w \\
 & + \text{light\_material}_w + \text{health}_w \\
 & + \text{cooking\_transition}_w + \text{female} + \text{african\_descent} \\
 & + \text{educ\_level} + \text{working}_w + \epsilon_w
 \end{aligned} \tag{3}$$

where  $\text{cooking\_transition}_w$  is a dummy representing a transition from polluting to clean cooking technologies or vice versa.

Finally, given that out-of-pocket health expenditures are not observed in every wave of the survey, we first estimate a household's probability of

incurring out-of-pocket health expenditures at a particular point of time, and then posteriorly, for households that have out-of-pocket health expenditures, we estimate the relationship between the amount spent and other household characteristics:



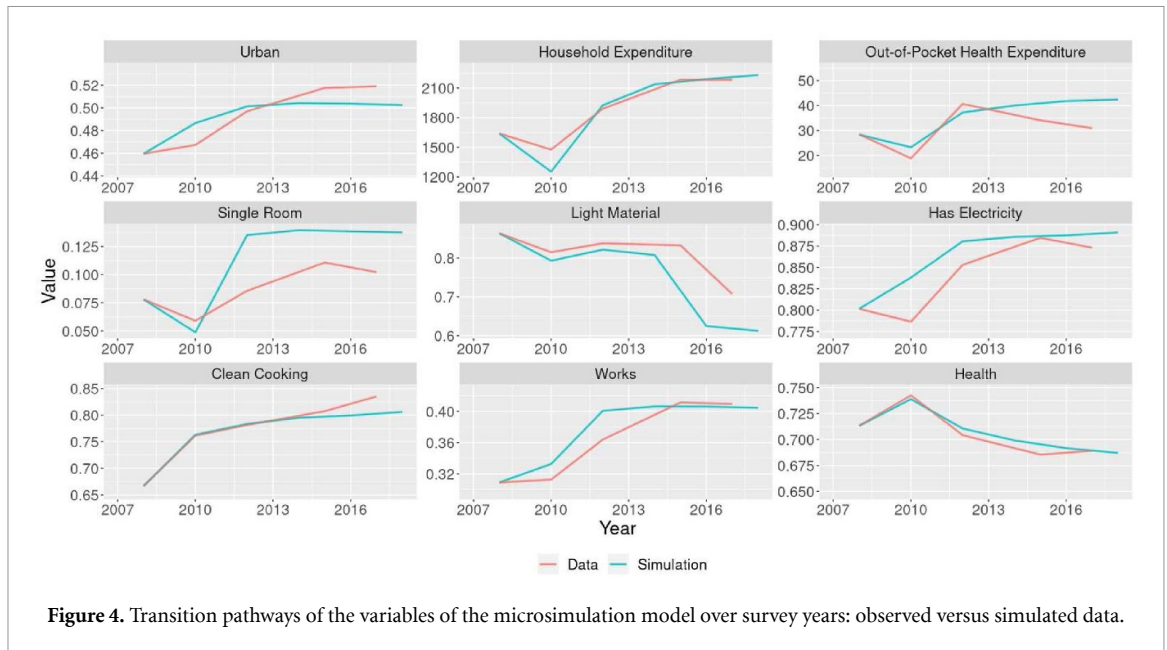


Figure 4. Transition pathways of the variables of the microsimulation model over survey years: observed versus simulated data.

$$\begin{aligned}
 P(hh\_health\_expenditure_w > 0)_w \sim & age_w + hh\_size_w + \log(hh\_expenditure_w) + urban_w \\
 & + has\_electricity_w + single\_room_w + light\_material_w \\
 & + clean\_cooking_w + health_w + female + african\_descent \\
 & + educ\_level + working_w + \epsilon_w
 \end{aligned} \tag{4}$$

$$\begin{aligned}
 \log\_hh\_health\_expenditure_w(| > 0) \sim & age_w + hh\_size_w + \log(hh\_expenditure_w) + urban_w \\
 & + has\_electricity_w + single\_room_w + light\_material_w \\
 & + clean\_cooking_w + health_w + female + african\_descent \\
 & + educ\_level + working_w + \epsilon_w.
 \end{aligned} \tag{5}$$

Using this simple parameterization, we are able to capture several of the characteristics of the observed data, although, in some cases, additional calibrations are necessary (e.g. tweaking the strength of the effect of the financial crisis). The estimated parameters are provided in table S4 in the supplementary information.

To assess the validity of our estimates, figure 4 shows the observed transition pathways of the modeled variables in comparison with the ones calculated using the microsimulation model. Although the fit is not perfect in every case, the trends are very well captured by the microsimulation model. Hence, we purposefully do not work further to improve the fit to avoid over-fitting, which can prevent generalizability of the results.

### 3. Microsimulation and results

We use our dynamic, discrete-time microsimulation model to project the impacts of alternative

cooking technologies on future health-age trajectories in South Africa for the period 2010–2030. We develop four simulation scenarios. The first is a baseline scenario, under which we assume current trends, without enforcing any new policies to encourage clean cooking and no alteration to the current trends for extending electricity access. We also develop two policy scenarios that assume interventions that provide immediate full access to either electricity or clean cooking to the entire sample population. We differentiate between a scenario of access to electricity and a scenario of access to clean cooking technologies, since, even though most access-related policies, not only in South Africa, but other countries in the Global South, are aimed at increasing electrification rates, many of those who have access to electricity still do not cook with electricity as it may be unaffordable to them (as it is also reflected in the recent IEA report (IEA 2020)). With this, we intend to quantify differences in health that may arise from electrification alone, from those that accompany the

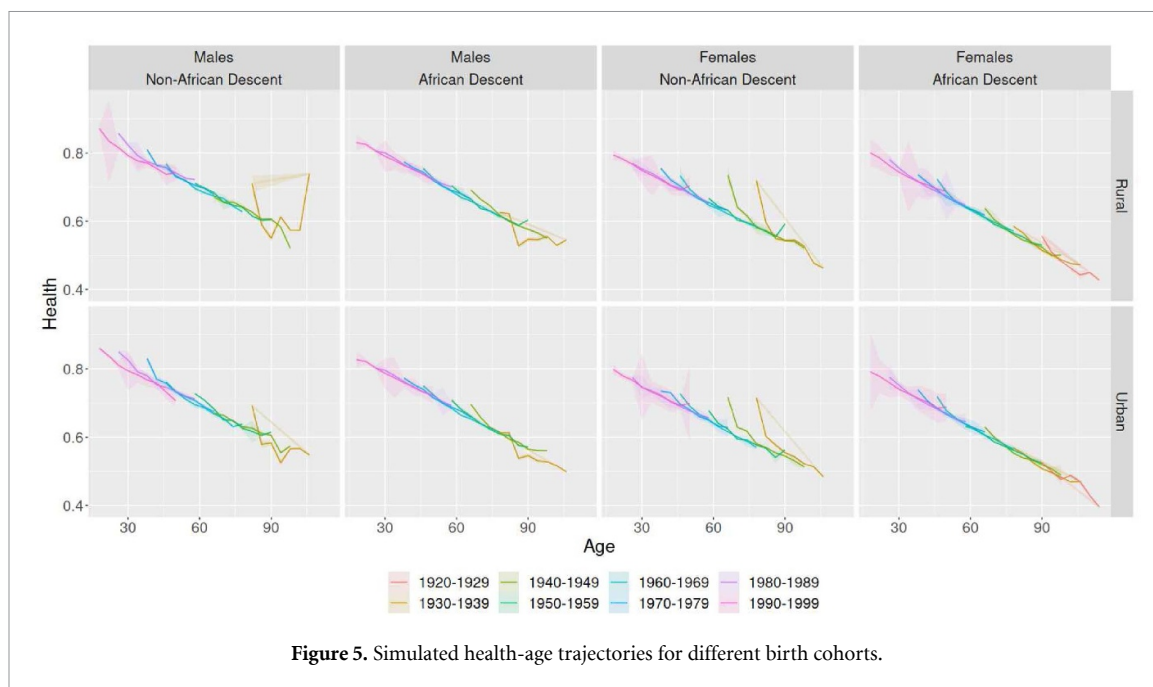


Figure 5. Simulated health-age trajectories for different birth cohorts.

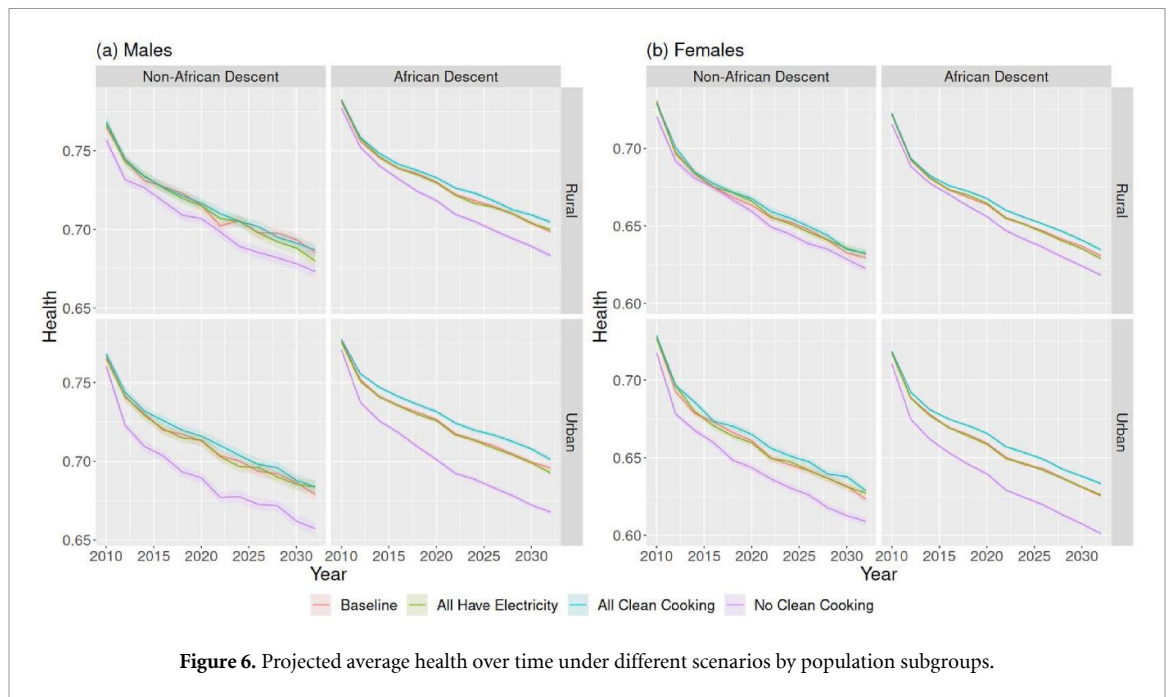
use of clean cooking technologies. However, given the relatively low rates of non-clean cooking in all of these scenarios, we develop an additional, extreme counterfactual scenario where we assume that no individual has access to clean cooking. This scenario presents an extreme contrast to better assess the long-term health impacts of clean cooking, that otherwise, would be hard to appreciate.

Here, we present the details of our microsimulation model. In our framework, the population consists only of within-sample individuals. Specifically, we use values from the estimation sample in 2008, corresponding to a total of 3638 individuals with no missing data for all the variables involved in both the estimation of the health metric and the microsimulation, as our starting values and then use the transition parameters estimated to simulate trajectories up to the year 2030. All characteristics of each individual (except gender, race and education level<sup>5</sup>) are updated at each step of the simulation period. Transition probabilities between different states are determined stochastically, where uncertainty arises from random draws of the distribution of the error terms in equations (2)–(5), and also, for probabilistic variables, from random draws of a uniform distribution. We do not include additional individuals at any point of the simulation (through births). However, individuals can leave the sample at any stage if they die according to the hazard rates estimated in equation (1). (i.e. we assume a closed population). We keep our simulation model simple and constrained to the NDIS sample on purpose, so that it can serve as an exploratory tool to compare the impact of different cook-stove and electrification policies on the health-age trajectories of a specific population.

We perform a total of one hundred Monte Carlo runs to assess the stability of our results<sup>6</sup>. The resulting health-age trajectories by cohort are seen in figure 5. We find a similar rate of decline in health for men and women, although women, regardless of their lower relative health, tend to live longer.

Besides this general result, our alternative ‘what if’ scenarios provide useful insights on the impact of different clean cooking access policy scenarios on future population health-age trajectories. In figure 6, we compare the evolution of population health over time in these scenarios with our baseline scenario assuming the continuation of current trends and no new access policies. Notice that, even under our baseline scenario, the rate of electrification and adoption of clean cooking technologies already improve significantly (as seen in the observed data, table 2 and in the simulation, figure 4). Therefore, the differences in health outcomes between the baseline and access scenarios are not as large as one would expect, e.g. as compared to a country where these developments happen at a much slower pace or from a lower base level. For example, we can see that for the population of non-African descent, the health effects are limited and for some years, statistically insignificant. However for the population of African descent, we find that access to clean cooking improves health over time. The introduction of the no clean counterfactual scenario is therefore relevant, as it markedly displays the stark differences in the pathways that would occur in the case that no one uses clean cooking technologies. We see that, not only is there an initial level drop

<sup>6</sup> Average values of the microsimulated variables in some selected years can be seen in table S5 in the supplementary information.



**Figure 6.** Projected average health over time under different scenarios by population subgroups.

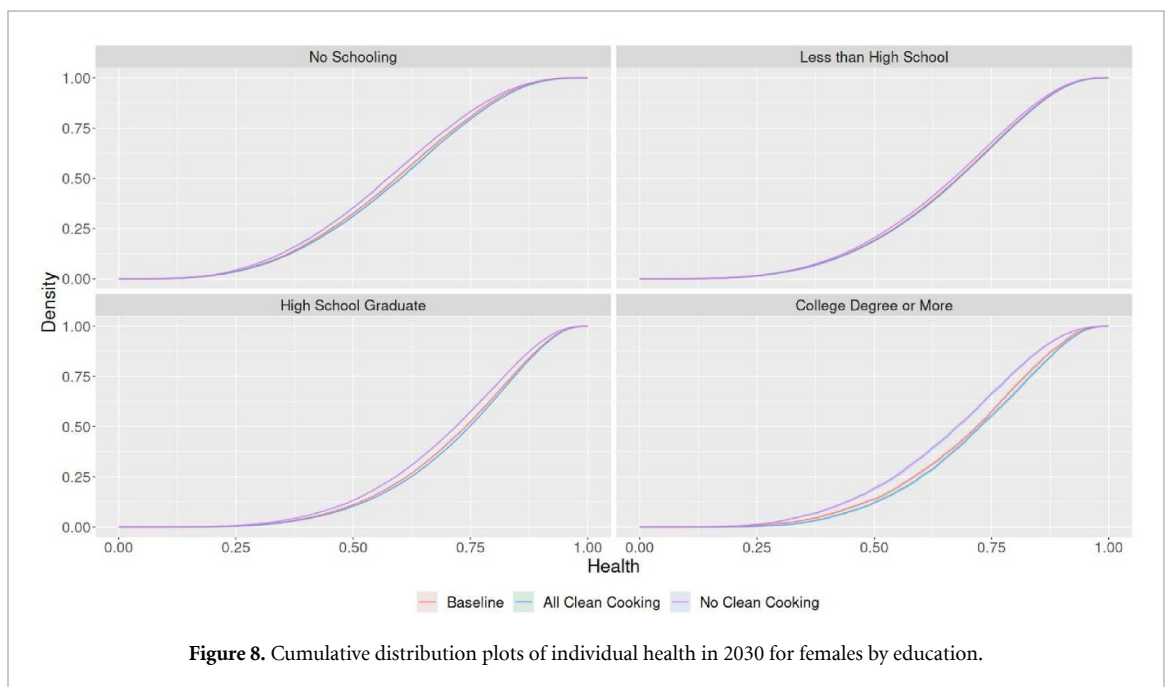
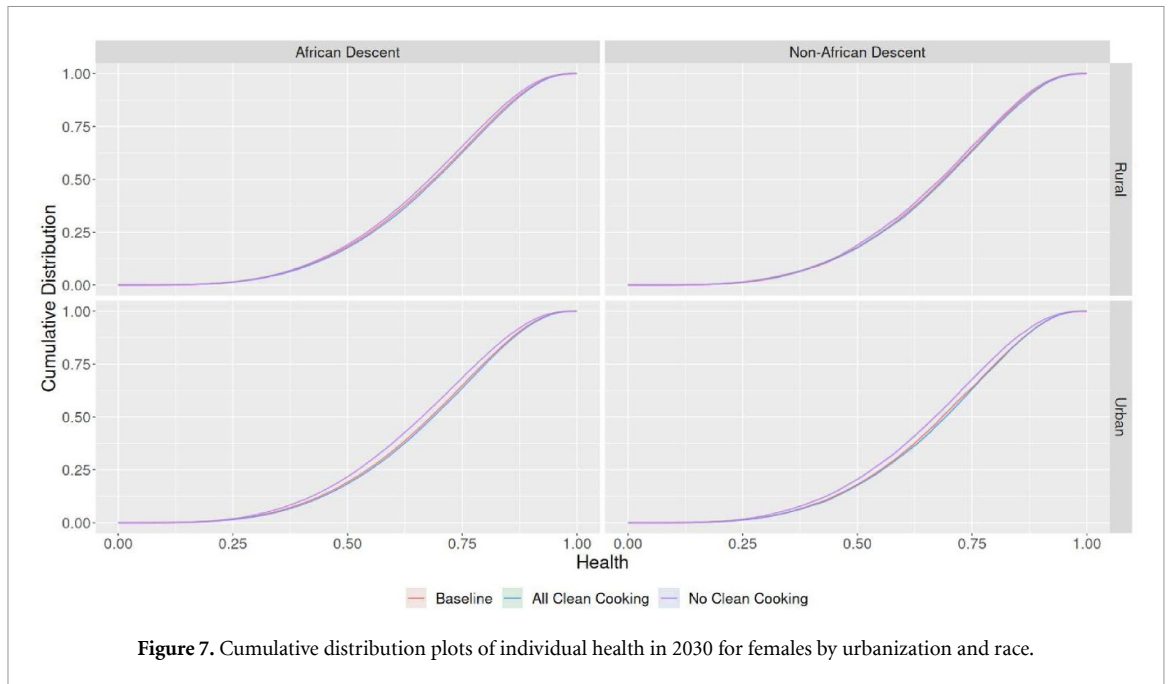
in average health, but that the difference greatly accumulates and amplifies over time for all groups, particularly for the African descent population. Finally, and of great interest, we find that electrification by itself does not lead to any health improvements when compared with the baseline scenario.

Also, we find that in urban areas, the differences in health are more significant and persistent. Here, there may be another effect at play that, unfortunately due to data limitations, could not be added to the model: in rural areas it may be more feasible to perform some cooking outdoors, which reduces exposure to pollution indoors. Besides, urban areas have higher levels of ambient air pollution, which can also increase the cumulative health effects of air pollution in homes.

This is also related to another phenomena we observe in the simulations, which is the positive relationship between health and affluence, particularly noticeable in the scenario with no clean cooking technologies, where expenditures are much lower by the end of the horizon (e.g. 16% lower on average compared to the baseline scenario, see table S5). This has two opposite effects on health: a direct, negative impact, by lowering health expenditures, but also an indirect positive impact associated, interestingly, with a worsening of living conditions. In particular, the number of individuals living in dwellings made out of light materials is higher in scenarios where expenditures are lower. This might be associated with lower pollution exposure for members of such households, as in such dwellings there is a higher probability that cooking may be performed outdoors or, even if cooking occurs inside the house, particulate matter can leave the dwelling faster if materials of the walls or ceiling are lighter and more porous (Dasgupta *et al* 2006, de la Sota *et al* 2018).

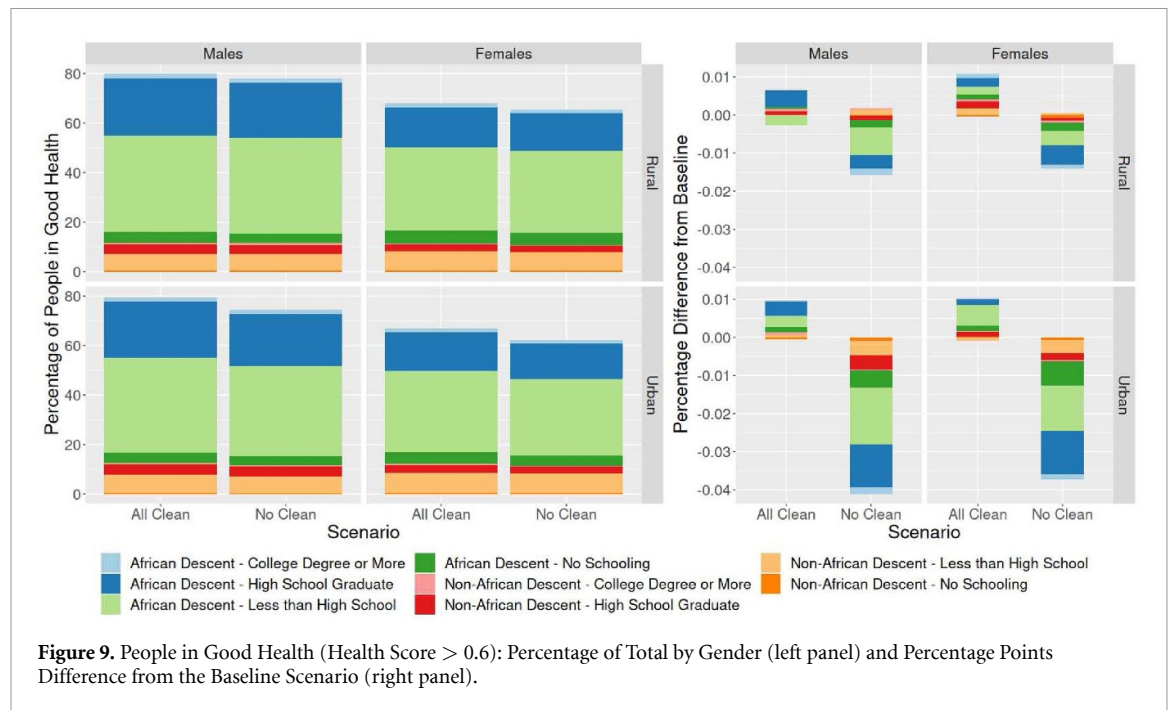
Considering that in most households males are not directly involved in cooking nor do they spend as much time indoors due to their participation in the labor force, we focus our analyses further on females specifically, as they are most likely to be affected by indoor air pollution from cooking. We also disregard the full electrification scenario in what follows, as we find it does not show any significant differences with the baseline. Therefore, we focus the rest of the analysis on the 'what-if' counterfactual scenarios related to the adoption of clean cooking technologies, in order to clearly evaluate its effects on health. Figure 7 shows the health distribution of females in the final year of our simulation, the year 2030. Notice that these curves represent the average of the distributions of health over all the bootstrap estimates (including their relatively small confidence bands), and therefore, the differences that arise due to randomization are minimized. Given this, differences between the baseline and the universal access to clean cooking scenario are subtle, and more evident for the African descent population. However, differences with the no clean cooking counterfactual are significant, pointing to the importance of clean cooking technologies for overall health.

The population subgroups depicted in the figure are not easily comparable, due to differences in income level and other overall living conditions. In this regard, it is known that the level of education, particularly of females, is an overarching proxy measure of these factors, as more educated people tend to have higher levels of income, as well as better housing conditions and health behaviors (Montgomery *et al* 2000). Hence, by examining health distributions by level of education, we may gain deeper insights into the simulated phenomena. Indeed, as



shown in figure 8, when controlling for education level, differences are clearly noticeable, specially with respect to the no clean cooking counterfactual scenario. For females with no college degree, clean cooking access creates a difference relative to the baseline scenario, albeit seemingly small, in particular for the less than high school educated group. However, we observe the largest differences in health outcomes among the highest educated group, as in addition to the gains from clean cooking, we expect that other improvements related to living conditions create important synergies for the health status of this group.

To quantify the effect beyond the shape of the distribution, we undertake one final exercise. We classify individuals as being in good health if their health level is above 0.6, as in Marois and Aktas (2021). This is also consistent with the median point of the health density of the entire population as presented in figure 2. Figure 9 shows, on the left panel, the percentage of population in good health by gender for each scenario and for all population subgroups. We can see that the scenario with universal clean cooking access again clearly shows the best health for both males and females, with a very sharp difference with the counterfactual scenario of no clean cooking. Also,



**Figure 9.** People in Good Health (Health Score > 0.6): Percentage of Total by Gender (left panel) and Percentage Points Difference from the Baseline Scenario (right panel).

in terms of percentage differences by subgroups, as presented on the right panel of figure 9, we see that almost all subgroups are winners except for a couple of subgroups with almost zero difference with respect to the baseline. In total, in the universal access scenario, we find a 2.1% points increase in the number of females in good health compared to the baseline scenario, while in the no clean cooking scenario there is a 5.1% decrease. This in gross terms implies that having universal access to clean cooking technologies increases the number of females in good health by around 7.2%. As for males, the estimates are slightly lower, with a 1.3% increase in the 'all clean' scenario, and a 5.5% decrease in the 'zero clean' scenario, for a total of 6.8% difference. Finally, in terms of population subgroups, less than high school and high school educated females of African descent living in urban areas are, in absolute terms, the most benefited groups, accounting for 3% of the 7.2% total difference for females between the 'zero-clean' and the 'all clean' cooking scenario.

#### 4. Conclusions and discussion

Our study provides additional evidence that providing households with access to clean energy sources, such as electricity, is not enough to create improvements in health, if reliance on polluting sources for cooking persist. These findings echo that of other studies, that show that simply providing access to clean energy and technologies may not be enough for households to also consistently use these (Malla and Timilsina 2014, Poblete-Cazenave and Pachauri 2018, Vigolo *et al* 2018, Kar *et al* 2019, 2020). We

also find that existing health disparities among population subgroups distinguished by gender, race and educational status in South Africa can be attenuated through the adoption of clean cooking technologies. In urban areas, we find that people are more likely to experience health impacts but with no major differences by race, though there are benefits across educational levels. However, and most importantly, current policies intended to increase electrification and to support the use of clean cooking technologies can bring population health to levels close to those under universal clean cooking access, even if current trends in access improvements continue.

The case of South Africa, which we study here, is distinct compared to many other nations in sub-Saharan Africa, as much of its population has access to clean cooking fuels and technologies already. If the example of South Africa can be mimicked by other countries in the region, where the use of traditional biomass is still widespread, a significant improvement in the health status of the population is feasible. The quantification of the health effects of cooking technologies, as carried out in this analysis, provides important evidence to drive policy makers to expand access to clean cooking more urgently.

Although there is ample evidence of the links between clean cooking technologies and health, quantifying these effects continues to be a challenge, since there is little consensus on how to quantify the health of individuals. Most existing approaches focus on exposure to particulate matter for mortality or morbidity (see e.g. Smith and Pillarsetti (2017) for a recent review of the literature), which is quantified

using CRA methodologies. Our contribution to the literature in this work is to develop a new metric of overall health, which we propose as an alternative way to measure these impacts. We then use this metric in microsimulations of alternative scenarios of clean cooking access.

The biggest advantage of the microsimulation methodology we employ is also its biggest constraint. It requires no additional external inputs for the creation of future scenarios, besides the single, large panel dataset. Unfortunately, panel studies such as the NDIS are not readily available in most regions, especially developing regions, which are the most impacted by transitions towards clean cooking. However, alternative approaches requiring additional external inputs can also be developed, as long as representative datasets, including health, energy, and socio-demographic variables such as the ones used in this study, are available (Aktas et al 2022).

The dataset used in this analysis is, however, not without limitations. For example, some important factors in determining the extent of exposure to indoor air pollution due to cooking are not directly available (e.g. the presence of appropriate ventilation, where is the cooking performed, who is the person mostly in charge of cooking duties, etc). Moreover, information on the health characteristics of children is also not included in the dataset, making our model unsuitable for any analysis of the effects of indoor air pollution on early life stages. Given the numerous channels through which exposure to pollution can affect children's health (Adaji et al 2019, Lee et al 2020, Islam et al 2021), this is an area which requires further research using datasets that capture the health characteristics of children.

Our study highlights the importance of policies to provide household access to clean cooking to improve individual health status. Our findings regarding the importance of clean cooking access for also reducing existing health inequalities by gender, race and education, clearly highlight that measures to encourage regular use of clean cooking fuels and technologies are particularly needed for socially disadvantaged and marginalized groups. In the case of such vulnerable populations, policies that also improve housing infrastructure, ventilation in cooking areas, and education levels can be important to improving the health of individuals. Our results support the need for concerted policies through collaboration across sectors to improve overall health and reduce inequalities in health, particularly since the determinants of ill-health often lie outside the health sector.

### Data availability statement

The data that support the findings of this study are available upon request from the authors.

### ORCID iDs

Arda Aktas  <https://orcid.org/0000-0001-5036-9525>

Miguel Poblete-Cazenave  <https://orcid.org/0000-0001-8280-910X>

Shonali Pachauri  <https://orcid.org/0000-0001-8138-3178>

### References

- Abtahi M, Koolivand A, Dobaradaran S, Yaghmaeian K, Mohseni-Bandpei A, Khaloo S S, Jorfi S and Saeedi R 2017 National and sub-national age-sex specific and cause-specific mortality and disability-adjusted life years (DALYs) attributable to household air pollution from solid cookfuel use (hap) in Iran, 1990–2013 *Environ. Res.* **156** 87–96
- Adaji E E, Ekezie W, Clifford M and Phalkey R 2019 Understanding the effect of indoor air pollution on pneumonia in children under 5 in low- and middle-income countries: a systematic review of evidence *Environ. Sci. Pollut. Res.* **26** 3208–25
- Aktas A, Pachauri S, Poblete-Cazenave M and Sanmartin M 2022 Scenarios for alleviating health disparities in India through the adoption of clean cooking technologies *Working Paper*
- Alberts S C, Archie E A, Gesquiere L R, Altmann J, Vaupel J W and Christensen K 2014 The male-female health-survival paradox: a comparative perspective on sex differences in aging and mortality *Sociality, Hierarchy, Health: Comparative Biodemography: A Collection of Papers* (National Academies Press)
- Arku R E et al 2020 Adverse health impacts of cooking with kerosene: a multi-country analysis within the prospective urban and rural epidemiology study *Environ. Res.* **188** 109851
- Arku R E, Ezzati M, Baumgartner J, Fink G, Zhou B, Hystad P and Brauer M 2018 Elevated blood pressure and household solid fuel use in premenopausal women: analysis of 12 demographic and health surveys (DHS) from 10 countries *Environ. Res.* **160** 499–505
- Bede-Ojimadu O and Orisakwe O E 2020 Exposure to wood smoke and associated health effects in sub-saharan Africa: a systematic review *Ann. Glob. Health* **86** 32
- Caballero F F, Soulis G, Engchuan W, Sánchez-Niubó A, Arndt H, Ayuso-Mateos J L, Haro J M, Chatterji S and Panagiotakos D B 2017 Advanced analytical methodologies for measuring healthy ageing and its determinants, using factor analysis and machine learning techniques: the ATHLOS project *Sci. Rep.* **7** 1–13
- Case A and Paxson C 2005 Sex differences in morbidity and mortality *Demography* **42** 189–214
- Cox D R 1972 Regression models and life-tables *J. R. Stat. Soc. B* **34** 187–202
- Das I, Pedit J, Handa S and Jagger P 2018 Household air pollution (HAP), microenvironment and child health: strategies for mitigating HAP exposure in urban rwanda *Environ. Res. Lett.* **13** 045011
- Dasgupta S, Huq M, Khaliqzaman M, Pandey K and Wheeler D 2006 Indoor air quality for poor families: new evidence from Bangladesh *Indoor Air* **16** 426–44
- Daskalopoulou C et al 2019 Healthy ageing trajectories and lifestyle behaviour: the Mexican health and aging study *Sci. Rep.* **9** 1–10
- de la Fuente J, Caballero F F, Sánchez-Niubó A, Panagiotakos D B, Prina A M, Arndt H, Haro J M, Chatterji S and Ayuso-Mateos J L 2018 Determinants of health trajectories in England and the United States: an approach to identify different patterns of healthy aging *J. Gerontol. A* **73** 1512–18

- de la Sota C, Lumbrellas J, Pérez N, Ealo M, Kane M, Youm I and Viana M 2018 Indoor air pollution from biomass cookstoves in rural senegal *Energy Sustain. Dev.* **43** 224–34
- Dutta S and Banerjee S 2014 Exposure to indoor air pollution and women health: the situation in urban India *Environ. Urbanization ASIA* **5** 131–45
- Edwards J H and Langpap C 2012 Fuel choice, indoor air pollution and children's health *Environ. Dev. Econ.* **17** 379–406
- GBD 2017 Risk Factor Collaborators 2018 Global, regional and national comparative risk assessment of 84 behavioural, environmental and occupational and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017 *The Lancet* **392** 1923–94
- Hystad P et al 2019 Health effects of household solid fuel use: Findings from 11 countries within the prospective urban and rural epidemiology study *Environ. Health Perspect.* **127**, 057003
- IEA 2020 SDG7: data and projections
- Islam S, Rana M J and Mohanty S K 2021 Cooking, smoking and stunting: effects of household air pollution sources on childhood growth in India *Indoor Air* **31** 229–49
- Kar A, Pachauri S, Bailis R and Zerriffi H 2019 Using sales data to assess cooking gas adoption and the impact of India's ujjwala programme in rural Karnataka *Nat. Energy* **4** 806–14
- Kar A, Pachauri S, Bailis R and Zerriffi H 2020 Capital cost subsidies through India's ujjwala cooking gas programme promote rapid adoption of liquefied petroleum gas but not regular use *Nat. Energy* **5** 125–6
- Kollia N, Caballero F F, Sánchez-Niubó A, Tyrovolas S, Ayuso-Mateos J L, Haro J M, Chatterji S and Panagiotakos D B 2018 Social determinants, health status and 10-year mortality among 10,906 older adults from the English longitudinal study of aging: the ATHLOS project *BMC Public Health* **18** 1–9
- Landrigan P J et al 2018 The lancet commission on pollution and health *The Lancet* **391** 462–512
- Lee K K et al 2020 Adverse health effects associated with household air pollution: a systematic review, meta-analysis and burden estimation study *Lancet Glob. Health* **8** e1427–34
- Lin W, Brunekreef B and Gehring U 2013 Meta-analysis of the effects of indoor nitrogen dioxide and gas cooking on asthma and wheeze in children *Int. J. Epidemiol.* **42** 1724–37
- Liu J, Hou B, Ma X-W and Liao H 2018 Solid fuel use for cooking and its health effects on the elderly in rural China *Environ. Sci. Pollut. Res.* **25** 3669–80
- Liu Z, Li J, Rommel J and Feng S 2020 Health impacts of cooking fuel choice in rural China *Energy Econ.* **89** 104811
- Malla S and Timilsina G R 2014 *Household Cooking Fuel Choice and Adoption of Improved Cookstoves in Developing Countries: A Review* vol 6903 (Washington, DC: The World Bank) (<https://doi.org/10.1596/1813-9450-6903>)
- Marois G and Aktas A 2021 Projecting health-ageing trajectories in europe using a dynamic microsimulation model *Sci. Rep.* **11** 1–15
- Montgomery M R, Gragnolati M, Burke K A and Paredes E 2000 Measuring living standards with proxy variables *Demography* **37** 155–74
- Owili P O, Muga M A, Pan W-C and Kuo H-W 2017 Cooking fuel and risk of under-five mortality in 23 sub-saharan African countries: a population-based study *Int. J. Environ. Health Res.* **27** 191–204
- Pimpin L et al 2018 Estimating the costs of air pollution to the national health service and social care: an assessment and forecast up to 2035 *PLoS Med.* **15** e1002602
- Poblete-Cazenave M and Pachauri S 2018 A structural model of cooking fuel choices in developing countries *Energy Econ.* **75** 449–63
- Schofield D, Zeppel M J, Tan O, Lymer S, Cunich M M and Shrestha R N et al 2018 A brief, global history of microsimulation models in health: past applications, lessons learned and future directions *Int. J. Microsimulation* **11** 97–142
- Shupler M, Balakrishnan K, Ghosh S, Thangavel G, Stroud-Drinkwater S, Adair-Rohani H, Lewis J, Mehta S and Brauer M 2018 Global household air pollution database: kitchen concentrations and personal exposures of particulate matter and carbon monoxide *Data Brief* **21** 1292–5
- Smith K R and Pillarsetti A 2017 *Household Air Pollution from Solid Cookfuels and Its Effects on Health* (Washington, DC: The Int. Bank for Reconstruction and Development / The World Bank) ch 7, pp 133–52
- Symonds P, Hutchinson E, Ibbetson A, Taylor J, Milner J, Chalabi Z, Davies M and Wilkinson P 2019 Microenv: a microsimulation model for quantifying the impacts of environmental policies on population health and health inequalities *Sci. Tot. Environ.* **697** 134105
- Verbrugge L M, Wingard D L and Features Submission H C 1987 Sex differentials in health and mortality *Women Health* **12** 103–45
- Verhagen J and Fox J-P 2013 Longitudinal measurement in health-related surveys. A Bayesian joint growth model for multivariate ordinal responses *Stat. Med.* **32** 2988–3005
- Vigolo V, Sallaku R and Testa F 2018 Drivers and barriers to clean cooking: a systematic literature review from a consumer behavior perspective *Sustainability* **10** 4322
- WHO 2018 *Household Air Pollution and Health* (Geneva: World Health Organization)
- WHO 2020 Deaths attributable to household air pollution
- Woolard I, Leibbrandt M and De Villiers L 2010 The South African national income dynamics study: design and methodological issues *Stud. Econ. Economet.* **34** 7–24
- Yu K et al 2018 Association of solid fuel use with risk of cardiovascular and all-cause mortality in rural china *JAMA* **319** 1351–61