Social Indicators Research (2022) 164:1297–1321
https://doi.org/10.1007/s11205-022-02983-w

ORIGINAL RESEARCH

Housing and occupant health: Findings from Vietnam

Phuong Thu Nguyen[1] · Preety Srivastava[1] · Longfeng Ye[1] · Jonathan Boymal[1]

Accepted: 22 July 2022 / Published online: 31 August 2022

Abstract
This paper investigates how poor-quality housing affects Vietnamese individuals' health, measured using the number of sick days, which directly affects productivity and economic growth. Our analysis addresses endogeneity issues generally ignored in previous studies. We provide evidence, robust to various alterations that, the absence of an indoor water tap and homeownership, harmful indoor cooking fuel type, and the poor physical structure of the dwelling have adverse impacts on individuals' health. Our findings indicate that the adverse effects of housing on health arise from a larger spectrum of housing issues in rural areas compared to urban areas. We also find that the intensity in deprived housing conditions has a non-linear relationship with health, suggesting that even some degree of housing assistance can have a beneficial impact on individuals' health. Our study provides useful policy guides and informs healthy living practices.

Keywords Health · housing quality · count data hybrid model

JEL Classification: I14 · I18 · R20

1 Introduction

In a bid to reform health care, the World Health Organization (WHO) has developed a set of housing and health guidelines that aim at 'informing housing policies and regulations at the national, regional and local level and are further relevant in the daily activities of implementing actors who are directly involved in the construction, maintenance and demolition of housing in ways that influence human health and safety' (WHO, 2018). Increasingly, a rising action consensus among most nations is that public investment in the housing sector is essential to improve health (Thomson et al., 2013). These initiatives and national policies

✉ Phuong Thu Nguyen
phuong.nguyen@rmit.edu.au

1 School of Economics, Finance and Marketing, RMIT University, Building 80 Level 11, GPO Box 2476, Melbourne, VIC 3001, Australia

are largely guided by recent research that has identified a number of structural, psychological and social pathways through which housing impacts on health (Shaw, 2004).

The existing evidence, however, is significantly more abundant in developed countries (e.g. Western Europe and North America) but scant in less developed nations (Wang et al., 2019), where housing conditions are generally much poorer (Cairncross et al., 1990; Fuller et al., 1993). Problems such as water supply contamination and household waste removal are often prevalent in the developing world (Cairncross et al., 1990). With the rapid growth of urban populations and impediments in developing infrastructures fast enough to keep abreast of this growth, exploring the effects of housing quality on health in developing nations becomes imperative (Fuller et al., 1993).

In this paper, we examine the adverse impact of poor housing on health in a developing country's context – Vietnam – where a large proportion of residents remain exposed to informal houses with substandard living conditions.[1] Nearly 67% of these informal houses across Vietnam are in semi-permanent, temporary or simple forms, built with flimsy construction materials, and often lack access to basic amenities (UN Habitat, 2014; World Bank, 2015). It is documented that housing conditions in rural areas are more severe: on average, around 9.5% of rural households were reported living in temporary and simple dwellings, compared to 3.1% of their urban counterparts over 2006–2016 (General Statistics Office of Vietnam, 2018). To provide insights into the geographic dimension of the housing-health relationship, we further conduct the analysis on separate samples of urban and rural households.

We aim to understand how housing quality (mainly construction materials, sanitary conditions, cooking fuel, homeownership, and crowding) affects its occupants' health after controlling for a range of potential confounders. We ask two specific questions: (1) which housing quality indicators play a crucial role in influencing occupants' health? and (2) are the health effects of housing quality different across urban and rural dwellers? As an extended analysis, we construct a score of adverse housing conditions to examine the impact of the intensity of deprived housing conditions on health, thus providing insights to policymakers on how health outcomes can be optimised subject to limited resources.

Our study provides supplementary empirical evidence to the growing literature on the relationship between housing and health in developing countries. Firstly, due to data limitations, most empirical studies focus on either indoor or outdoor housing quality indicators; we include both to provide a more comprehensive picture of the housing-health nexus. Secondly, instead of using an aggregated housing deprivation index commonly used in empirical studies, the detailed household survey data enables us to measure various dimensions of housing separately, and identify their effects individually. Lastly, the complex interrelationships between poverty, deprived housing, and poor health make it challenging to guarantee adequate control for confounders since those residing in unsatisfactory housing tend to experience many other deprivations (Wilkinson, 1999). To alleviate any potential estimation bias, we make use of a rich household survey dataset that allows us to control for a host of individual, household and neighbourhood confounders.

[1] In the context of Vietnam, the term 'informal housing' is used for dwellings that are built mainly informally by households and emerging micro-developers who do not possess an official Land Use Right Certificate (LURC), or the Building Ownership and Land Use Certificate (BOLUC). Informal housing also refers to dwellings that are self-built or incrementally constructed without a complete set of permissions (e.g., Building Construction Permit) from the government authorities (UN Habitat, 2014; World Bank, 2015).

Our study is closely related to previous work by Herrin et al., (2013) on Uganda. The main points of departure consist of: (1) we use individual-level data for the analysis as opposed to household-level data; (2) we control for neighbourhood characteristics, including the distance of the dwelling to the nearest hospital/healthcare centre; and (3) our main outcome variable is the number of sick days of each household member in the past 365 days, in contrast to the number of sick days for the "sickest person" of the household in the past 30 days, which we believe is likely to exaggerate the magnitude of the housing-health relationship. Our data is drawn from the Vietnam Access to Resources Household Survey (VARHS), which provides rich biennial longitudinal information on Vietnamese households. We use a hybrid count data regression approach to estimate the empirical model. We provide a thorough discussion on any potential endogeneity we could encounter and the corresponding strategies to address them.

Our main finding is that the absence of homeownership and access to basic sanitary conditions, such as an indoor water tap, have the largest detrimental effects on individuals' health in both rural and urban Vietnam. We also find that other housing quality indicators, such as cooking fuel type, and the physical structure of the dwelling, also impact on individuals' health. Lastly, we find that the intensity in deprived housing conditions has a non-linear relationship with health. From a policy perspective, it is interesting to note how even some degree of housing assistance can have a beneficial impact on individuals' health. For example, assisting individuals with one of their deprived housing conditions can result in about 6 fewer sick days on average in a year, and 8 days if they get assistance with multiple conditions. A set of robustness tests lend confidence to our findings. It is evident that poor housing quality adversely affects health such that promoting better housing can generate health improvements. Therefore, to achieve better health outcomes, the government should also focus on policies and programs that contribute to improving the physical structure of poorly built houses as well as their indoor conditions, including access to water and non-toxic cooking facilities.

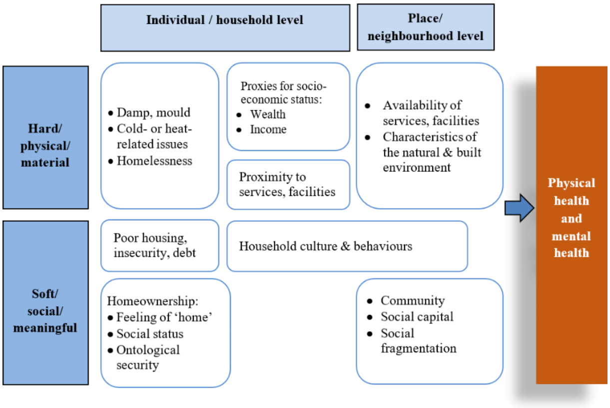
The rest of this paper is organised as follows: Section 2 reviews the literature. Sections 3 and 4 describe the data and estimation techniques, respectively. Section 5 presents the empirical results. Section 6 concludes the paper.

2 Literature Review

2.1 Conceptual model

Taking into account multiple dimensions of housing, Shaw's (2004) conceptual model illustrates how housing can influence health. According to the author, housing quality indicators can be clustered into hard/physical/material versus soft/social/meaningful factors. Figure 1 illustrates how the 'hard/physical/material' related factors (e.g., materials used in construction) are related to health. For example, poor building materials that cause damp, mould, and insufficient heating are often linked with respiratory diseases such as asthma, rhinitis, wheezes and coughs. The 'soft/social/meaningful' housing factors, predominantly homeownership, can affect health through a sense of security, feeling of 'home', and social status versus renting a home where residents may lack a sense of 'control and mastery' (Shaw,

Fig. 1 Ways in which housing can affect health, *Source: Author's adaptation from* Shaw (2004*)*



2004, p.408). This is because housing is perceived as a constituent of general well-being, ontological security, and social status in both individuals as well as community contexts.

Shaw also segregates the factors at individual/household and neighbourhood levels. At the neighbourhood level, factors such as the availability of services, features of the natural and built environment, predominant cultures and behaviours in the surrounding neighbourhood. The sense of community and shared values, and the level of social capital are also likely to affect health.

The conceptual model developed by Shaw forms the basis of our research. Our focus is on the physical conditions of housing, encompassing dwellings' physical conditions and housing tenure, all of which reflect housing quality and which we believe are crucial determinants of health in developing countries. Note however, estimating the whole suite of determinants underlying the conceptual model is beyond the scope of the paper, mostly due to data unavailability. To motivate our choice of housing determinants, we next highlight the key differences in housing quality indicators used in studies in developed and developing countries. To contextualise our study, we also provide a selection of previous research of relevance below. We then formulate our two main hypotheses.

2.2 Indicators of housing quality

The literature on the relationship between housing quality and health mainly draws on developed countries (Evans, 2006). These studies often use a wide range of housing quality indicators that encompass various aspects of housing including physical/structural conditions, space, internal housing facilities and access to amenity/facilities, housing tenure, or they construct a composite quality index based on these multiple indicators (Pevalin et al., 2008; Navarro et al., 2010; Angel & Bittschi, 2017; Špirková et al., 2017). Constrained by inadequate data on housing characteristics and a lack of consensus on appropriate measures of housing quality, there is a relatively smaller number of studies that analyse housing quality and/or its health effects in developing countries (Fiadzo et al., 2001; Montgomery & Hewett, 2005; Meng & Hall, 2006; Ssewanyana & Younger, 2008; Arku et al., 2011; Herrin et al., 2013). A wide review of research in 28 developing countries by Bradley & Putnick (2012) shows that housing quality indicators used in these studies ranged from provision of water, sanitation facilities, cooking facilities, food storage, to construction materials. Most studies had access to only a selection of the indicators.

The main contribution of our study lies in the wide range of housing quality measures that we utilise in the context of a developing country. Our housing quality indicators encompass the structural aspects of a dwelling, sanitary conditions, and housing tenure. By examining the health impact of each separate housing quality indicator, we are able to identify those factors that are crucial for health.

2.3 Effects of indoor housing conditions on health

2.3.1 Dampness and cold homes

Poor building materials that cause damp, mould, and insufficient heating can have adverse impact on both physical and mental health. On the one hand, condensation caused by insufficient air movement or poor ventilation and damp houses aid in the proliferation of mould, fungi, and other potentially harmful bacteria that lead to a number of health conditions including respiratory infections, asthma and allergic conditions (Kasl, 1990; Matte & Jacobs, 2000; Shaw, 2004; Angel & Bittschi, 2017). On the other hand, the structural aspects of the home that result in coldness, dampness, leaky roof, rot in wood, mould, pest infestation, and peeling paint or holes in the floor could also cause neurotic conditions, psychological distress, and mental disorders (Baker et al., 2016; Evans et al., 2003).

2.3.2 Indoor air pullant, and absence of indoor water tap

A number of studies identify the negative effects of indoor air pollution on health, often in the context of less-developed nations.[2] For instance, Dasgupta et al., (2006) demonstrate that cooking indoors with biomass fuel has an effect on indoor air pollution in Bangladesh and emphasise the relative importance of ventilation vis-a-vis fuel choices. In Uganda, Herrin et al., (2013) find that exposure to the burning of biomass for cooking has the largest negative impact on health. In Peru, Gajate-Garrido (2013) finds a significant effect of indoor air pollution on boys' respiratory health.

Lack of indoor tab water access is prevalent in developing countries and has significant health implications. Firstly, an absence of clean water impacts health via water related diseases. Krieger & Higgins (2002), for example, show that an absence of safe drinking water and a lack of hot water for washing are contributing factors to the spread of infectious diseases. Zhang (2012) finds that introducing village-level access to water from water plants in China led to a reduction in the incidence of illness among adults and an increase in weight-for-height among both adults and children. Increased access to safe piped water is found to contribute to child mortality decline (Galiani et al., 2005; Gamper-Rabindran et al., 2010). Secondly, access to clean water also impacts on general wellbeing. Households in developing countries often spend substantial amount of time obtaining water, which can generate

[2] The use of biomass for cooking is not only popular in less developed countries but also exists in poor areas of developed countries (Herrin et al., 2013). For instance, indoor burning for cooking and heating have been found to increase the likelihood of lung cancer in the UK and Eastern Central Europe (Lissowska et al., 2005).

considerable stress and tension for all household members.[3] By decreasing the time burden of water collection, households can spend more time on additional leisure or production and eliminate an important source of stress and tension (Devoto et al., 2012).

2.4 Effects of homeownership on health

Ownership is an important economic and social dimension of housing. As discussed by Munford et al., (2020), in theory it can affect health through various channels including housing conditions, a sense of physical and emotional security, household wealth, labour market and production of social capital. Empirically, as reviewed by Dietz & Haurin (2003), evidence on the link between homeownership and health originated from the record of better health outcomes of home-owners relative to renters in developed countries (notably UK), for example, lower mortality rates (Filakti & Fox, 1995), suicide rates (Ferrada-Noli, 1997), general practitional consultation rates (Benzeval & Judge, 1996), and health conditions (Kind et al. 1998). The effect remained significant even after controlling for potential economic and social cofounders such as incomes (Macintyre et al., 1998). Home-owners also tend to report lower levels of psychological distress and depressions than renters (Kearns and Smith, 1993; Evans et al., 2003; Bloze and Skak, 2012; Bentley et al., 2016). More recently, using exogenous variation in subsidies in England, Munford et al., (2020) reinforced the evidence of a causal effect of homeownership on health.

A number of studies have emerged in less developed countries looking at the relationship between homeownership and health. Arku et al., (2011) investigate the social and economic dimension of housing (e.g. ownership, the fear of being ejected from one's residence) in Ghana, and find that a lack of control on residence has a negative effect on mental health. A similar association between housing and mental health is found by Xie (2019) for rural migrants in urban China. Homeownership has also been associated with better subjective wellbeing (see, Hu 2013; Cheng et al., 2016; Anderson, 2019).

2.5 Neighbourhood effects

Bilger & Carrieri (2013) show that pollution, noise, and crime are health-damaging, and that the impact of these factors is even stronger than the effect of poor economic circumstances. Several studies have also examined the role of geographic proximity to facilities and services (e.g. health care centres and hospitals) in utilisation of health services and mortality. For example, Lavela et al., (2004) find that the further away was the place of residence from health care facilities, the less frequently veterants in the US utilise both outpatient and inpatient health care services. Similarly, Roghmann & Zastowny (1979) conclude that distance to a hospital is a critical determinant of the frequency of utilisation as well as the choice of facilities, even in metropolitan areas. Okwaraji & Edmond (2012) provide a systematic review of 13 studies from low- and middle-income countries and note that the proximity to health facilities is a key predictor of mortality among children under five years old, particularly in the perinatal and neonatal periods.

[3] A study by Kremer et al., (2011) estimates that on average a rural household in Western Kenya makes around seven water-fetching trips per day, with each trip typically requiring a 20-minute walk. Families who depend on public water taps in urban Morocco, devote over seven hours a week to collect water (Devoto et al., 2012).

In summary, a large body of evidence on the health-housing nexus comes from developed countries. Nonetheless, in recent times, a growing number of studies in developing countries have emerged attempting to understand this relationship. Using multi-dimensional indicators of housing quality, we examine the health impact of housing, and identify housing factors that are cucial for health. We also conduct an analysis on separate samples of urban and rural households to test if the health effects of housing are larger and/or arise from more extensive housing conditions in rural areas due to more severely deprived housing conditions. Our two hypotheses are: (1) poor housing quality, as measured by deprived housing conditions and the absence of homeownership, is associated with poorer health; (2) the adverse effects of housing on health are larger and/or arise from a larger spectrum of housing issues in rural areas compared to urban areas.

3 Data

We use the five most recent waves of the VARHS data from 2008 to 2016, co-administered by the United Nations University's World Institute for Development Economics Research (UNU-WIDER) and the Vietnamese Statistical Office. The survey is conducted biennially and targeted to households (predominantly from the Kinh ethnic majority) living in 12 provinces, which provide a good representation of the eight administrative regions of Vietnam.[4] The survey provides a wide range of demographic and socio-economic information on all household members. This allows us to conduct analysis at an individual level.

3.1 Dependent variable

Our outcome of interest is individual health which we measure using the number of sick days (SD) over the last 12 months. This measure proxies for both physical and mental health and since we do not have information on the nature of the sickness, we cannot separate the two health effects. Each household member aged six years and over in the survey is asked the following question: "During the past 12 months, how many days were you not able to perform normal activities due to sickness?" The days of illness range from 0 to 365 days. We restrict sick days to values less than and equal to 60 (i.e. dropping approximately the top 2%) because its distribution is heavily right-skewed.[5]

3.2 Housing variables

We construct several measures of housing quality. First, to capture problems associated with the *physical structure* of a dwelling, we use dummy variables to indicate whether the house has: (i) flimsy surrounding walls, and (ii) a flimsy roof. These are likely to cause damp, mould, hot indoor temperature in the summer, or cold indoor temperature in the winter. Table 1 describes which materials are considered flimsy and sturdy, following the

[4] The 12 provinces include Ha Tay (which was subsumed into the metropolitan area of Hanoi in 2008), Lao Cai, Dien Bien, Lai Chau, Phu Tho, Nghe An, Quang Nam, Khanh Hoa, Dak Lak, Dak Nong, Lam Dong, and Long An.

[5] Accordingly, we only examine $SD \leq 60$ in the main regression. As robustness checks, we use the regression results of the full sample of SD and $SD \leq 30$ (i.e. equivalent to removing 3% of observations).

Table 1 Means of housing variables and income

	Sturdy (S) or Flimsy (F)	Household gross income ('000 VND)	% of households
Surrounding walls			
Earth	F	36,099.69	6.85
Leaves/branches/bamboo	F	41,112.14	3.39
Wood	S	49,870.24	31.58
Galvanized iron	S	66,873.13	0.83
Brick (fired or unfired)	S	85,467.25	54.82
Concrete	S	88,698.47	2.30
Others	S	60,346.28	0.23
Roof	F	39,221.25	3.73
Straw, leaves	F	42,056.38	20.71
Sheeting/panel	F	47,928.30	0.31
Canvas, tar paper	S	52,027.73	0.94
Wood	S	67,164.36	32.91
Tile	S	77,236.43	27.78
Galvanised iron	S	113,440.70	12.65
Concrete, cement	S	48,481.58	0.97
Others			

Source: VARHS 2008–2016

classification method of the General Statistics Office of Vietnam (GSO) in the 2009 Population and Housing Census. For instance, the roofing material of a house that is made of iron sheets/panels often makes the rooms hot in summer but cold in winter, unless electricity is available and a household can afford the high electricity cost (UN Habitat, 2016). To provide insight on how the housing quality relates to individuals' economic status, we also report the corresponding average household gross income in the table. As expected, flimsy or substandard construction materials are more prevalent in houses occupied by households with lower incomes.

Second, a dummy variable is used to indicate whether the household uses *a harmful source of fuel for cooking*. In the survey, cooking sources that are identified as detrimental to health include firewood, charcoal, and coal as they generate toxic substances such as SO_2, NO_2 or CO_2 that directly affect human health. As most Vietnamese households cook indoors, the inhalation of such toxic gases raises the risk of respiratory infections and other health conditions such as pneumonia, asthma, lung cancer. The use of charcoal and coal for cooking is highly prevalent in rural Vietnam.

Third, we construct indicators of the *sanitary conditions* of the household using whether the household has: (i) an indoor water tap, and (ii) an indoor septic/semi-septic toilet. Fourth, we use housing tenure status (i.e. being a homeowner versus a renter) to indicate the willingness and/or ability of households to upgrade their physical housing conditions. Finally, we use housing area per person to capture any overcrowding in the house. The effect of crowding has been found to have both positive and negative effects on health (Fuller, 1993; Pevalin et al., 2008; Navarro et al., 2010; Angel & Bittschi, 2017).

3.3 Control variables

To alleviate time-variant unobserved heterogeneity, socio-economic variables that are commonly found to affect health are controlled for. Specifically, gross household income (deflated using CPI 2010 as the base year), educational attainment, marital status, and working status are included in the model. Working status is particularly useful because those who are retired or fulfilling domestic tasks are likely to be more sensitive to poor housing conditions than employed individuals who spend more time outside the house (Angel & Bittschi, 2017). Other individual characteristics such as age, gender, ethnicity, and a lifestyle variable indicating whether an individual is a current smoker, are also included in the model. Regarding neighbourhood characteristics, two variables are used to measure the proximity to facilities and services, specifically the distance from the house to a: (i) public health centre, and (ii) hospital. These variables capture the effect of infrastructure on health. Table 2 presents the summary statistics of all variables used in the analysis. Our panel sample consists of information on 16,453 urban and rural dwellings and 79,381 observations.

3.4 Preliminary analysis

As a precursor to the empirical analysis, Table 3 looks at some correlations between housing quality indicators and health outcome considered. It can be seen that those with poor quality housing tend to have more sick days within a year compared to those living in good quality housing. These relationships are clearly exhibited in indicators including cooking source, indoor flush toilet, water tap, and housing tenure. In contrast, it is less evident in indicators of the quality of outer walls and roof. This preliminary analysis provides important initial insights into which housing variables are likely to be correlated with number of sick days. However, because of confounders such as income and education, we next estimate the relationship using an econometric model.

4 Model specification

4.1 Econometric framework

Following the conceptual model discussed in Sect. 2, an empirical model is specified to estimate the relationship between individual health and a range of indoor and outdoor housing quality indicators, as follows:

$$Y_{iht} = \alpha + X\prime_{iht}\beta + \phi_i + \phi_h + \phi_t + u_{iht} \tag{1}$$

where Y_{iht} denotes the number of sick days over the last 12 months for the i^{th} individual living in the h^{th} household in year *t*. X_{iht} is a set of explanatory variables which include the housing variables and other individual- and household-level controls; β is a *k x 1* parameter vector. ϕ_i, ϕ_h and ϕ_t denote individual, household and year fixed effects, respectively, and u_{iht} is the idiosyncratic error term.

Given the count nature of the dependent variable, we can estimate Eq. (1) using the Poisson regression model (Cameron & Trivedi, 2013). In particular, the Poisson estimates the *expected* number of sick days in the last 12 months for the i^{th} individual living in the h^{th} household in year *t*, $E[Y_{iht}|X_{iht}]$, conditional on a set of independent variables X_{iht}. By definition:

$$\mu_{iht} = E[Y_{iht} = y\,|X_{iht}] = \exp(X'_{iht}\beta + \phi_i) \tag{2}$$

Table 2 Summary statistics

Variables	Obs.	Mean	Std. Dev.
Number of sick days (SD)			
SD≤60	73,369	4.49	8.41
Full sample	74,793	8.29	32.74
Flimsy outer wall (=1 if true; =0 otherwise)	73,828	0.10	0.30
Flimsy roof (=1 if true; =0 otherwise)	73,813	0.25	0.43
Harmful cooking source (=1 if true; =0 otherwise)	73,825	0.64	0.48
No flush toilet (=1 if true; =0 otherwise)	73,807	0.58	0.49
No water tap (=1 if true; =0 otherwise)	73,820	0.24	0.42
Housing tenure (=1 for renters; =0 for homeowners)	73,808	0.987	0.11
Housing area per capita (m^2)	73,806	16.56	13.56
Household gross annual income ('000 VND)	73,610	69,205	100,840
Working status			
Employed full-time by someone	79,381	0.24	0.43
Doing agricultural work of the household	79,381	0.40	0.49
Doing non-agricultural economic activities	79,381	0.03	0.16
Activities using common property resources[1]	79,381	0.01	0.09
Doing domestic tasks	79,381	0.10	0.30
Pupil, student, retired & other	79,381	0.22	0.42
Educational attainment			
No degree	71,081	0.87	0.34
High school & vocational training	71,081	0.10	0.30
College & above	71,081	0.03	0.18
Marital status			
Single	71,199	0.41	0.49
Married	71,199	0.53	0.50
Widowed / Divorced	71,199	0.06	0.25
Gender (=1 if male; =0 if female)	79,198	0.43	0.49
Ethnicity (=1 if Kinh majority; =0 if ethnic minorities)	73,825	0.51	0.50
Age categories			
Age≤20	79,381	0.39	0.49
21≤Age≤30	79,381	0.16	0.37
31≤Age≤40	79,381	0.13	0.33
41≤Age≤50	79,381	0.13	0.33
51≤Age≤60	79,381	0.09	0.29
Age≥61	79,381	0.10	0.30
Currently smoker			
Yes	72,069	0.59	0.49
No, but used to	72,069	0.22	0.41
Never	72,069	0.20	0.40
Distance to public health centre (km)	73,779	3.08	12.39
Distance to hospital (km)	73,782	16.84	22.61

Source: VARHS 2008–2016

[1]Examples are hunting, fishing in the sea or lakes not on household's property, gathering honey and berries, gathering forestry products

.This parameterisation ensures that the average number of sick days is strictly non-negative.

Table 3 Housing characteristics and health

Housing characteristics	Average sick days[1]	
	Yes	**No**
Whether the house has flimsy outer walls	4.450	4.486
Whether the house has a flimsy roof	4.089	4.611
Uses harmful source for cooking	4.633	4.216
Not having a flush toilet	4.571	4.358
Not having a water tap	4.548	3.838
Rented house	6.091	4.462

Source: Author's calculation from the VARHS 2008–2016

[1]Average number of sick days is restricted to ≤60

Suppose the number of occurrences for Y_{iht} given X_{iht}, is Poisson distributed with density:

$$\Pr\left(Y_{iht} = y|\mu_{iht}\right) = \frac{e^{-\mu_{iht}}{\mu_{iht}}^{y}}{y!} \quad (3)$$

An important feature of the Poisson distribution is the equality of the conditional mean and conditional variance (equi-dispersion). In practice, count data may turn out to be over-dispersed as a consequence of either unobserved heterogeneity, event dependence over time, or excess zeroes. Over-dispersion has implications on the significance of the coefficients, that is, the estimated standard errors are likely to be small. The most common model allowing for over-dispersion is the negative binomial model (NB). The NB model accounts for over-dispersion through an additional parameter $\theta_i \geq 0$ (assumed constant over time for a given individual) that follows a Gamma distribution, such that:

$E[Y_{iht}|X_{iht}] = \mu_{iht}$ and $Var[Y_{iht}|X_{iht}] = \mu_{iht} + \left(\frac{1}{\theta_i}\right)\mu_{iht}^2$ (4)

Given the panel nature of our data, fixed effects regression allows us to control for unobserved variables that are constant over time. However, the unconditional fixed effects estimator for the NB model produces inconsistent estimates because of the incidental parameters problem (see, Allison & Waterman 2002; Greene, 2005).[6] An alternative approach suggested by Allison (2009) is to estimate a random effects NB model with all time-varying covariates expressed as deviations from the household-specific means, termed as the *hybrid* model:

$$Y_{iht} = \beta_0 + \beta_1(x_{iht} - \bar{x}_{ih}) + \beta_2 w_{ih} + \beta_3 \bar{x}_{ih} + \phi_i + \phi_h + \phi_t + u_{iht} \quad (5)$$

In Eq. 5, all variables have been decomposed into a between ($\bar{x}_{ih} = \frac{1}{T}\sum_{t=1}^{T} x_{iht}$) component and a within $(x_{iht} - \bar{x}_{ih})$ component, also referred to as deviation scores. Here, β_1 gives the within-effect or the fixed effects estimates and β_2, the between effects. However, for the estimate of β_2 to be unbiased, $E\left(\phi_i|x_{iht}, w_{ih}\right) = 0$ and $\phi_i|x_{iht}, w_{ih} \sim N(0, \sigma^2)$ has to hold. Following closely the codes developed by Schunck (2013), we estimate the model on Stata 15. Standard errors are clustered at the household level to account for the use of multiple respondents from the same household cluster.

[6] According to Allison & Waterman (2002), the conditional NB model for panel data, proposed by Hausman et al., (1984), and available on Stata, is not a true fixed-effects method and does not control for time-invariant unobserved factors.

4.2 Potential endogeneity

The estimation of Eq. (5) is subject to potential endogeneity arising from both the housing variables and some control variables, such as income and work status. The key coefficients on the housing variables can be biased if individuals' health also influences housing choices directly, and indirectly (via incomes). For instance, poor health may directly prevent individuals from regular maintenance of walls or roofs. Indirectly, poor health reduces income earned, which in turn affects housing choices.

However, there are several reasons why the above arguments do not pose a significant concern for our estimation. In the context of a developing country like Vietnam, and particularly in rural dwellings, multiple household members within a household (often two to three generations) work collectively both to generate income and to maintain housing conditions (Glewwe et al., 2004). Sometimes even relatives outside the household will also help with house maintenance. Thus, one individual's health status is unlikely to affect the housing conditions of the household he/she belongs to. In addition, the survey data encompasses all individuals within a household rather than the household head. This further moderates the potential influence on housing choices resulting from the poor health of the household head, if one considers the household head as the main bread earner.

Turning to the social-economic variables such as income and work status, they are also likely to endogenous. Following Angel & Bittschi (2017) and Krieger & Meierrieks (2019), we use lagged values of household income and work status as instruments. In addition, we use income quintiles (rather than reported income) which has the added advantage of reducing endogeneity resulting from any potential measurement error. Lastly, unobserved time-invariant factors such as genetics, preference, and taste are controlled for using the hybrid modelling approach.

5 Results

5.1 Relationship between housing and health

We estimate the hybrid Negative Binomial (NB) model specified in Eq. 5 above. The hybrid model estimates a set of within-effects (or fixed effects) and between-effects for time varying variables, and can also identify the effects of time invariant variables. Due to space constraint and since our main interest is in the fixed effects, we report only the within-effects coefficients in column 1 in Table 4. The full set of results are available on request. Note that the variables here represent deviation scores [i.e $(x_{iht} - \bar{x}_{ih})$]. However, for simplicity we use the given variable names.[7] Along with the coefficients, we also report the incident rate ratios (IRRs) in column 2 which make interpretation easier. The IRR expresses the incident rate as the relative change in the dependent variable caused by a unit change in the j^{th} independent variable holding all else constant.

Housing tenure appears to have the largest effect on health. The IRR indicates that the number of sick days is 1.47 times larger for a renter than a home owner. In other words, renters are associated with 47% more sick days compared to owner-occupiers. This is in line

[7] A Chi-squared test for over-dispersion rejects the null hypothesis of 'no over-dispersion' indicating that the Negative Binomial model is preferred to the Poisson model.

Table 4 Housing characteristics and health. Panel Analysis

Variables[a]	Dependent variable: Sick Days Negative Binomial (hybrid model)	
	(1)	**(2)**
	Coeff.	**IRR[b]**
Household-level		
Flimsy outer walls	0.119 (0.036)***	1.127 (0.040)***
Flimsy roof	0.081 (0.025)***	1.084 (0.028)***
Harmful cooking source	0.039 (0.022)*	1.040 (0.023)*
No water tap	0.237 (0.033)***	1.268 (0.042)***
No flush toilet	-0.055 (0.022)**	0.947 (0.020)**
Renter	0.387 (0.088)***	1.472 (0.130)***
Housing area per capita	0.004 (0.001)***	1.004 (0.001)***
Income quintiles (Lagged) - Ref. cat.: Quintile 1		
Quintile 2	-0.006 (0.024)	0.994 (0.024)
Quintile 3	0.035 (0.024)	1.036 (0.025)
Quintile 4	0.120 (0.026)***	1.127 (0.029)***
Quintile 5	0.218 (0.028)***	1.243 (0.035)***
Individual-level		
Working status (lagged) - Ref. cat.: Full-time worker		
Agricultural work of the household	0.003 (0.018)	1.003 (0.018)
Non-agricultural economic activities	-0.038 (0.044)	0.963 (0.043)
Activities using common property resources	0.011 (0.077)	1.011 (0.078)
Domestic tasks	0.145 (0.027)***	1.156 (0.031)***
Pupil, student, retired & other inactive person	0.095 (0.024)***	1.100 (0.026)***
Education - Ref. cat.: No degree		
High school & vocational training	0.009 (0.024)	1.009 (0.024)
College & above	-0.022 (0.043)	0.978 (0.042)
Marital status - Ref. cat.: Single		
Married	0.028 (0.028)	1.028 (0.029)
Widowed/divorced	0.176 (0.040)***	1.192 (0.047)***
Male[d]	-0.096 (0.013)***	0.909 (0.012)***
Kinh ethnic majority[d]	-0.141 (0.020)***	0.869 (0.018)***
Age - Ref. cat.: Age ≤ 20		
Age between 21 & 30	0.000013 (0.027)	1.000 (0.027)
Age between 31 & 40	0.016 (0.034)	1.174 (0.040)
Age between 41 & 50	0.359 (0.035)***	1.433 (0.050)***
Age between 51 & 60	0.661 (0.037)***	1.938 (0.071)***
Age 61 and above	1.070 (0.038)***	2.915 (0.110)***
Smoking status - Ref. cat.: Never smoked		
Yes	0.023 (0.021)	1.023 (0.021)
No, but used to	0.071 (0.023)***	1.073 (0.025)***
Neighbourhood		
Distance to health care centre	-0.003 (0.002)*	0.997 (0.002)*
Distance to hospital	-0.006 (0.001)***	0.994 (0.001)***
Individual fixed effects	Yes	
Year fixed effects	Yes	
Constant	-1.282 (0.088)***	0.278 (0.024)***
Observations	48,240	
Log likelihood	-109,725.3	

Table 4 (continued)

Variables[a]	Dependent variable: Sick Days Negative Binomial (hybrid model)	
	(1)	**(2)**
	Coeff.	**IRR[b]**
Dispersion χ^{2c}	200,930.4	
p-value	0.00	
Wald χ^2	845.18	
p-value	0.00	

Source: VARHS 2008–2016

Notes: Dependent variable, sick days ≤ 60. We present all the within and between estimates in Table R1 in the *Appendix*

[a] We estimate the models using the hybrid approach. The variables (except for gender and ethnicity) represent deviation scores and the corresponding coefficients are fixed effects estimates. Robust standard errors clustered at the household level in parentheses. *** p < 0.01; ** p < 0.05; * p < 0.1

[b] IRR shows the relative effect of a unit change in the explanatory variable on $E[Y_{iht}|X_{iht}]$.

[c] Under Ho: $\frac{1}{\theta_i} = 0\,(nooverdispersion)$.

[d] The hybrid model also allows us to estimate the effects of time-invariant variables. For these to be unbiased, the random-effects assumption of orthogonality between time-invariant observable and unobservables must hold

with the broad findings of existing studies that the most beneficial type of housing tenure from a health outcome perspective is owner-occupation (Smith et al., 2004), from which various health benefits are derived (see, e.g. Cairney & Boyle 2004; Zavisca & Gerber, 2016). Notably, homeownership can positively influence both the quality of dwellings and the ontological security (Bentley et al., 2016). Owner-occupiers have better motivation to maintain their houses than renters who may ignore damage as they are more mobile (Dietz & Haurin, 2003), or frequently ejected from rented homes (Arku et al., 2011). Homeownership could also provide people with a sense of physical and emotional security (Shaw, 2004), control over their life and safety which may reflect social comparisons with those who did not make it to the housing ladder, as well as social capital via greater participation in community organisations, neighbourhood and block associations and socialisation (see further discussion in Munford et al., (2020)).

The second largest effect on number of sick days results from access to tap water. Unlike Herrin et al., (2013) who find no significant relationship between bad water and number of sick days, our NB estimates indicate that a dwelling without an indoor water tap is associated with 27% higher number of sick days. Using the mean number of sick days (i.e. 4.49 days) reported in Table 2, the corresponding IRR of 1.27 implies an estimated 1.21 extra sick days per year for a dwelling without an indoor water tap.[8]

In line with several studies which have found an adverse effect of toxic cooking fuels on health, our results suggest that dwellings that use harmful sources for cooking (e.g., firewood or coal) are associated with 4% more sick days per year. In regard to toilet facilities, the evidence in the literature is mixed. While Fuller et al., (1993) find that individuals who have access to exclusive toilet facilities are less likely to visit a health professional, others find no relationship between illness and poor toilet conditions (Ssewanyana & Younger, 2008; Herrin et al., 2013). Interestingly we find a negative and statistically significant effect

[8] The computation is the mean sick days x ($IRR_{flimsy\ walls}$ – 1), or [4.49×(1.27–1) = 1.21].

of flush toilet on sick days. A plausible explanation is that a flush toilet is installed in the house as a consequence of poor health.

As expected the physical structure of a dwelling has an adverse effect on health. Living in a house with flimsy surrounding walls and roof is found to be associated with 13% and 8% more sick days, respectively, than living in a house with sturdy walls and roof. This result suggests that, independent of other sanitary conditions, there is something detrimental associated with poor building materials. These findings are consistent with studies which find that living in a dwelling with a leaky roof affects self-reported health (Navarro et al., 2010) and increases the likelihood of being in poor health (Angel & Bittschi, 2017).

Interestingly, individuals who live in a house with a larger 'housing area per capita' (i.e. less crowded) tend to have more sick days per year. Possibly, this finding suggests that people sharing a dwelling could provide care to sick members of the household, thus reducing the duration of an episode of ill health. Similar findings have been reported in other studies.[9] For example, individuals residing in more crowded households are less likely to visit health professionals in Thailand (Fuller et al., 1993), and are significantly related to lower rates of catastrophic expenditure in Vietnam (Minh et al., 2013).[10]

Following the literature, we control for other related household and individual characteristics which might potentially be associated with illness to help isolate the effect of housing. In this study, we divide lagged income into quintiles to model any *non-linear* relationship between health and income. Also, the use of quintiles has the advantage of reducing the potential measurement error bias, whilst using income at household level reduces the risk of endogeneity (Angel & Bittschi, 2017). Although studies have long demonstrated a positive relationship between income and health - which is in line with the viewpoint that a rise in income increases investments in health-enhancing goods, and that health is a normal good (Grossman, 1972) – some suggest that health may be affected by income distribution within the society (Wilkinson, 1996; Kawachi & Kennedy, 1999). Our results demonstrate that the poorest-income quintile (the base category) is associated with fewer sick days compared to the richest-income quintiles (Q4 and Q5), in contrast to the positive income-health gradient.

Individual characteristics are also significant determinants of health. Specifically, compared to a full-time worker (the base category), those fulfilling domestic tasks and non-economic activities (i.e. pupil, student, retired & other inactive person) have 16% and 10% more sick days, respectively. Widowed/divorced individuals are associated with 19% more sick days per year than single individuals. In terms of age effects, we find only those age above 40 years to have a significant association with sick days. Relative to people aged 20 and under (the base category), those in the age groups 41–50, 51–60 and 61 and above are associated with 43%, 94% and 192% more sick days, respectively. While there is no

[9] Several studies have examined the effect of 'overcrowding' (generally measured by persons per room) on health (Fuller, 1993; Pevalin et al., 2008; Navarro et al., 2010; Angel & Bittschi, 2017). The findings are mixed.

[10] An anonymous reviewer suggested categorising the housing area per capita variable to capture any non-linearity. We estimate the model using 5 categories of housing area per capita. Our results indicate that relative to the lowest category, those living in a house with larger areas per person (thus less crowded) tend to have more sick days per year, with effects on sick days estimated at 10%, 18%, 26%, and 41%, for the respective categories. The effects are all statistically significant at 1% level. The reviewer also suggested looking at the effect of single-person households with the assumption that loneliness may affect their mental health. We find that the effect of 'loneliness' (measured using a dummy indicator for single-person households) is significant, indicating that number of sick days is 49% higher. Note however the share of single-person households only accounts for 0.95% of the sampled households. Both set of results are available on request.

significant association between current smokers and health, we find that those who quit smoking report 7% more sick days relative to those who never smoked, consistent with a lagged impact on health. The hybrid model also allows us to estimate the effects of time-invariant variables, namely gender and ethnicity. Our estimates suggest that males and those from the Kinh ethnic majority are associated with 9% and 13% less sick days than their respective counterparts. Regarding neighbourhood characteristics, it is unclear why a house located further away from a hospital is correlated with less sick days, although the magnitude of its effect is negligible. One plausible explanation is that households with frequently ill member(s) are likely to choose to live in a house located near a hospital for the sake of convenience.

5.2 Rural vs. urban areas

There is evidence that rural residents experience a housing disadvantage, with more severe socio-economic differences compared to urban residents. Specifically, rural residents, despite having more space to live in, tend to face more deprived living standards compared to their urban counterparts (General Statistics Office of Vietnam, 2018). In this section, we evaluate differences in the impact of deprived housing on number of sick days between rural and urban Vietnamese.

Equation (1) is estimated separately for urban and rural households using the NB hybrid model. Results reported in Table 5 provide important insights on the differences. We find that in rural areas all indicators of poorer housing quality are associated with more sick days except for the absence of a flush toilet. On the other hand, in urban areas, while the effects are stronger, the only housing quality indicators that are associated with sick days are flimsy outer walls, absence of tap water and per capita housing area, with health effects of 107%, 164% and 0.006%, respectively.

Homeownership continues to have the largest health effect in rural areas (49.5%), while it is statistically insignificant in urban areas. Possibly, in rural areas of developing countries, homeownership has often been found to be in the form of “de facto tenure”. Even though occupants may not have the legal title or certificate, they would have the actual control of the property. As discussed by Van Gelder (2010), the longer is the length of occupancy and the larger is the size of the settlement, the higher is the level of legitimacy and protection. Given that residents perceive residence as an investment, security and life goal for the family and their heirs (Arima, 1997; Wachter & Megbolugbe, 1992), this de-facto ownership is expected to positively impact on health via the quality of dwellings and ontological security (Bentley et al., 2016). Rural houses in Vietnam are also very likely to be characterised by a similar type of tenure. At the same time housing is more affordable in rural areas (General Statistics Office of Vietnam, 2008) such that those who opt for rental housing usually belong to the lowest income group and live in the cheapest and poorest housing conditions.

Conversely, in urban areas, driven by rapid urbanization, population growth and diverse demands, rental properties are a mix of different types of housing, including those of high quality, supplied by high-end developers. Moreover, the legal property rights in urban Vietnam are in multiple forms (Kim, 2004), akin to the full and partial ownership concepts which differ in security and liquidity in urban China (Cheng et al., 2016). The mix of housing types and ownerships imply that the quality of self-owned dwellings is not necessarily better than those in the rental market. Furthermore, some self-owned houses may not compete with

Table 5 Urban vs. rural differences in the housing-health relationship

Variables [a]	Sick days Negative Binomial (hybrid model)			
	Rural		**Urban**	
	(1)	(2)	(3)	(4)
	Coeff.	IRR	Coeff.	IRR
Flimsy outer walls	0.127***	1.135***	0.729**	2.074**
	(0.036)	(0.041)	(0.288)	(0.598)
Flimsy roof	0.079***	1.082***	0.064	1.066
	(0.026)	(0.028)	(0.149)	(0.159)
Harmful cooking source	0.044**	1.045**	-0.081	0.922
	(0.023)	(0.024)	(0.129)	(0.119)
No water tap	0.195***	1.215***	0.970***	2.637***
	(0.034)	(0.042)	(0.135)	(0.356)
No flush toilet	-0.054**	0.948**	0.010	1.010
	(0.022)	(0.021)	(0.138)	(0.139)
Renter	0.402***	1.495***	-0.266	0.766
	(0.091)	(0.136)	(0.37)	(0.285)
Housing area per capita	0.004***	1.004***	0.006***	1.006***
	(0.00056)	(0.00056)	(0.0022)	(0.0022)
Control variables included	Yes		Yes	
Individual fixed effects	Yes		Yes	
Year fixed effects	Yes		Yes	
Constant	-1.31***	0.27***	-1.68***	0.18***
	(0.091)	(0.025)	(0.45)	(0.084)
Observations	46,456		1,778	

Source: VARHS 2008–2016

Notes: Dependent variable, sick days $\leq$60.

[a] We estimate the model using the hybrid approach. The variables represent deviation scores and the corresponding coefficients are fixed effects estimates. All control variables are same as in Table 4. Robust standard errors clustered at the household level in parentheses. *** $p<0.01$; ** $p<0.05$; * $p<0.1$

rental housing in terms of access to transportation/employment and are likely to negatively impact health via labour markets (Blanchflower & Oswald, 2013; Laamanen, 2017).

5.3 Intensity of housing problems

As shown in the previous sections, poor housing quality are generally associated with poorer health of its residents. In this section, we further explore the relationship by examining the intensity of deprived housing quality, focusing on the physical conditions of the house. In particular, if the government were to provide assistance to improve houses' conditions, by how much will it reduce individuals' sick days in a year? This crude exercise provides important insights to policymakers about how health outcomes can be optimised subject to limited resources. For simplicity, we do not assign any weight or rank to the housing problems. An even more insightful but complex exercise can be conducted by looking at different combinations of the problems.

Here, the 'intensity' variable takes value 0 if a household does not report any problem with their physical indoor and outdoor conditions (i.e. either flimsy walls, flimsy roof, harmful cooking sources, no water tap, or no flush toilet), and values 1 to 5 respectively, for the number of housing problems they report out of the five types. Table 6 shows the distribution of the reported housing problems faced by all households in the sample. Around one-third

Table 6 Observed intensity of deprived housing problems

Intensity (Number of dwelling problems)	Number of households	Frequency (%)
0	804	4.89
1	4,370	26.56
2	3,533	21.47
3	4,534	27.56
4	2,348	14.27
5	864	5.25
Total	16,453	100%

Source: VARHS 2008–2016

Note: Missing values are not reported in the table

of households report at least one housing problem, and just over 5% report all five housing problems. Only about 5% of the sample report not having any of the five issues with their dwellings.

Table 7 reports the regression results, utilising the same estimation approach used earlier. The effects of the number of housing problem on health are estimated between 33% and 82% of additional sick days. We conclude that an increase of housing problems does not necessarily linearly increase sick days. Alternatively, even one housing problem can affect individuals' health as badly as, for example, four housing problems. From a policy perspective, it is interesting to note how even some degree of housing assistance can have a large beneficial impact on individuals' health. For example, if households were to receive assistance with one such problem, this can potentially reduce individuals' number of sick days by 6 on average; an assistance with all 5 problems will result in 8 fewer sick days in a year.[11]

5.4 Robustness checks

We use some tests to evaluate the robustness of our results which we report in Table 8. First, we vary the cut-off point of the dependent variable 'number of sick days', by (1) using the full sample, and (2) removing very ill people who report more than 30 sick days in a year. The results are generally consistent with our main results, with the IRRs being quite similar in magnitudes across all samples.

Second, we test whether our results were driven by provinces where housing deprivation is most prevalent and severe (Panel B). Of the 12 provinces in the sample, Lao Cai and Dien Bien provinces are in the top two on nearly all housing deprivation indicators and are thus omitted from the estimation sample. Once again, our results are generally consistent with the main results. These sensitivity tests lend robustness to our main findings and allows us to generalise our results to a broader population group.

6 Conclusion

Numerous studies have raised concerns that deprived housing quality can adversely affect occupants' health. This paper uses data from the VARHS for the period 2008–2016 to explore the relationship between housing quality (with a focus on construction materials, sanitary conditions, cooking fuel, homeownership, and crowding) and the health of its occupants. To

[11] The numbers of sick days are calculated at the SD mean value of 4.49 days.

Table 7 The impact of the intensity of housing problems on health

Variables [a]	Sick Days	
	Negative Binomial (hybrid model)	
	(1)	(2)
	Coeff.	IRR
Intensity (Ref. cat.: No housing problems)		
Any one problem	0.363***	1.437***
	(0.044)	(0.064)
Any two problems	0.285***	1.329***
	(0.047)	(0.062)
Any three problems	0.363***	1.438***
	(0.050)	(0.071)
Any four problems	0.359***	1.431***
	(0.055)	(0.079)
All five problems	0.596***	1.816***
	(0.065)	(0.119)
Control variables included	Yes	
Individual fixed effects	Yes	
Year fixed effects	Yes	
Constant	-1.155***	0.315***
	(0.094)	(0.030)
Observations	48,245	

Source: VARHS 2008–2016 for SD

Notes: Dependent variable, sick days ≤ 60.

[a] We estimate the model using the hybrid approach. The variables represent deviation scores and the corresponding coefficients are fixed effects estimates. All control variables are same as in Table 4. Robust standard errors clustered at the household level in parentheses. *** p < 0.01; ** p < 0.05; * p < 0.1

the best of our knowledge, this is the first study in Vietnam to examine the health-housing nexus. Moreover, this study departs from most existing literature by employing a different measure of health outcome, namely the number of sick days, which has direct implications on productivity and economic growth.

Using a hybrid Negative Binomial model approach, we find that renting a house (versus owning it) and lack of access to water tap have the largest detrimental health effects. Our finding is consistent with the belief that homeownership provides ontological security and owner-occupiers have better motivation to maintain their houses than renters. Moreover, access to indoor water tap is vital in providing households with an adequate water source for cooking, drinking and personal hygiene, and in reducing the potential health risks of frequently carrying water containers. We also find that other housing quality indicators such as cooking fuel type, and the construction materials also have adverse impacts on individuals' health. A separate analysis on rural and urban households shows that while the health of rural area residents is negatively impacted by most of the housing quality indicators, in urban areas only houses' wall conditions, lack of indoor tap water and crowding have an adverse impact on health. In contrast to our prior expectation, we find that the effects of these factors are larger in urban areas.

Lastly, we conduct a crude exercise to assess the impact of cumulative housing problems on health, focusing on housing physical conditions. Using a score to measure the intensity of poor housing conditions, we find a non-linear impact of the intensity of deprived housing conditions on health. For example, assisting individuals with one of the five poor physical conditions will result in about 6 fewer sick days in a year while assistance with all 5 deprived housing physical conditions will lead to 8 fewer sick days. This basic exercise

Table 8 Robustness tests

Variables [a]	Dependent variable: Sick Days Negative Binomial (Hybrid model)			
	(1)	(2)	(3)	(4)
	Coeff.	IRR	Coeff.	IRR
Panel A: Varying the cut-off points for SD				
	Sick days - Full sample		**Sick days≤30**	
Flimsy outer walls	0.100***	1.105***	0.097***	1.102***
	(0.036)	(0.039)	(0.036)	(0.040)
Flimsy roof	0.078***	1.081***	0.066**	1.068**
	(0.025)	(0.028)	(0.026)	(0.028)
Harmful cooking source	0.031	1.032	0.032	1.033
	(0.022)	(0.022)	(0.023)	(0.023)
No water tap	0.222***	1.48***	0.231***	1.259***
	(0.033)	(0.041)	(0.034)	(0.043)
No flush toilet	-0.038*	0.963*	-0.067***	0.935***
	(0.021)	(0.021)	(0.022)	(0.021)
Renter	0.360***	1.433***	0.376***	1.457***
	(0.088)	(0.126)	(0.090)	(0.131)
Housing area per capita	0.004***	1.004***	0.005***	1.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: Excluding Lao Cai and Dien Bien provinces				
Flimsy outer walls	0.027	1.028		
	(0.046)	(0.048)		
Flimsy roof	0.041	1.042		
	(0.030)	(0.031)		
Harmful cooking source	0.058**	1.060**		
	(0.023)	(0.024)		
No water tap	0.269***	1.308***		
	(0.036)	(0.047)		
No flush toilet	-0.084***	0.919***		
	(0.023)	(0.021)		
Renter	0.339***	1.403***		
	(0.097)	(0.136)		
Housing area per capita	0.005***	1.005***		
	(0.001)	(0.001)		

Source: VARHS 2008–2016

[a] We estimate the model using the hybrid approach. The variables represent deviation scores and the corresponding coefficients are fixed effects estimates. Robust standard errors clustered at the household level in parentheses. All control variables are same as in Table 4. *** p<0.01; ** p<0.05; * p<0.1

provides important insights to policymakers about how health outcomes can be optimised subject to limited resources.

Findings from this study can be used to inform health care policy. It is evident that adverse housing quality have a negative impact on health such that promoting better housing can generate health improvements. Therefore, to achieve better health outcomes, the government should focus on policies and programs that contribute to improving the physical structure of poorly built houses as well as their indoor conditions, including access to water, and non-toxic cooking facilities. For instance, the government could also provide financial support to poor households living in rural areas with dwellings made of substandard materials to rebuild or upgrade their conditions, and develop infrastructure so that communities can benefit from indoor tap water. Our research findings also emphasise the importance of creating awareness in households to cook outdoors if they need to burn biomass. Finally, our findings suggest that even some degree of housing assistance can have a beneficial impact on individuals' health.

While our study can provide useful policy guides and inform healthy living practices, it bears some limitations. First, the health indicators are self-reported and may suffer from recall bias. Second, we only examine the prevalence of poor housing quality and not the extent of the poor conditions. Third, we conduct a basic analysis of the intensity of housing problems based on a simple score measuring the number of deprived housing conditions; future research can examine the effect of different combinations of housing problems on health. This will help develop more tailored prescriptive measures to address housing issues and improve population health.

7 Appendix

Table R1 Full set of main results: Housing characteristics and Health. Panel Analysis

Variables[a]	Dependent variable: Sick Days Negative Binomial (hybrid model)	
	(1)	(2)
	Coeff.	IRR[b]
Within effects		
Household-level		
d_Flimsy outer walls	0.119 (0.036)***	1.127 (0.040)***
d_Flimsy roof	0.081 (0.025)***	1.084 (0.028)***
d_Harmful cooking source	0.039 (0.022)*	1.040 (0.023)*
d_No water tap	0.237 (0.033)***	1.268 (0.042)***
d_No flush toilet	-0.055 (0.022)**	0.947 (0.020)**
d_Renter	0.387 (0.088)***	1.472 (0.130)***
d_Housing area per capita	0.004 (0.001)***	1.004 (0.001)***
d_Income quintiles (Lagged) - Ref. cat.: Quintile 1		
Quintile 2	-0.006 (0.024)	0.994 (0.024)
Quintile 3	0.035 (0.024)	1.036 (0.025)
Quintile 4	0.120 (0.026)***	1.127 (0.029)***
Quintile 5	0.218 (0.028)***	1.243 (0.035)***
Individual-level		
d_Working status (lagged) - Ref. cat.: Full-time worker		
Agricultural work of the household	0.003 (0.018)	1.003 (0.018)
Non-agricultural economic activities	-0.038 (0.044)	0.963 (0.043)
Activities using common property resources	0.011 (0.077)	1.011 (0.078)
Domestic tasks	0.145 (0.027)***	1.156 (0.031)***
Pupil, student, retired & other inactive person	0.095 (0.024)***	1.100 (0.026)***
d_Education - Ref. cat.: No degree		

Table R1 Full set of main results: Housing characteristics and Health. Panel Analysis

Variables[a]	Dependent variable: Sick Days Negative Binomial (hybrid model)	
	(1)	(2)
	Coeff.	IRR[b]
High school & vocational training	0.009 (0.024)	1.009 (0.024)
College & above	-0.022 (0.043)	0.978 (0.042)
d_Marital status - Ref. cat.: Single		
Married	0.028 (0.028)	1.028 (0.029)
Widowed/divorced	0.176 (0.040)***	1.192 (0.047)***
Male[d]	-0.096 (0.013)***	0.909 (0.012)***
Kinh ethnic majority[d]	-0.141 (0.020)***	0.869 (0.018)***
d_Age - Ref. cat.: Age≤20		
Age between 21 & 30	0.000013 (0.027)	1.000 (0.027)
Age between 31 & 40	0.016 (0.034)	1.174 (0.040)
Age between 41 & 50	0.359 (0.035)***	1.433 (0.050)***
Age between 51 & 60	0.661 (0.037)***	1.938 (0.071)***
Age 61 and above	1.070 (0.038)***	2.915 (0.110)***
d_Smoking status - Ref. cat.: Never smoked		
Yes	0.023 (0.021)	1.023 (0.021)
No, but used to	0.071 (0.023)***	1.073 (0.025)***
Neighbourhood		
d_Distance to health care centre	-0.003 (0.002)*	0.997 (0.002)*
d_Distance to hospital	-0.006 (0.001)***	0.994 (0.001)***
Between effects		
Household-level		
m_Flimsy outer walls	-0.090 (0.006)	0.914 (0.055)
m_Flimsy roof	-0.251 (0.045)***	0.778 (0.035)***
m_Harmful cooking source	0.216 (0.046)***	1.241 (0.057)***
m_No water tap	0.242 (0.056)***	1.274 (0.071)***
m_No flush toilet	0.139 (0.045)***	1.149 (0.052)***
m_Renter	0.197 (0.132)	1.218 (0.161)
m_Housing area per capita	0.001 (0.001)	1.001 (0.001)
m_Income quintiles (Lagged) - Ref. cat.: Quintile 1		
Quintile 2	0.014 (0.063)	1.014 (0.064)
Quintile 3	0.023 (0.065)	1.023 (0.067)
Quintile 4	-0.036 (0.067)	0.964 (0.064)
Quintile 5	-0.021 (0.066)	0.979 (0.064)
Individual-level		
m_Working status (lagged) - Ref. cat.: Full-time worker		
Agricultural work of the household	0.021 (0.074)	1.001 (0.074)
Non-agricultural economic activities	-0.220 (0.149)	0.803 (0.119)
Activities using common property resources	0.013 (0.289)	1.013 (0.293)
Domestic tasks	0.380 (0.126)***	1.462 (0.185)***
Pupil, student, retired & other inactive person	-0.323 (0.113)***	0.724 (0.082)***
m_Education - Ref. cat.: No degree		
High school & vocational training	-0.190 (0.088)**	0.827 (0.073)**
College & above	-0.463 (0.121)***	0.630 (0.076)***
m_Marital status - Ref. cat.: Single		
Married	-0.020 (0.108)	0.980 (0.106)

Table R1 Full set of main results: Housing characteristics and Health. Panel Analysis

Variables[a]	Dependent variable: Sick Days Negative Binomial (hybrid model)	
	(1)	(2)
	Coeff.	IRR[b]
Widowed/divorced	0.399 (0.138)***	1.491 (0.206)***
m_Age - Ref. cat.: Age≤20		
Age between 21 & 30	-0.047 (0.120)	0.955 (0.114)
Age between 31 & 40	0.392 (0.158)**	1.479 (0.234)**
Age between 41 & 50	0.412 (0.151)***	1.510 (0.228)***
Age between 51 & 60	0.521 (0.136)***	1.684 (0.229)***
Age 61 and above	1.201 (0.128)***	3.323 (0.424)***
m_Smoking status - Ref. cat.: Never smoked		
Yes	-0.075 (0.049)	0.927 (0.046)
No, but used to	0.078 (0.060)	1.081 (0.065)
Neighbourhood		
m_Distance to health care centre	-0.003 (0.002)	0.997 (0.002)
m_Distance to hospital	-0.002 (0.001)*	0.998 (0.001)*
Individual fixed effects	Yes	
Year fixed effects	Yes	
Constant	-1.282 (0.088)***	0.278 (0.024)***
Observations	48,240	
Log likelihood	-109,725.3	
Dispersion χ^2	200,930.4	
p-value	0.00	
Wald χ^2	845.18	
p-value	0.00	

Source: VARHS 2008–2016

Notes: Dependent variable, sick days≤60. All deviation scores start with a 'd_' and all mean scores start with a 'm_'

[a] We estimate the models using the hybrid approach. The variables (except for gender and ethnicity) represent deviation scores and the corresponding coefficients are fixed effects estimates. Robust standard errors clustered at the household level in parentheses. *** p<0.01; ** p<0.05; * p<0.1

[b] IRR shows the relative effect of a unit change in the explanatory variable on $E[Y_{iht}|X_{iht}]$

[c] Under Ho: $\frac{1}{\theta_i} = 0\,(nooverdispersion)$.

[d] The hybrid model also allows us to estimate the effects of time-invariant variables. For these to be unbiased, the random-effects assumption of orthogonality between time-invariant observable and unobservables must hold

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing Interests: The authors have no competing interests to declare that are relevant to the content of this article.

References

Allison, P. D. (2009). *Fixed effects regression models*. SAGE publications

Allison, P. D., & Waterman, R. P. (2002). Fixed–effects negative binomial regression models. *Sociological methodology*, *32*(1), 247–265

Anderson, J. E. (2019). Real estate ownership and life satisfaction in transition countries.Journal of European Real Estate Research

Angel, S., & Bittschi, B. (2017). 'Housing and health'. *Review of Income and Wealth*, *77*(2), 72–81

Arima, B. C. (1997). The determinants of housing tenure Choice in Ibadan, Nigeria. *Urban Studies*, *34*(1), 105–124

Arku, G., Luginaah, I., Mkandawire, P., Baiden, P., & Asiedu, A. B. (2011). *'Housing and health in three contrasting neighbourhoods in Accra, Ghana'* (pp. 1864–1872). Social Science a Medicine

Baker, E., Lester, L. H., Bentley, R., & Beer, A. (2016). 'Poor housing quality: Prevalence and health effects'. *Journal of Prevention and Intervention in the Community*, *44*(4), 219–232

Bentley, R. J., Pevalin, D., Baker, E., Mason, K., Reeves, A., & Beer, A. (2016). 'Housing affordability, tenure and mental health in Australia and the United Kingdom: a comparative panel analysis'. *Housing Studies*, *31*(2), 208–222

Benzeval, M., & Judge, K. (1996). 'Access to health care in England: continuing inequalities in the distribution of GPs'. *Journal of public health*, *18*(1), 33–40

Bilger, M., & Carrieri, V. (2013). 'Health in the cities: When the neighborhood matters more than income'. *Journal of Health Economics*, *32*(1), 1–11

Blanchflower, D. G., & Oswald, A. J. (2013). *Does high home-ownership impair the labor market? No. w19079*. National Bureau of Economic Research

Bradley, R. H., & Putnick, D. L. (2012). Housing quality and access to material and learning resources within the home environment in developing countries. *Child development*, *83*(1), 76–91

Cairncross, S., Hardoy, J. E., & Satterthwaite, D. (1990). *The poor die young: housing and health in Third World cities*. London: Earthscan

Cairney, J., & Boyle, M. H. (2004). 'Home ownership, mortgages and psychological distress'. *Housing Studies*, *19*(2), 161–174

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (53 vol.). Cambridge University Press

Cheng, Z., King, S. P., Smyth, R., & Wang, H. (2016). 'Housing property rights and subjective wellbeing in urban China'. *European Journal of Political Economy*, *45*, 160–174

Dasgupta, S., Huq, M., Khaliquzzaman, M., Pandey, K., & Wheeler, D. (2006). 'Who suffers from indoor air pollution? Evidence from Bangladesh'. *Health Policy and Planning*, *21*(6), 444–458

Devoto, F., Duflo, E., Dupas, P., Parienté, W., & Pons, V. (2012). 'Happiness on tap: Piped water adoption in urban Morocco'. *American Economic Journal: Economic Policy*, *4*(4), 68–99

Dietz, R. D., & Haurin, D. R. (2003). 'The social and private micro-level consequences of homeownership'. *Journal of Urban Economics*, *54*, 401–450

Evans, G. W., Wells, N. M., & Moch, A. (2003). 'Housing and mental health: a review of the evidence and a methodological and conceptual critique'. *Journal of social issues*, *59*(3), 475–500

Evans, G. W. (2006). 'Child development and the physical environment'. *Annual Review of Psychology*, *57*, 423

Ferrada-Noli, M. (1997). Health and socioeconomic indicators in psychiatric catchment areas with divergent suicide rates. *Psychological reports*, *81*(2), 611–619

Fiadzo, E. D., Houston, J. E., & Godwin, D. D. (2001). 'Estimating housing quality for poverty and development policy analysis: CWIQ in Ghana'. *Social Indicators Research*, *53*(2), 137–162

Filakti, H., & Fox, J. (1995). Differences in mortality by housing tenure and by car access from the OPCS Longitudinal Study.Population Trends, (81), pp.27–30

Fuller, T. D., Edwards, J. N., Sermsri, S., & Vorakitphokatorn, S. (1993). 'Housing, Stress, and Physical Wellbeing: Evidence from Thailand'. *Social Science and Medicine*, *36*(11), 1417–1428

Gajate-Garrido, G. (2013). 'The impact of indoor air pollution on the incidence of life threatening respiratory illnesses: evidence from young children in Peru'. *The Journal of Development Studies*, *49*(4), 500–515

Galiani, S., Gertler, P., & Schargrodsky, E. (2005). 'Water for life: The impact of the privatization of water services on child mortality'. *Journal of Political Economy*, *113*(1), 83–120

Gamper-Rabindran, S., Khan, S., & Timmins, C. (2010). 'The impact of piped water provision on infant mortality in Brazil: A quantile panel data approach'. *Journal of Development Economics*, *92*(2), 188–200

General Statistics Office of Vietnam. (2018). Results of Vietnam Household Living Standard Survey 2016. (Ed.) *Sect. 8: Housing, electricity, water and sanitation facilities*. Hanoi, Vietnam: Statistical Publishing House. General Statistical Office of Vietnam

Glewwe, P., Agrawal, N., & Dollar, D. (2004). *Economic Growth, Poverty, and Household Welfare in Vietnam, World Bank Regional and Sectoral Studies*. Washington, D.C.: The World Bank

Greene, W. (2005). Functional form and heterogeneity in models for count data. *Foundations and Trends in Econometrics*, *1*, 113–218

Grossman, M. (1972). 'On the concept of health capital and the demand for health'. *Journal of Political economy*, *80*(2), 223–255

Hausman, J. A., Hall, B. H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, *52*, 909–938

Herrin, W. E., Amaral, M. M., & Balihuta, A. M. (2013). 'The relationships between housing quality and occupant health in Uganda'. *Social Science and Medicine*, *81*, 115–122

Hu, F. (2013). Homeownership and subjective wellbeing in urban China: Does owning a house make you happier? *Social Indicators Research*, *110*(3), 951–971

Kasl, S. V. (1990). Quality of the residential environment, health, and well being. *Bulletin of the New York Academy of Medicine*, *66*(5), 479

Kawachi, I., & Kennedy, B. P. (1999). 'Income Inequality and Health: Pathways and Mechanisms'. *Health services research*, *34*(1 Pt 2), 215–227

Kim, A. M. (2004). 'A market without the "right" property rights'. *Economics of Transition*, *12*(2), 275–305

Kind, P., Dolan, P., Gudex, C., & Williams, A. (1998). Variations in population health status: results from a United Kingdom national questionnaire survey. *British Medical Journal*, *316*(7133), 736–741

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2011). 'Spring cleaning: Rural water impacts, valuation, and property rights institutions'. *Quarterly Journal of Economics*, *126*(1), 145–205

Krieger, J., & Higgins, D. L. (2002). *Housing and Health: Time Again for Public Health Action*. American Journal of Public Health

Krieger, T., & Meierrieks, D. (2019). 'Income inequality, redistribution and domestic terrorism'. *World Development*, *116*, 125–136

Laamanen, J. P. (2017). 'Home-ownership and the labour market: evidence from rental housing market deregulation'. *Labour Economics*, *48*, 157–167

Lavela, S. L., Smith, B., Weaver, F. M., & Miskevics, S. A. (2004). 'Geographical proximity and health care utilization in veterans with SCI&D in the USA'. *Social Science and Medicine*, *59*(11), 2387–2399

Lissowska, J., Bardin-Mikolajczak, A., Fletcher, T., Zaridze, D., Szeszenia-Dabrowska, N., Rudnai, P., Fabianova, E., Cassidy, A., & Mates, D. (2005). 'Lung Cancer and Indoor Pollution from Heating and Cooking with Solid Fuels: The IARC International Multicentre Case-Control Study in Eastern/Central Europe and the United Kingdom'. *American Journal of Epidemiology*, *162*(4), 326–333

Macintyre, S., Ellaway, A., Der, G., Ford, G., & Hunt, K. (1998). Do housing tenure and car access predict health because they are simply markers of income or self esteem? A Scottish study. *Journal of Epidemiology and Community Health*, *52*(10), 657–664

Matte, T. D., & Jacobs, D. E. (2000). Housing and health—current issues and implications for research and programs. *Journal of Urban Health*, *77*(1), 7–25

Meng, G., & Hall, G. B. (2006). 'Assessing housing quality in metropolitan Lima, Peru'. *Journal of housing and the built environment*, *21*(4), 413–439

Minh, V. H., Phuong, N. T. K., Saksena, P., James, C. D., & Xu, K. (2013). 'Financial burden of household out-of pocket health expenditure in Viet Nam: Findings from the National Living Standard Survey 2002–2010'. *Social Science and Medicine*, *96*, 258–263

Montgomery, M. R., & Hewett, P. C. (2005). 'Urban poverty and health in developing countries: Household and neighborhood effects'. *Demography*, *42*(3), 397–425

Munford, L. A., Fichera, E., & Sutton, M. (2020). 'Is owning your home good for your health? Evidence from exogenous variations in subsidies in England'. *Economics and Human Biology*, *39*, 100903

Navarro, C., Ayala, L., & Labeaga, J. M. (2010). 'Housing deprivation and health status: evidence from Spain'. *Empirical Economics*, *38*, 555–582

Okwaraji, Y. B., & Edmond, K. M. (2012). 'Proximity to health services and child survival in low-and middle-income countries: a systematic review and meta-analysis'. *British Medical Journal Open*, *2*(4), e001196