Fuel-choice and indoor air quality: a household-level perspective on economic growth and the environment*

Shubham Chaudhuri and Alexander S.P. Pfaff
Department of Economics
and
School of International and Public Affairs
Columbia University

January 2003

Abstract

The fuel-use decisions of households in developing economies, because they directly influence the level of indoor air quality that these households enjoy (with its attendant health effects), provide a natural arena for empirically assessing latent preferences towards the environment and how these evolve with increases in income. Such an assessment is critical for a better understanding of the likely effects of aggregate economic growth on the environment. Using household data from Pakistan we estimate Engel curves for traditional (dirty) and modern (clean) fuels. Our results provide empirical support for a household production framework in which non-monotonic environmental Engel curves can arise quite naturally. Under plausible assumptions about the emissions implied by fuel use, our estimates yield an inverted-U relationship between indoor air pollution and income, mirroring the environmental Kuznets curves that have been documented using aggregate data. We then demonstrate, through a simple voting model, that this household-choice framework can generate aggregate EKCs even in a multi-agent setting with heterogeneous households and purely external environmental effects.

JEL codes: D11, H41, O12, Q25

---

[*] We would like to thank the World Bank and the Pakistan Federal Bureau of Statistics for allowing us access to the Pakistan Integrated Household Survey data. We would also like to thank Steve Cameron and Matt Kahn, as well as participants at AEA, NBER, NEUDC and Harvard Environmental Economics and Policy seminars for comments. Corresponding author: Shubham Chaudhuri, Columbia University MC-3308, New York, NY 10027. E-mail: sc301@columbia.edu.

1. Introduction

Cross-country empirical analyses by the World Bank (1992), Grossman and Krueger (1995) and others have brought to the fore the possibility of non-monotonic relationships between income and environmental quality.[1] In particular, the aggregate cross-country evidence seems to suggest that while economic growth might initially degrade the environment, continued growth reverses any initial adverse effects. Not surprisingly, these 'Environmental Kuznets Curves' (or EKCs) have generated considerable debate. Many authors have questioned the robustness of these initial, aggregate findings, arguing that the relationship between environmental degradation and economic growth is sensitive to the pollutants studied and the data used.[2] Others have noted the range of potentially confounding effects not incorporated in these analyses. For instance, if environmental outcomes for other, linked regions such as trading partners are not taken into account, interpretations based on results for one location can be misleading.[3]

From a policy perspective, it is clearly important to establish the validity and generality of these aggregate empirical findings. But it is also crucial to obtain a better understanding of the channels through which the claimed reversal might arise. The most common (though by no means the only) story is one that focuses on the role of environmental regulations.[4] In this account, income growth is seen as having two opposing effects on environmental quality. On the one hand, the higher levels of output associated with income growth lead, all else equal, to higher levels of pollution. On the other hand, as incomes rise, there is increased public demand for environmental quality. With reasonably well-functioning democratic institutions, this results in increased environmental regulation, which, if effective, has the potential to reverse the environmental degradation that would otherwise accompany higher levels of output. If at low levels of income the first (negative) effect dominates, while at higher income levels the second (positive) effect dominates, the result is an aggregate EKC.

Though this account of how an aggregate EKC might arise seems intuitive enough, neither its theoretical nor its empirical underpinnings have been developed in much detail in the existing literature. The aim of this paper is to flesh out some of these details. On the empirical front, central to the story is the implicit assumption that households' preferences are such that at low levels of income the demand for environmental quality is outweighed by the demand for consumption of other goods and services, with the relative weights being reversed at higher levels of income. There is little direct evidence of this in the existing literature and not surprisingly so, since the empirical analyses have almost exclusively been based on aggregate cross-country data.[5]

---

[1] See Selden & Song (1994), Shafik (1994) and Holtz-Eakin & Selden (1995).

[2] See the special issues of *Environment and Development Economics* (November,1997) and *Ecological Economics* (May, 1998), Millimet and Stengos (1999), Harbaugh et al. (2000), and Taskin and Zaim (2000).

[3] See, e.g., Saint-Paul (1995). Pfaff (2001) discusses New England forests, which fell significantly with economic growth and then returned. That agriculture shifted to the Midwest, timber shifted to the Lake States and Northwest, and both were then exported to New England seems crucial.

[4] Frankel and Rose (2002), though it deals with a separate (but related) question, provides a nice summary of the conventional wisdom regarding EKCs.

[5] There are exceptions, for instance, the innovative studies by Deacon and Shapiro (1975), Fischel

In terms of the theory, there is an extensive literature in the neoclassical growth tradition that explores the optimal intertemporal tradeoffs between current consumption, investment in productive capital and pollution control.[6] But, because these papers rely on a representative agent framework they do not, naturally, develop a political-economic mechanism through which degradation might in reality be reversed, for instance, through voting on environmental regulations by heterogeneous voters.[7]

In this paper we try to fill these gaps. We do two things. First, using household-level data we provide evidence that household preferences regarding environmental quality are consistent with those implicitly assumed in the conventional account summarized above. Second, we demonstrate, through a simple illustrative voting model, how such preferences might generate aggregate EKCs in a setting with purely external environmental effects and multiple heterogeneous agents.

The main difficulty with empirically assessing how households' marginal valuations of the environment evolve with increases in household income is that in most cases, the environmental impacts of household choices are external to the household. Observed household choices will therefore not reflect those impacts. A household might value cleaner air but not curb its pollution, no matter its income, given that this pollution has a vanishingly small impact on the air quality the household enjoys. This makes it difficult to estimate the effect of income on household demand for environmental quality.

Our empirical innovation, given this difficulty, is to focus upon the fuel-choice decisions of households in a poor developing economy, in particular the choice between dirtier traditional fuels (wood, dung and other biomass) and cleaner modern fuels such as kerosene and natural gas. These choices directly influence the levels of indoor air quality enjoyed by the households, and because indoor air quality is first and foremost a private good, they can be expected to reflect pollution impacts. Though most published indicators of air pollution are measures of outdoor or ambient air quality, studies suggest that perhaps eighty percent of world population exposure to particulates occurs indoors in developing countries as a result of emissions from biomass fuels used for cooking (Smith 1993, p.551). While smoke does exit the household, with external impact, even in settings where ambient air quality is low, indoor air pollution is often the dominant source of exposure because of the

---

(1979) and Kahn and Matsusaka (1997), which use voting data on environmental referenda to explore how the demand for environmental quality varies with income. These studies provide an useful complementary approach to the one we pursue in this paper.

[6] A typical example is Gruver (1976), which extends the standard neoclassical growth model by incorporating the portfolio choice between investments in productive capital and pollution control capital. Under certain parameter configurations the optimal growth path is shown to be unbalanced. The emphasis in the initial stages of growth is on the accumulation of productive capital, which implies increasing levels of output and pollution. However, once a target stock of productive capital is reached, savings are shifted towards pollution control capital. See also John & Pecchenino 1994, Selden & Song 1995, Stokey 1998, Chimeli 2001, and for related work also Plourde 1972, Keeler et al. 1972, D'Arge & Kogiku 1973, Forster 1973, Gruver 1976, Stevens 1976, Asako 1980, Becker 1982 and Tahvonen & Kuuluvainen 1993.

[7] Others have shown that changes in the composition of goods consumed and techniques of production can matter (Copeland and Taylor 1995, Jaeger 1998, Grossman 1995.) Still others have focused on single-actor stories about preferences and abatement technologies, which to yield EKC predictions require explicit aggregation through identified mechanisms.(Andreoni and Levinson 2001; Pfaff, Chaudhuri and Nye 2002b, which explicitly models a voting mechanism).

close proximity and extended exposure of the individuals to the sources of indoor air pollution.[8] And such exposure has severe health consequences.[9] With health effects of this magnitude, and other disutility from fuels' emissions (such as smoke in one's eyes), there is good reason to expect to observe relevant income effects upon fuels choices, since households do vary in income and can be expected to value health gains from lower emissions.

Using household-level data from Pakistan, we estimate Engel curves for traditional (dirty) and modern (clean) fuels. The estimation approach we adopt is motivated by the household choice framework detailed in Pfaff, Chaudhuri and Nye (2002a), which emphasizes two distinctive features of indoor air quality (and most environmental amenities): that households begin with an *endowment* of indoor air quality, and that indoor air quality *cannot be directly purchased* but is instead degraded through the consumption of marketed fuels, with some fuels being cleaner than others. What this framework makes clear is that Engel curves for marketed fuels may well be non-monotonic, and that the decisions about whether to use a fuel, and how much to use, may react quite differently to income. Specifically, higher income lowers the likelihood of using dirtier fuels but conditional on it being used, raises the quantity used whereas in the case of cleaner modern fuels, increases in income raise both the probability of use and the quantity used.

To allow for these differential effects we adopt a "generalized Tobit" approach to the estimation of the Engel curves. Consistent with previous evidence on "energy ladders", we find that as incomes rise, households transition from traditional biomass fuels to cleaner modern fuels. More importantly, we find evidence of the differential effects of income during the transition that are suggested by the theory. Further, we find that the influence of household size and composition on fuel choice and fuel use decisions is broadly consistent with our claim that these decisions reflect concerns about indoor air quality.

We then turn to directly estimating the effects of income on indoor air quality. Because we do not observe indoor air quality we resort to a hybrid strategy of estimation with partially simulated data. Under plausible assumptions on the ratio of traditional-fuel to modern-fuel emissions we find an inverted-U relationship between levels of indoor air pollution and income. In other words, increases in income appear

---

[8] Smith (1993, p.541) makes this point clear: coal-fired power plants in the United States produce 1.6 kg of particulates per person, while cigarettes produce only 50g per person; however, as cigarette smoke is released so close to the lungs, and is often trapped in the same volume of indoor air breathed by many people for many hours, it is thousands of times more likely to reach people's lungs, and produces exposures eighty times as high. Smith (1987, p.145) points out: "...the exposures and nominal doses of major pollutants found in biofuel smoke rival or exceed those received by active smokers for some pollutants."

[9] For instance, from Smith 1987, p.vii: "...every day...14,000 children die from respiratory infection. The majority of these deaths result from severe acute respiratory illness (ARI)...those who survive ARI...will be more susceptible to respiratory disease throughout their lives and are more likely to suffer chronic obstructive lung diseases. ...exposure to emissions from biomass cooking and heating fuels is an important contributing factor". See also Wilson & Spengler (1996). For instance, in a chapter in this volume, Dockery and Pope estimate that daily mortality increases 0.5-1.6% for each 10 microgram per cubic meter increase in particulate concentrations. Pope (1989) describes a case in which during the winter after a labor dispute had resulted in the closure of a local steel mill—which had been the largest single source of particulate air pollution—PM10 concentrations averaged 51 units compared with a mean of 90 the winter before, and children's hospital admissions for respiratory disease dropped by more than 50% compared with the previous year.

to be associated *initially* with a *deterioration* in indoor air quality as consumption of energy services rises. Only after household income crosses a threshold do subsequent increases in income lead to investments in cleaner fuels that yield reductions in emissions and improvements in indoor air quality.

The implications of this household-level empirical evidence are obvious in the case of environmental amenities (such as indoor air quality) that have a significant private component. If the relationship between income and environmental quality is non-monotonic at the household level, clearly, aggregate income growth and the accompanying changes in the cross-sectional distribution of income can yield a non-monotonic relationship between aggregate environmental quality and aggregate income.[10] But the significance of our findings extends beyond this narrow class of environmental amenities. We demonstrate this by developing a simple illustrative voting model, where we retain the basic elements of the household-choice framework that motivated the empirical analysis, but extend the framework to a multi-agent setting with heterogeneous agents and purely external environmental effects. We show that the characterization of preferences for which we find empirical support can generate an aggregate EKC even in this setting.

The paper is organized as follows. Section 2 briefly summarizes the household-production model from Pfaff, Chaudhuri and Nye (2002a) that seems appropriate for this setting. Section 3 describes the data and presents descriptive statistics. The data come from the Pakistan Integrated Household Survey (PIHS), which includes an energy module that permits study of fuel choices and fuel use as a function of income and other household characteristics. Section 4 details our econometric strategy, discusses various issues that arise in the estimation of fuel-use Engel curves using data from a developing economy, and presents the estimates of fuel-use and indoor air quality Engel curves. Section 5 then demonstrates the utility of such household-level results with an illustrative model of how household preferences could aggregate through voting for environmental policies. Section 6 concludes.

2. Household production model

We begin with the observation that many environmental services (including indoor air quality) cannot be directly purchased. Rather, households are endowed with positive amounts of these amenities, which are then degraded due to the use of marketed commodities. In many developing economies, the consumption of firewood or kerosene results in the joint production of services that households value (e.g., cooking, heating and lighting) and reductions in indoor air quality. We formalize this point within a household-production, or characteristics framework.[11] We use the simplest possible model to demonstrate that non-monotonic environmental Engel curves may arise, and to motivate our estimation approaches.

Let s denote a household's consumption of cooking services, and a denote the level of indoor air quality. Neither s nor a can be directly purchased. Instead,

---

[10] There is, of course, no reason why such a relationship has to be U-shaped, and it might in fact take on any number of shapes depending on the ways in which income growth is distributed. That turns out to be an attractive property given the contested aggregate evidence.

[11] Classic early references in the household production literature include Gorman (1980), Becker (1965), and Lancaster (1966a and 1966b). As our model is not the focus here and is the subject of Pfaff, Chaudhuri and Nye (2002a), see that paper for extended discussions.

they are jointly produced (in the case of a, degraded) through the use of marketed fuels. Consider a situation in which households have a choice between two marketed fuels, a "dirty" (more environmentally destructive) traditional fuel t and a "clean" modern fuel m. Assuming that s is generated linearly from the use of these fuels, we redefine the units in which the fuels are measured so that total cooking services are given by:

$$s = q_t + q_m \tag{2.1}$$

where q_t and q_m are the quantities used of the dirty and clean fuels respectively. We assume that the level of indoor air quality enjoyed by the household is given by:

$$a = \overline{a} - b \tag{2.2}$$

where $\overline{a} > 0$ is the initial *endowment* of air quality and b is the level of indoor air pollution. In this simple model (though not in our later empirical work), we assume b is linear in total emissions, which are themselves linear in the quantities used of the purchased fuels:

$$b = \alpha q_t + \beta q_m \tag{2.3}$$

Here $\alpha > \beta > 0$, i.e., the use of either fuel leads to pollution, but fuel t is dirtier.

The household chooses the marketed q to maximize (2.4) subject to (2.5):

$$U(s, a) \tag{2.4}$$

$$p_t q_t + p_m q_m = y \tag{2.5}$$

where y is household income and p_t and p_m are, respectively, the per-unit (of cooking services) prices of the dirty and clean goods, with $p_t < p_m$. We assume $U(.)$ is increasing and concave in both arguments and that preferences are such that, were households able to directly purchase both s and a, the demands for both would be normal. We assume, moreover, that $\lim_{s \to 0} U_s(s, a) = +\infty$.

With these assumptions, it is straightforward, though somewhat tedious, to show that the household's optimal choice of q will be such that at low levels of income only the dirty fuel is used, at intermediate levels of income, both fuels are used with the share of the cleaner fuel rising with income, and finally at high levels of income, only the clean fuel is used.[12] That immediately raises the possibility that the relationship between household income and environmental quality may be non-monotonic since the slope of the household Engel curve for indoor air quality a is given by:

$$\frac{da(q^*(y))}{dy} = \sum_j \left(\frac{\partial a(q^*)}{\partial q_j}\right) \frac{\partial q_j^*}{\partial y}(y) \tag{2.6}$$

Clearly, with dirty fuels being inferior after a certain income, while clean fuels are normal throughout, it is possible that the Engel curve for indoor air quality will be U-shaped at least over a range of incomes. More generally, the household-level relationship between income and indoor air quality could take on any number of shapes, but this indeterminacy is arguably an attractive property given the contested aggregate evidence.[13]

---

[12] Details are available in Pfaff, Chaudhuri and Nye (2002a).

[13] Note, in particular, that in this simple model, once a household specializes completely in the clean fuel, subsequent increases in income can only lead to a deterioration of air quality.

Figure 1 provides some intuition for how the non-monotonicities arise. The endowment ($s = 0$, $a > 0$) is at the upper left. Each dashed ray depicts the combinations of indoor air quality a and cooking services s attainable through exclusive use of one of the fuels. The solid lines connecting the rays are budget constraints; larger budgets are further from the endowment. The budget slopes indicate the relative shadow price of air quality and cooking, i.e. the rate at which households can trade these two off, given the underlying technologies and the prices of the marketed fuels. The negative slope reflects our assumption that dirtier fuels are cheaper than cleaner fuels per unit of cooking service produced. The shape of the indifference curves comes from the concavity of the utility function.

Figure 1 shows the optimal consumption points of the household at six levels of income. The two transitions from point A to point C involve degradation of environmental quality, at first through increased use of only the dirty fuel and then while the clean fuel is also used. Juxtaposing the indifference curves with the budget sets shows why in the lowest income transition from A to B, while the household could substitute, it does not desire any of the clean fuel. Because the endowment is so skewed towards air quality, moving as rapidly as possible to greater balance of s and a is preferable. This dictates using only the dirty fuel. However, as income continues to rise, the household does begin to use the clean fuel, at first in combination with the dirty fuel and then exclusively. As the share of the clean fuel rises, environmental quality, which had been deteriorating, begins to improve. Eventually the household transitions into exclusive use of the modern fuel, at which point further increases in income must yield reductions in environmental quality.

What does this behavior imply for the estimation of fuel-use Engel curves? Consider Figure 2, which illustrates the fuel-use Engel curves for a single household. In estimating such Engel curves from the observed fuel-usage levels of a cross-section of households, we would obviously need to incorporate a household-specific random error term that potentially shifts around the threshold income levels at which various transitions occur. From Figure 2 it is clear then that if this framework provides a reasonable model of actual behavior, the likelihood of using the cleaner fuel ought to rise with income, as should the quantity used. On the other hand, at higher incomes the likelihood of using the dirtier fuel ought to be lower, but conditional on it being used, there should be a range of income where the quantity used *rises with income*. Therefore, the strategy used to estimate the Engel curves ought to be flexible enough to permit this differential effect of income on the fuel-choice and fuel-use decisions.

Lastly, anticipating the more general model that we present in Section 5, we note here that fuel switching (from cheaper dirtier traditional fuels to more expensive but cleaner modern fuels) represents a particular form of (pollution) abatement by households, with the "abatement expenditures" being the increased cost of obtaining services from cleaner fuels. This can be seen quite clearly by rewriting the household's budget constraint, (2.5), in the following form:

$$p_t s + x = y \tag{2.7}$$

where $x = [p_m - p_t]q_m$ is the level of abatement expenditures that the household chooses to incur. The level of indoor air quality enjoyed by the household can, correspondingly be rewritten as:

$$\begin{aligned} a(s,x) &= \overline{a} - b(s,x) \\ &= \overline{a} - \alpha s + \left[\frac{\alpha - \beta}{p_m - p_t}\right] x \end{aligned} \tag{2.8}$$

and the household's problem therefore becomes one of choosing s and x to maximize $U(s, a(s, x))$ subject to (2.7) and (2.8) and the relevant non-negativity constraints.

3. Data source and descriptive statistics

The data for this study come from the Pakistan Integrated Household Survey (PIHS) 1991.[14] The PIHS is a national survey, although it oversamples urban households.[15] In addition to the the standard socioeconomic and demographic variables found in most household surveys—i.e., age, educational attainment, employment status of household members, household size and composition, consumption expenditures and income—the PIHS provides detailed information on the sources and levels of energy consumption. For each of the 4,800 households the PIHS indicates whether the household uses any of what we will call *traditional* fuels (dung, firewood, and biomass) and/or any of what we will call *modern* fuels (kerosene, LPG, and natural gas), and in particular whether these are used for cooking.[16] Moreover, several questions permit calculation of the quantities of these fuels used for cooking (see Appendix A for details).

Table 1 presents some general descriptive statistics. Means are presented for the whole sample, as well as conditional on living in an urban or rural area, or on being relatively rich or poor (above or below the median per-capita expenditure level). Household incomes are over 75% higher in urban areas than in rural. Urban homes have more rooms, and are less likely to be windowless. In terms of fuel use (at all and for cooking), urban dwellers are more likely to use electricity, natural gas and LPG, and less likely to use firewood, dung and biomass. Finally, urban households are much less likely to collect their fuels (either wood or dung).

Richer households are smaller than poorer (due to fewer children). Their homes have more rooms, and are less likely to be windowless. They are four times as likely to use natural gas, and twice as likely to use LPG, but less likely to use firewood, dung, and biomass (either at all or for cooking). Finally, richer households are significantly less likely than poorer households to collect their fuels.[17]

---

[14] The PIHS was designed and implemented jointly by the Federal Bureau of Statistics, Government of Pakistan and the World Bank, and is one of a number of Living Standards Measurement Study (LSMS) household surveys conducted in various developing economies with the assistance of the World Bank. The purpose of these studies is to provide policy makers and researchers with individual, household, and community level data that facilitate analysis of the impact of policy initiatives on household living standards.

[15] The sample was drawn using a multi-state stratified sampling procedure from the Master Sample Frame developed by the Federal Bureau of Statistics (FBS) based on the 1981 census. It covers all four provinces, and according to the FBS, the areas excluded contain only about 4% of the national population. The sample frame consists of three main domains (self-representing cities, other urban areas, and rural areas), which are exclusively and exhaustively divided into primary sampling units.

[16] As seen in Table 1, note that while many households use electricity, almost none use it for cooking (it is used mostly for lighting). Coal and charcoal were also not used for cooking.

[17] Missing from these statistics is indoor air quality. That is because it is not observed directly

Tables 2 and 3 provide more detailed descriptive statistics on the use of fuels for cooking, which is the focus of our empirical analysis. The top panel of Table 2 indicates that for each of the traditional fuels, the proportion of housholds using the fuel declines monotonically from the poorer to the richer quartiles.[18] In contrast, for each of the modern fuels, the incidence of use increases from the poorer to richer quartiles though there is some variation across the fuels. Kerosene usage rises at earlier incomes and then levels off, while LPG and natural gas rise more in the upper quartiles, suggesting that kerosene might be a "transitional" fuel.

The bottom panel of Table 2 presents information on mean fuel-usage levels, conditional on a particular fuel being used at all. For all the modern fuels, these conditional means generally rise going from the poorer to the richer quartiles, mirroring the pattern of fuel-use incidence. In the case of the traditional fuels, however, the pattern is the opposite of that in the top panel—conditional on use, the average quantity used rises with household expenditure.

Table 3 provides statistics on fuel use, by aggregated fuel categories and expenditure deciles. Using conversion factors from HESS (1993) we converted our various fuel-quantity measures into a common energy unit, megajoules (or BTUs), and then summed up the BTUs deriving from all fuels in each of the two aggregated fuel categories: traditional and modern. These aggregated statistics present much the same picture as in Table 2: the incidence of traditional fuel use declines monotonically from the poorest to the richest expenditure decile while that of modern fuels rises. But, conditional on use, mean levels of consumption of both modern and traditional fuels rise as households grow richer. Moreover, the proportion of households that use both traditional and modern fuels is highest in the middle deciles. Overall, these summary statistics provide prima facie evidence consistent with the framework in Section 2.

4. Econometric issues and estimation

This section details our econometric strategy and presents our empirical results. We begin by estimating Engel curves for modern and traditional fuels. To allow for the differential effects of income on the probability of use and the level of use, we adopt a "generalized Tobit" approach to the estimation of the Engel curves. We then turn to directly estimating the effects of income on indoor air quality. Because we do not observe indoor air quality we resort to a hybrid strategy of estimation with partially simulated data.

Though fairly straightforward in principle, the practical implementation of our empirical strategy is complicated by a number of issues, at least some of which are specific to the Pakistani context and to the data we use. We discuss these issues before proceeding to the estimation.

(and thus is not in the survey). It is produced through household choices such as of fuel type and quantity, and must be estimated from the fuel-use and engineering-technologies information.

[18] The fact that even in the richest quartile 48% of the households continue to use firewood probably indicates the lack of access to modern fuels in many rural areas of Pakistan. Of the 34% of the households in this quartile who reside in rural areas, nearly 90% use firewood.

4.1. Complications

4.1.1. Defining access

In our data the use of modern fuels for cooking is extremely rare among rural households and among the poor. For instance, only 1% of rural households use natural gas for cooking, while only 4% use LPG (i.e., cylinder) gas. The issue here is that *access* to modern fuels is limited in many parts of Pakistan, especially rural areas that simply lack natural gas connections or supplies of LPG. Some households, regardless of income, do not have the option of using modern fuels.[19] And in some of the largest cities, traditional fuels such as firewood and dung may not be readily available. Because the inclusion of households who lack access can potentially obscure the effect of income, these households need to be identified and dropped.

There is, unfortunately, no clean way of identifying them in our data. The survey has a question on access and availability for each fuel but a large number of those responses are missing. In the case of electricity, only 17% of the households reported having access—the other responses being missing or negative—whereas 76% of households are observed using electricity. Simply dropping households who do not use a particular fuel is not a solution, since of course there will be households who have access to that fuel but simply choose not to use it.

We resolve the issue of access by dropping those households who live in 'areas' where no households use the fuel in question. Our definitions of 'area' are based on the sampling frame for the survey. Loosely speaking, in rural areas adjoining villages in the same division of a province are classified as being in the same area. In large cities, wealthy, middle income, and poor neighborhoods are treated as separate areas. In the case of other urban areas (i.e., towns with populations between 5,000 and 500,000), adjoining towns in the same division of a province are aggregated into 'areas'.

4.1.2. Controlling for prices

A second data-related difficulty we face is that we do not have direct independent data on fuel prices. Moreover, though we have data on *aggregate* energy expenditures, for most households, at the level of *disaggregated* fuels, we only observe *quantities* used, not expenditure levels. And hence, we are not able to use unit values as imperfect proxies for prices.[20]

We cannot obviously ignore the fact that we do not observe prices. Omitting prices from the Engel curves would, if prices and incomes are correlated, introduce biases in the estimated income coefficients. In this context, the sign of any likely correlation is not immediately clear. On the one hand, a positive correlation between price and income might arise from unobserved quality variation. And if higher income households use higher quality fuels that cost more, they may cut down on the quantity consumed, biasing downwards the estimated effect of income. On the other hand, a negative correlation might arise because "access costs" are lower in

---

[19] Households can in principle choose where to live. Thus 'lack of access' in some sense reflects household preferences. However, as the residential location choice of households is influenced by so many other factors, treating 'access' as exogenous here seems to us a reasonable assumption.

[20] Even if we were able to construct unit values, it is not at all clear that these would be suitable proxies for prices, especially in a context where there might well be unobserved variation in quality.

richer areas and this would result in an upward bias. In either case, with incomplete controls for prices, income would pick up some of the price effect, biasing its coefficient.

To at least partially address this issue, we use area and month fixed effects to control for unobserved variation in the *ratio* of modern and traditional fuels prices. Areas are defined as mentioned above when discussing the issue of access. The month fixed effects control for the month in which the household was interviewed, and this may partially control for seasonal price variations.

4.1.3. Fuel collection, household production, and externalities from deforestation

Finally, in the case of traditional fuels, and in particular, firewood, two other complications arise. The first is that in poor agrarian economies, traditional fuels are often either produced within the home (as a byproduct of other activities) or "collected", rather than purchased. As Bardhan et al (2002) point out in a very careful analysis of firewood collection behavior in Nepali villages, in such settings, consumption expenditures and fuel-use choices are likely to be *jointly* determined by household-level time-allocation decisions including decisions about the time devoted to "collection". The standard approach to estimation of Engel curves, which treats household consumption or income levels as being exogenous, is hence, unlikely to be appropriate.

Table 1 indicates that "collection" of firewood and dung is quite common in our sample, but is almost exclusively a *rural* phenomenon. So, because we do not have any plausible instruments for household incomes (or expenditure levels), we estimated fuel-use Engel curves for both the sample as a whole, and then separately for the sample of urban households. The results did not differ discernibly and hence, except in one instance, we only report the results for the entire sample.

A second complication stems from the fact that within rural communities there are clearly inter-household externalities associated with the deforestation that results from collection of firewood from common forest areas. Bardhan et al (2002) and Foster et al (2002) point out that these externalities are likely to give rise to community-specific effects in the household-level firewood-use functions. To the extent that the area effects we include are at a coarser level of resolution than the "communities" within which these externalities are relevant, our coefficient estimates may be biased. But here again, the problem is largely a rural one.

4.2. Estimation of fuel-use Engel curves

4.2.1. Specification

We estimate the following equations for each of the two aggregated fuel categories, traditional and modern:

$$\begin{aligned} q_h &= \alpha + \beta_1 y_h + \beta_2 d_h + \Sigma_m \gamma_m M_h^m + \Sigma_k \delta_k A_h^k + \epsilon_h \qquad (4.1) \\ q_h &= \alpha + \beta_1 y_h + \beta_2 (y_h)^2 + \beta_3 (y_h)^3 + \beta_4 (y_h)^4 + \beta_5 d_h + \Sigma_m \gamma_m M_h^m + \Sigma_k \delta_k A_h^k + \epsilon_h \end{aligned}$$

For each category, the dependent variable, q_h, is the total quantity (measured in megajoules) of fuels consumed in that category per month per capita by household

h. The key independent variable is y_h, monthly per-capita expenditures, while d_h represents household size and M_h^m and A_h^k are dummy variables for month and area (i.e., spatial) fixed effects.[21] The first specification provides a benchmark, while the second, in which we include a polynomial in y_h to allow for possible non-linearities, is our preferred specification.

Given the frequent censoring at zero of fuel use, even for the aggregated-fuels categories, our first estimation approach is a maximum-likelihood Tobit procedure. Since Tobit estimation with fixed effects requires at least one non-zero observation per fixed-effect unit, the inclusion of area effects automatically forces us to exclude all households from areas where none of the households in our sample use the fuel in question. As noted, this also is how we deal with the issue of access.

These Tobits function as a benchmark, especially for the traditional fuels. Tobit's single index function (single $X\beta$) assumes that the effects of explanatory factors on the probability of use are the same as upon the quantity used given any use. Both the theory and the descriptive statistics highlighted above suggest this is not likely to be appropriate for traditional fuels. Higher income lowers the probability of using traditional fuels, but conditional upon use it raises the quantity, as richer households consume more fuels services. For modern fuels this does not arise, since the expectation (supported by the descriptive statistics) is that higher income raises both the probability of use and the quantity used conditional upon use.

To get around the excessive restrictiveness of the Tobit specification, we implement, as our preferred strategy, the "generalized Tobit" model suggested by Cragg (1971). In this approach, the probability of a limit observation (the fuel choice decision) is estimated as a Probit, and a *separate* truncated regression model is estimated for the non-limit observations (the fuel quantity conditional on use decision). Previewing Table 4, for traditional fuels a likelihood ratio test comparing the Cragg generalized Tobit model with the basic Tobit model always soundly rejects the latter. The gains from this additional flexibility are not surprising, given the clear theoretical predictions and the descriptive statistics for traditional fuels.

4.2.2. Estimates

Table 4 presents our estimated fuels-usage Engel curves, for traditional and modern fuel aggregates. Recall, the dependent variables are quantity used per capita, in a unit of measure common to the two fuels categories, total megajoules (denoted "BTUs" in the table). The top half of the table presents regressions for traditional-fuel BTUs, while the bottom half concerns modern fuels. All runs include month and spatial fixed effects. Both the linear benchmark and the more general polynomial specifications are presented. Within each specification (linear to the left, polynomial to the right), three runs are shown. The first is the Tobit, which estimates a single

---

[21] We use per-capita expenditure rather than income because incomes can be quite variable in this setting. The literature on intertemporal consumption behavior suggests that household consumption decisions (of which fuel choices are a subset) are more likely to reflect long-run average income. If households smooth consumption in the face of income fluctuations, expenditures will be a better proxy for average income than actual income in any given period. We have also estimated these equations using per-capita income as the key explanatory variable. The results were not qualitatively different. Both the household income and the household expenditures measures we use below were created as part of the PIHS. Its construction is discussed in the *PIHS Basic Information* document.

index function for both the probability of use and the quantity conditional on use. The others are Cragg's probit and truncated regressions, the combination of which permits the use and quantity decisions to react differently to income.

The major findings are: i) evidence of a clear transition from traditional to modern fuels as per-capita household expenditure rises; and ii) evidence that, consistent with what the basic theory would have led us to expect, the more flexible estimation approach of Cragg (1971) reveals more clearly the nature of the transition.

The Cragg results indicate that for modern fuels, both the probability of use and the quantity used separately rise with expenditure. Not surprisingly, therefore, the standard Tobit estimate (blending use with quantity) also indicates significant positive effects of per-capita household expenditure on modern fuels.[22]

But does higher income lead households to drop traditional fuels? Here the Tobit evidence is ambiguous, a significant negative effect in the linear specification but a positive one in the fourth-order polynomial. This ambiguity is not surprising, since the single set of Tobit coefficients must blend the falling likelihood of using traditional fuels with the rising quantity used conditional on non-zero use. This difficulty is made clear by the Cragg generalized Tobit results. The Probit regressions for traditional fuel use yield consistently significant negative expenditure effects. On the other hand, the truncated regressions indicate that conditional on the use of traditional fuels, the quantity used rises with per-capita expenditure. Likelihood ratio tests confirm that for both the linear and the quartic specifications, the Cragg generalized Tobit model provides a better fit than the standard Tobit model.

Turning to the other controls, for all the estimation runs the area and month effects are (highly) jointly significant, indicating systematic spatial and seasonal variation in access and fuel prices. More interestingly, household size appears to be, statistically, a very important influence on fuel-choice and fuel-use decisions. Recall that the dependent variable is *per capita* use of fuels. Thus the fact that larger households will, in general, use more fuel does not (in the absence of within-household economies of scale, which we discuss below) imply a positive prior for the household-size variable.

Our basic results regarding the influence on fuels choices of household size are as follows: controlling for per-capita household expenditure, larger households are less likely to use traditional fuels and also use lower quantities per-capita when they do use it. On the other hand, larger households are more likely to use modern fuels but, as with traditional fuels, use lower quantities per-capita.

4.2.3. Interpreting the evidence: household size

How should we interpret these effects? Within the context of our model above (and natural extensions), three explanations come to mind. First, it is possible and perhaps even likely that there are economies of scale in the generation of cooking and other services from fuels. It clearly does not require five times as much fuel (and

---

[22] For the quartic specification, this is true for the income range we observe in the data. We calculated the predicted expected value of fuel use at the average estimated area and month effect, and for various household sizes. We held these constant as we varied expenditure to trace out an Engel curve. Thus our predictions do not include a forecast of changes in access or more generally, any changes in unobserved price components as per-capita expenditure (and income) rises (which would be reflected in changes in the area effects). Nor did we allow for systematic changes in household size with increases in per-capita expenditure levels.

energy) to cook for a household of five than it does to cook for a single individual.[23] Such economies of scale, combined with the fact that the number of equivalent adults is unlikely to increase one-for-one with the number of household members including children (our measure of household size), would explain why, controlling for per-capita household expenditure and conditional on use, per-capita quantities of fuel use decline with household size for both modern and traditional fuels.

Second, larger households may also realize scale economies in other types of consumption activities. That would imply a positive "income" effect of increases in household size on quantities of fuel use, though also a negative "substitution" effect relative to other types of consumption. Were the substitution effect to dominate, again the quantities per-capita of both modern and traditional fuels would be lower for larger households (an effect that can not be distinguished from the story just above). However, were the income effect to dominate—and whether it does will, in general, depend on the magnitude of the scale economies realized in the generation of cooking services relative to those in other consumption activities, and on the relevant income and price elasticities of demand—per-capita quantities of both types of fuel use would rise with household size.[24] The negative coefficients on household size in the fuel-quantity regressions suggest that such an income effect does not dominate both the substitution effect and any economies of scale in fuel-service provision.

Our third explanation extends our model in Section 2 above by noting that indoor air quality is a *local (i.e., intra-household) public good.* Therefore, the larger the household, the greater the benefits of improving indoor air quality by switching to cleaner modern fuels.[25] Unlike the two stories just above, this suggests a clear prior for the effects of household size in the Probit, probability-of-use regressions. Controlling for per-capita expenditure, larger households should be less likely to use traditional and more likely to use modern fuels, as we find in Table 4.[26]

This third, environmental story also has implications for the per-capita quantity regressions. Concern about indoor air quality may dampen quantities of use of both traditional and modern fuels, reinforcing the negative effects of any economies of scale in the generation of cooking services, or a dominant substitution effect of economies of scale in other consumption activities. However, we would expect that for this third story, the dampening of fuel use might be stronger for traditional fuels, because each additional unit of traditional fuel leads to a larger deterioration in indoor air quality. That would explain the greater magnitude and significance of the negative household size effect in the traditional fuels regressions in Table 4.

---

[23] While our regressions focus on cooking, more generally this sort of dissipation effect is relevant for fuels services. For instance, a fire generates heat and light, of which a significant share goes directly to the empty spaces of the household, not enjoyed by anybody. The more people around to absorb those benefits, the lower the share of dissipated services. Thus, larger households do not need as much fuel per capita for a given level of services per capita.

[24] Deaton and Paxson (1998), which looks at the relationship between per-capita food expenditures and household size, provides a detailed discussion of how economies of scale in consumption might interact with household size and composition.

[25] While for space reasons we do not do so here, it is simple enough to add this feature to our model in Section 2, arriving at this comparative static for the number of people in the household.

[26] To the extent that a switch to modern fuels increases the overall fuel bill of the household, such behavior would provide one potential explanation for the paradoxical finding highlighted by Deaton and Paxson (1998), namely that , controlling for per-capita household expenditure, per-capita food expenditure levels decline noticeably with household size.

4.2.4. Interpreting the evidence: household composition

The discussion above suggests that the estimated fuel-use Engel curves are consistent with the simple model we presented earlier in which we assumed that households value indoor air quality and that this concern influences fuel choice and use decisions. But they might also be consistent with some other explanation, in particular, one in which latent preferences for indoor air quality are *not* part of the story.

The leading candidate for such an explanation centers on possible fixed costs in the use of modern fuels, coupled with lower per-service-unit costs of modern fuels. In an environment characterized by credit market imperfections, the presence of such fixed costs could explain why the likelihood of modern fuel use increases with the level of per-capita household expenditure. The results regarding the influence of household size on fuel-choice could similarly be reconciled with a pure fixed costs story in that the effective per-capita price (inclusive of fixed costs) of modern fuels, is, by definition, lower in larger households, implying that controlling for per-capita expenditure, larger households should be more likely to use modern fuels.

Empirically speaking, fixed costs are likely to be relevant. To use LPG cylinders for cooking, households need to buy a gas burner, whose price can be a significant portion of average household income. They may also have to pay upfront deposits for cylinders. Likewise, to use kerosene households need to replace traditional mud ovens with kerosene stoves. Thus this explanation bears consideration.

However, two observations suggest that fixed costs *alone* cannot explain the behavior we observe. First, the empirical validity of the other essential ingredient of a pure fixed-costs explanation—that the variable costs (per unit of cooking services) of modern fuels be lower than those of traditional fuels—is questionable. We do not directly observe these costs. But in this setting, where many traditional fuels are collected and the opportunity costs of the time collecting are likely to be low in terms of foregone income because of involuntary unemployment and the use of child and female labor, the per-service-unit cost of modern fuels may well be higher for many households. Second, if fixed costs were the sole driver of the observed fuel-choice transition, we should not observe households using both traditional and modern fuels, but 12% of our sample do so (and some use multiple modern fuels).

We can also provide some direct evidence that concerns about indoor air quality do play a role in household fuel-choice decisions. For instance, Smith (1987) cites studies in Guatemala, Nepal, and India, which report that in post-adoption surveys households indicated that smoke exposure reductions were an important element in their decision to adopt cleaner-burning stoves. In our data, of the households who responded, 69% reported being irritated by smoke from cooking activities. If nothing else, these survey findings indicate that households are aware of the implications of their fuel choice decisions for indoor air quality.

Table 5 provides more formal evidence that the observed fuel-choice and fuel-usage behavior of households is at least partially driven by concerns about indoor air quality. We report there the results of probit regressions of traditional and modern fuel use that are similar to the fifth column of Table 4, except that now the household size variable is disaggregated into separate counts of adult males, adults females, boys and girls (under age 15) in the household. The top panel reports estimates generated using the full sample, while the bottom panel reports estimates from the sub-sample of urban households as a further robustness check.

The results indicate that more than household size, household composition matters for the choice of fuels used by the household. In particular, controlling for per-capita expenditure, the greater the number of women in the household, the less likely it is that the household uses traditional fuel and the more likely it is to use modern fuels.[27] The only other significant effect (at the 5% or 10% level) is the impact of an increase in the number of girls, which reduces the likelihood of traditional fuel use. The presence of more men in the household also makes it less likely that the household uses traditional fuels, but the effect is not significant.

It is difficult to reconcile these composition effects with a straightforward pure fixed-costs-based explanation. The presence of fixed costs suggests a direct role in influencing fuel choices only for total household size. But if households care about indoor air quality, household composition effects may arise, should the valuation of this intra-household public good vary across household members. In this setting, women do most if not all of the cooking, and as a result are more directly exposed to the indoor air pollution that results from fuels use.[28] Therefore, it should not be surprising that the greater the number of women in the household, the greater the value placed on improving indoor air quality, and the more likely it is that the household will switch from traditional fuels to modern fuels.

4.3. Predicting the income-indoor air quality relationship

What do the observed fuels choice imply about the relationship between per-capita expenditure levels and indoor air quality? We do not observe indoor air quality at the household level. Thus, for this analysis we resort to a hybrid strategy of estimation with partially simulated data. We use the data on the quantities (q_h) of modern and traditional fuels used by the households to construct alternative versions of an index that reflects the level of indoor air pollution for the household.

Let q_{th} and q_{mh} denote, respectively, the quantities of traditional and modern fuels used by household h. We assume that the function linking these levels of fuel use to the indoor air pollution experienced by the household, b_h, takes the form:[29]

$$b_h = (\rho q_{th})^{\theta} + (q_{mh})^{\theta} \tag{4.2}$$

where ρ is a parameter indicating the ratio of emissions (of pollutants) from traditional fuels to those from modern fuels, and θ is a parameter indicating the degree and direction of non-linearity with which emissions accumulate and translate into pollution within the household.[30] The higher the value of ρ, the more polluting

---

[27] Note that neither result necessarily implies the other because of the presence of households who use both types of fuels.

[28] Moreover, a number of studies (e.g., Thomas (1990), Behrman (1997)) suggest that mothers care more about the health of their children than do fathers. Because the health of children, and especially daughters who may assist the mother in cooking is adversely effected by indoor air pollution, that would—assuming a larger number of women implies an increase in the influence of women in household decisions—provide another explanation for our finding that the presence of women and girls decreases the likelihood of traditional fuel use.

[29] Note that when $\theta = 1$, this reduces to the simple specification, (2.3), we assumed in Section 2.

[30] We also considered a related functional form in which we assumed that any non-linearity in the function translating emissions into pollution applied to the *sum* of the emissions from the two different sources. Because the results we obtained were not substantively different, we do not report them here.

traditional fuels are assumed to be, relative to modern fuels. We consider values of ρ ranging from 5 to 200. A value of $\theta > 1$ (< 1) implies that the emissions function is convex (concave), i.e., that the marginal increase in pollution associated with each additional unit of fuel use rises (falls) with the level of fuel use. Whether emissions functions are concave or convex is likely to vary by setting, and is largely an empirical matter. From a theoretical perspective, convexity turns out to be a sufficient condition for non-monotonic environmental Engel curves (see Pfaff, Chaudhuri and Nye, 2002b) and that is what we assume in the theoretical voting example in the next section. For the empirical results of this section, however, it seems important to consider a full range of values of θ, so below we use from 0.25 to 1.5.

For each of the combinations of ρ and θ that we consider, we use the implied values of indoor-air pollution to estimate the following equation:

$$b_h = \alpha + \beta_1 y_h + \beta_2 (y_h)^2 + \beta_3 (y_h)^3 + \beta_4 (y_h)^4 + \beta_5 d_h + \Sigma_m \gamma_m M_h^m + \Sigma_k \delta_k A_h^k + \epsilon_h \quad (4.3)$$

where, as in (4.1) above, y_h is the per-capita level of household expenditure, d_h is household size, and M_h^m and A_h^k are sets of month and area effects. We use the estimates to predict the relationship between income and indoor air pollution.

Figures 3-6 plot these predicted relationships for different combinations of ρ and θ.[31] Figure 3 displays the predicted relationship (and associated standard error band) for the combination $\rho = 100$, $\theta = 0.5$. Indoor air pollution rises initially with increases in household expenditure but quickly levels off, and is more or less constant for a wide intermediate range of per-capita expenditure levels. Beyond a point, however, further increases in household expenditure levels are associated with a decline in the level of pollution. For this particular combination of parameters, therefore, the predicted relationship at the household level is an inverted-U.

Figures 4, 5 and 6 indicate that, with exceptions, a similar inverted-U relationship emerges to a lesser or greater extent for the combinations of parameters that we explored, in the range of household expenditures we observe in our sample. Figures 4 plots the predicted relationships under alternative assumptions regarding the ratio of emissions from traditional fuels to those from modern fuels, for two separate values of θ, 0.5 and 1. Holding constant the degree of non-linearity in the function linking emissions to indoor air pollution, the smaller the assumed difference in the emissions generated by the two types of fuels, the higher the household expenditure level at which indoor air pollution starts to fall, and the greater the initial increase in indoor air pollution before the subsequent decline.

Figures 5 and 6 plot the predicted relationship under different assumptions regarding the degree and direction of non-linearity in the generation of pollution from emissions, holding fixed the ratio of emissions from traditional fuels to those from modern fuels.[32] In Figure 5, this ratio is assumed to be 100, while in Figure 6, it

---

[31] In generating the predicted relationships, we fixed the household size at its mean for our sample, 7, and set the month and area effects to their estimated average values. Note also that the predicted levels of indoor air pollution at different levels of per-capita household expenditure have been normalized relative to the level at the lowest level of household expenditure.

[32] In addition to the 4 values of θ for which the results are displayed in Figure 5, we considered values of θ ranging up to 1.5. The shape of the predicted relationship in all these cases mirrors that for the case where θ is assumed to be 1.1, but with much larger initial increases and subsequent declines. To keep the scale of the vertical axis comparable to those of the other figures, we do not display the results for these higher values of θ.

is assumed to be only 5. These figures indicate that greater concavity (i.e., lower values of the non-linearity parameter, θ) dampens the initial increase in pollution levels and widens the range of household expenditures for which pollution levels remain more or less constant or even increase slightly. In fact, in Figure 6, the two cases which assume the greatest concavity ($\theta = 0.25$ and $\theta = 0.10$), are the two we alluded to earlier, for which we do not observe an inverted-U.

To sum up, our quasi-simulations indicate that for a wide range of plausible parameter values, the fuel-choice and fuel-use decisions of the households in our sample imply an inverted-U relationship between levels of indoor air pollution and per-capita household expenditures. This finding, while interesting in that it mirrors some of the empirical findings at the aggregate level, is however, itself less important than the evidence we present in Tables 4 and 5 which indicates that households do care about indoor air quality and can be expected to take actions to improve it.

5. Aggregating household preferences through voting

Household fuel choice decisions provided a natural arena for our *empirical* analysis because the environmental impacts of these decisions are largely internalized, offering some hope that we might uncover latent preferences towards the environment. From a theoretical perspective, however, the particular characterization of household preferences we adopted and provide empirical support for is quite general and suggests the building blocks for a theory of aggregate EKCs *even when environmental impacts are purely external and households are heterogeneous.* We demonstrate this in this section by means of a simple illustrative model.

Consider an economy with a large number of households with varying income levels, each of whom faces the (reformulated) decision problem we presented at the end of Section 2, i.e., to choose s and x to maximize $U(s, a)$, where s is the household's consumption of a composite good, x is the level of abatement expenditures it chooses to incur, and a is the level of environmental quality the household enjoys. We assume, as before, that $U(.,.)$ is an increasing and concave function of s and a, and that $\lim_{s \to 0} U_s(s, a) = +\infty$.

However, in contrast to Section 2 where we assumed that a depended only on the household's own consumption level and abatement expenditures, here we allow for purely external environmental effects by assuming that a varies with the *aggregate* level of consumption and abatement expenditures. Moreover, in this more general context we adopt a more general representation of the abatement technology and abatement expenditures than we did in Section 2. In particular, we assume that environmental quality is given by:

$$a(X, S) = \overline{a} + \gamma X - \delta S \qquad \gamma, \delta > 0 \tag{5.1}$$

where $\overline{a}$ is, as before, the initial endowment of environmental quality, S denotes overall consumption of services in the economy, and X denotes aggregate abatement expenditures.

If the number of households is large enough so that each individual household ignores its effects on environmental quality—i.e., treats X and S as fixed in choosing s and x—it is clear that no single household will choose to independently incur any abatement expenditures, instead choosing to devote its entire income to consumption. And in that case, in the absence of any collective choice mechanism, as incomes grow and consumption levels increase, environmental quality will continually and monotonically decline.

However, the literature on local public goods shows that voting mechanisms can coordinate individual decisions. We therefore consider the simplest possible voting scheme, a majority voting procedure on a proportional income tax rate, where it is understood that the proceeds of the tax will be used to finance public abatement expenditures.

Imagine, then, that each household calculates its preferred tax rate by solving:

$$\max_{0 \leq t \leq 1} U([1-t]y, a(tY, [1-t]Y)) \tag{5.2}$$

where y is the household's income, t is the proportional tax rate, and Y is aggregate income. Implicit in this formulation of the household's decision problem is our assumption that the household recognizes that the choice of a tax rate, t, not only determines its own disposable income (and hence consumption level), $[1-t]y$, but also affects the level of environmental quality because it determines the level of aggregate (tax-financed) abatement expenditures, $X = tY$, as well as the aggregate level of consumption, $S = [1-t]Y$. Thus, though *once a tax rate t is selected*, households, no matter how strong their preferences regarding the environment, ignore the external effects of their consumption decisions, *ex-ante* these preferences are reflected in the voting on the tax rate.

The assumptions we make above ensure that (5.2) yields an unique maximum, i.e., that preferences regarding the tax rate are single-peaked. Let $t^*(y;Y) \in [0,1]$ denote the preferred tax rate of a household with income y, given an aggregate income level of Y. It can easily be shown that there exists a threshold income level, $\widehat{y}(Y)$, such that $t^*(y;Y) = 0$ for all $y \leq \widehat{y}(Y)$. The intuition behind this result exactly parallels that in the single-agent-no-external-effects setting we described earlier. At low levels of income, and with an initial positive endowment of environment quality, households are unwilling to pay taxes to finance abatement—the marginal utility from additional consumption exceeds the marginal utility of the improvement in environmental quality that would be possible from a positive tax rate. Moreover, if we make the additional assumption that:

$$\frac{-sU_{ss}}{U_s} > 1 \qquad \text{for all } s \tag{5.3}$$

it can be shown that for $y > \widehat{y}(Y)$, the preferred tax rate rises with household income. This additional assumption essentially ensures that preferences for consumption are sufficiently elastic.[33] Then, as incomes and pollution increase, households are willing to devote a higher share of their incomes to abatement, by voting for higher proportional income tax rates.

With single-peaked preferences and monotonically increasing preferred tax rates, the median voter theorem applies. Thus, the tax rate that will emerge from the simple majority voting procedure will be the tax rate preferred by the median voter, in this case the household with the median level of household income. Letting y_m denote the median household income, the prevailing tax rate will therefore be given

---

[33] Stokey (1998) also makes this assumption in deriving sufficient conditions for an EKC in a representative agent framework, given a specific abatement technology.

by $t^*(y_m;Y)$ and the level of environmental quality by:

$$\begin{aligned} a(X,S) &= a(t^*(y_m;Y)Y\,,\,[1-t^*(y_m;Y)]Y) \\ &= \overline{a}+\gamma t^*(y_m;Y)Y-\delta[1-t^*(y_m;Y)]Y \end{aligned} \tag{5.4}$$

From (5.4) it is clear that the link between economic growth—increases in aggregate income, Y—and the environment is likely to be multifaceted, varying with the specifics of the growth process. In particular, the magnitude and direction of the shifts in y_m that accompany changes in Y matter crucially, generating a rich set of comparative static effects that in turn have obvious implications for empirical analyses using aggregate data.[34] For instance, we might ask, what happens to environmental quality when aggregate income increases but growth is concentrated in the top half of the income distribution—i.e., the median level of income remains unchanged. Alternately we might be interested in the effects on environmental quality of an increase in the median level of household income, holding fixed the overall level of income, an exercise which is suggestive of (but obviously need not guarantee) an increase in income equality. Or we might consider the impact of income growth when it is equiproportionately distributed, implying that the ratio of y_m to Y remains unchanged.

To convey the sorts of results this framework can yield, we briefly work through this last comparative static. Let aggregate income, Y, be a multiple, N of the median household income, y_m, in the economy at some initial point, and assume that all subsequent growth is equi-proportionately distributed, or in other words, that the ratio N remains unchanged over time.

Consider first the effect of income growth on the preferred tax rate. Recall that $\widehat{y}(Y)$ is the threshold level of income below which a household will prefer a tax rate of zero, given an aggregate income of Y. Given our assumptions about preferences, it can be easily shown that as Y falls to zero, $\widehat{y}$ goes to infinity, while as Y goes to infinity, $\widehat{y}$ monotonically declines to zero. Keeping in mind that $Y = Ny_m$, this implies that there exists a y_l such that for $y_m < y_l$, a tax rate of zero is preferred by the median household, i.e., $t^*(y_m;Ny_m) = 0$.

To characterize the impact of increases in y_m beyond the threshold y_l, note that with $Y = Ny_m$ the *total* effect of economic growth on the preferred tax rate of the median household is the sum of two *partial* effects:

$$\frac{dt^*(y_m;Ny_m)}{dy_m} = \frac{\partial t^*(y_m;Ny_m)}{\partial y_m} + N\frac{\partial t^*(y_m;Ny_m)}{\partial Y} \tag{5.5}$$

The first is the marginal effect of the increase in median income, holding constant aggregate income, while the second is the marginal effect of the increase in aggregate income, holding fixed median income. From the earlier discussion, we know that the first effect is non-negative. It is easy to show that if $\gamma t - \delta[1-t] < 0$, as will be the case at low values of t, the second effect is also positive, i.e., the preferred tax rate for a household with income y rises with aggregate income Y. The intuition behind this result is as follows. From (5.4) it can be seen that $\gamma t - \delta[1-t]$ represents

---

[34] There are a few existing papers that explore the multi-faceted nature of the link between economic growth and the environment. For instance, Cropper and Griffiths (1994) examines how population density factors in, while Frankel and Rose (2002) consider the influence of variation in the degree of openness to trade.

the direct impact of an increase in aggregate income Y on environmental quality, given an initial tax rate of t. When tax rates are low, even if their own incomes are unchanged, households recognize that aggregate incomes have increased and that, at the low existing tax rate, increased aggregate income has a negative net impact on the environment. Each household is therefore willing to at least partially offset the deterioration in environmental quality through an increase in the tax rate.

Therefore, from $t = 0$ at the threshold y_l, the preferred tax rate rises with increases in income until income reaches an upper threshold y_u defined implicitly by:[35]

$$t^*(y_u; Ny_u) = \frac{\delta}{\gamma + \delta} \tag{5.6}$$

Whether, beyond this point, the preferred tax rate continues to rise monotonically depends on the particular specification of preferences and values of γ and δ. However, the preferred tax rate cannot fall below the level in (5.6) as long as y_m remains above y_u.

Next we consider the impact of increasing median income on environmental quality. We know from above that when $y_m < y_l$ there are no tax-financed abatement expenditures. The only effect of increasing income, therefore, is to raise consumption levels and lower environmental quality. The situation is somewhat more complicated once y_m crosses this lower threshold. *Holding fixed aggregate income*, as y_m rises, the median household prefers a higher tax rate, which unambiguously improves environmental quality, by lowering aggregate consumption and increasing public abatement expenditures. But an increase in y_m also implies an increase in aggregate income, since $Y = Ny_m$. And this has two effects on environmental quality.

The first is the direct impact, $\gamma t - \delta[1 - t]$, of an increase in aggregate income. The second is the indirect impact realized through the change in the preferred tax rate (discussed above) induced by the increase in Y. When y_m is below the upper threshold y_u, the two effects are in opposing directions and hence, the net impact is ambiguous. However, it is easily verified that once y_m crosses the upper-threshold y_u, subsequent increases in y_m (and hence, Y) unambiguously improve environmental quality.

Thus, except for an intermediate region where the relationship between income growth and environmental quality is indeterminate, the overall relationship broadly mirrors the non-monotonic relationship emphasized by the empirical work on environmental Kuznets curves. That is, there exist two thresholds, y_l and y_u where $0 < y_l < y_u$, such that environmental quality falls with income when median income is below y_l and rises when median income is above y_u.

6. Conclusion

Indoor air quality is not only a major health issue in developing countries but also a window on households' valuation of environmental quality. Because it is foremost a private good, one whose degradation will be internalized to a significant extent, its value is reflected in the household choices which drive indoor air pollution.

---

[35] At this threshold, the net impact of a marginal increase in aggregate income is exactly equal to zero, i.e., $\gamma t - \delta[1 - t] = 0$.

This paper has characterized fuels choices, for insights into behavioral linkages between household preferences, levels of income, and environmental degradation. First we showed in a household-production model why even if indoor air quality is a normal good, the *possibility* of "household-level EKCs" (i.e., non-monotonic Engel curves for indoor air quality) arises naturally. Then for traditional and modern fuel aggregates, we provided evidence of a transition as household income rises from traditional fuels (dung, wood, other biomass) to modern (kerosene, LPG, natural gas). For a wide range of plausible assumptions regarding the emissions implied by fuel use, these observed household behaviors implied a U-shaped relationship between indoor air quality and household income. That is, increases in income initially lead to a deterioration in air quality, but later lead to increased air quality.

This evidence is consistent with the implications of our household model. The model extended the literature by predicting 'N-shaped' relationships between income and environmental degradation in fuels-choice and analogous abatement settings. Pollution will rise, later fall, but then rise again as income continues to rise, because further degradation is inevitable once a household is using only the cleanest fuel. Along these lines, one direction for research concerns the variety of marketed goods (such as fuels) that produce services but degrade the environment. In particular, as rising incomes lead more people to use only the least degrading existing good, and thus perhaps to be constrained by the lack of an even cleaner (less degrading) good, there will be demand for 'clean innovation' to generate such new goods.

Another avenue for further research concerns the need for additional explanation of environmental regulations, which appear to be an important factor in the level of air quality in developed countries. This paper has explicitly modeled one way that aggregate-level EKCs could arise from changed household voting for regulations as household incomes rise. Thus, our findings on household valuation of environmental quality have implications beyond the narrow class of private environmental amenities. However, how different political-economic settings and scenarios would affect the aggregation of these preferences is worth further exploration.

Appendix: details about the data

1. Fuel Quantities

Here we discuss information in the PIHS and how we formed our fuel-quantity estimates. Several questions permit calculations of quantities. Note that different questions were asked of males and females. In general, we believe that a number of assumptions must be made in order to arrive at quantity estimates.

a. Traditional fuels quantities

For wood, females are asked how many kilograms of wood are used per day, and how many days wood is used per month. We multiply these responses to get kilograms of wood used per month. However, this follows somewhat arbitrary correction of what appear to be miscodings of some observations, in grams instead of kilograms. The processes for dung and charcoal were the same, and also faced miscoding issues. The process for "biomass"(i.e. other biomass fuels) was also the same, after an additional step in which all of the other-biomass fuels are aggregated.

b. Modern fuels quantities

Here, quantity estimates can be constructed in at least two ways based on questions asked of males. If we can use males' answers about hours of usage to proportionally indicate quantity (applying a constant flow per hour), we could just use hours themselves.However, we are not sure whether to put much faith in those answers. They tend to have missing values for more than half of the households in which females report positive hours of usage and for a large fraction of households in which males themselves report some useage. Thus it is hard to know what subsample of answers is non-missing. Another difficulty in the natural-gas case is that the quantity responses seem to be generated from utility bills which seem to arrive infrequently, even quarterly, and to apply to a varying number of days for different households. The bills are also usually for multiple households.

Thus, our hours variables are meant to be estimates of hours of use of the fuel in question for cooking and other related purposes, and are created using questions asked of females. For each fuel (kerosene, LPG, and natural gas), females are asked how many days per month they use each of a number of appliances which make use the fuel in question, as well as how many hours a day they use the appliance when they use it. These answers can be multiplied to estimate hours per month of appliance useage (we add together only the appliances used for cooking and related services; for example, for kerosene we count stove hours, but not room heater or lamp hours). Where appropriate (in particular for stoves), females are also asked how many burners the appliance has. We use that response to estimate the number of burners used in an average use of that appliance, and then use that number to estimate "burner-hours" per month. If a constant flow per burner-hour can be meaningfully applied, then this number differs only in scale from the true quantity.

2. Income, consumption and other household characteristics

The PIHS has measures of household income and household expenditure, both resultign from quite detailed calculations. We use expenditures to better reflect long-run income (see Section 4), but have also made use of income, as a robustness check. We do not generate our own measures of these variables. As mentioned above, the PIHS also contains information that we can use on household size, number of adults, and number of children in the household, household head's age and education, number of rooms in the house, and whether the house has windows.

References

[1] Andreoni, J. and A. Levinson (2001). "The simple analytics of the environmental Kuznets curve," *Journal of Public Economics* 80:269-286.

[2] Asako, K. (1980). "Economic Growth and Environmental Pollution under the Max-Min Principle". *Journal of Environmental Economics and Management* 7:157-183.

[3] Bardan, P. et al. (2002), "The environmental impact of poverty: evidence from firewood collection in rural Nepal," *mimeo*, Department of Economics, University of California, Berkeley, September.

[4] Becker, G. (1965), "A theory of the allocation of time," *Economic Journal*, Vol. 75, pp.493-517.

[5] Becker, R.A. (1982). "Intergenerational Equity: The Capital-Environment Trade-Off". *Journal of Environmental Economics and Management* 9:165-185.

[6] Behrman, Jere R. (1997). "Intrahousehold distribution and the family," in *Handbook of population and family economics*, Vol.1A, eds. Mark R. Rosenzweig and Oded Stark, Amsterdam: North-Holland, pp.125-187.

[7] Chimeli, A.B.(2001). "Optimal Dynamics of Environmental Quality in Economies in Transition,". *Mimeo*, Department of Economics, University of Illinois at Urbana-Champaign.

[8] Copeland, B.R. and M.S. Taylor (1995). "Trade and Transboundary Pollution". *American Economic Review* 85(4):716-737.

[9] Cragg, J. (1971). "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods". *Econometrica* 39:829-844.

[10] Cropper, M. and C. Griffiths (1994), "The interaction of population growth and environmental quality," *American Economic Review*, 84(2), May, 250-254.

[11] D'Arge, R.C. and K.C. Kogiku (1973), "Economic growth and the environment," *Review of Economic Studies*, Vol. 40, pp. 61-77.

[12] Deacon, R. and P. Shapiro (1975). "Private preference for collective goods revealed through voting on referenda," *American Economic Review*, 65(5): 943-955, December.

[13] Deaton, A. and J. Muelbauer (1980). *Economics and Consumer Behavior,* Cambridge University Press, Cambridge.

[14] Deaton, A. and C. Paxson (1998). "Economies of scale, household size, and the demand for food," *Journal of Political Economy*, 106(5), October, pp. 897-930.

[15] Dockery, Doug and A. Pope. "Epidemiology of Acute Health Effects: Summary of Time-Series Studies." In Wilson, R and J. Spengler, eds.(1996). *Particles in Our Air: Concentrations and Health Effects.* Harvard University Press.

[16] Fischel, W. (1979), "Determinants of voting on environmental quality: a study of a New Hampshire pulp mill referendum," *Journal of Environmental Economics and Management*, 6(2): 107-118, June.

[17] Forster, B.A. (1973), "Optimal capital accumulation in a polluted environment," *Southern Economic Journal*, Vol. 39, pp.544-547.

[18] Foster, A., M. Rosenzweig, and J. Behrman (2002), "Population growth, income growth and deforestation: management of village common land in India," *mimeo*, Department of Economics, Brown University.

[19] Frankel, J. and A. Rose (2002), "Is trade good or bad for the environment? Sorting out the causality," *National Bureau of Economic Research Working Paper* #9201, September.

[20] Gorman, W.M. (reprinted 1980), "A possible procedure for analysing quality differentials in the egg market," *Review of Economic Studies*, Vol. 47.

[21] Grossman, G. and A. Krueger (1995), "Economic Growth and the Environment," *Quarterly Journal of Economics* 110(2):353-377.

[22] Grossman, G.M. (1995). "Pollution and Growth: what do we know?", in Goldin, Ian and L. Alan Winters, eds., *The Economics of Sustainable Development.* Cambridge University Press for the OECD and Centre for Economic Policy Research.

[23] Gruver, G.W. (1976), "Optimal investment in pollution control in a neoclassical growth context," *Journal of Environmental Economics and Management*, Vol. 3, pp.165-177.

[24] Harbaugh, W., A. Levinson and D. Wilson (2000)."Re-examining the Empirical Evidence for and Environmental Kuznets Curve". NBER Working Paper #7711, May.

[25] HESS (1993). *Household energy demand handbook*, Household Energy Strategy Study: Islamabad, Pakistan.

[26] Holtz-Eakin, D. and T. Selden (1995). "Stoking the fires? CO_2 emissions and economic growth. *Journal of Public Economics* 57(1):85-101.

[27] Jaeger, William K. (1998). "Growth and Environmental Resources: A Theoretical Basis for the U-shaped Path". *mimeo* 10/14/98, Williams College, and presentation at an AERE/ASSA session, New York, January 1999.

[28] John, A. and R. Pecchenino (1994). "An Overlapping Generations Model of Growth and the Environment". *The Economic Journal* 104:1393-1410.

[29] Jones, L.E. and R.E. Manuelli (2000)."Endogenous Policy Choice: The Case of Pollution and Growth." *Review of Economic Dynamics.*

[30] Kahn, M.E.(1998), "A Household Level Environmental Kuznets Curve," *Economics Letters*

[31] Kahn, M.E. and J.G. Matsusaka (1997), "Demand for environmental goods: evidence from voting patterns on California initiatives," *Journal of Law and Economics*, XL, April, 137-173.

[32] Keeler, E., M. Spence and R. Zeckhauser (1972), "The optimal control of pollution," *Journal of Economic Theory*, Vol. 4, pp.19-34.

[33] Kuznets, S. (1955). "Economic growth and income inequality," *American Economic Review*, Vol. 65, pp. 1-28.

[34] Lancaster, K.J. (1966a), "A new approach to consumer theory," Journal of Political Economy, Vol. 74, pp. 132-157.

[35] _______(1966b), "Change and innovation in the technology of consumption," *American Economic Review*, Vol. 56, pp. 14-23.

[36] Lipsey, R.G. and G. Rosenbluth (1971), "A contribution to the new theory of demand: a rehabilitation of the Giffen good," *Canadian Journal of Economics*, Vol. 4, pp. 131-163.

[37] Millimet, D.L. and T. Stengos (1999). "A Semiparametric Approach to Modeling the Environmental Kuznets Curve Across U.S. States". *mimeo*, Southern Methodist University.

[38] Pfaff, A.S.P. (2000), "From Deforestation to Reforestation in New England, USA," in M. Palo and H. Vanhanen, eds. *Global Prospects of Deforestation and Forest Transition*, Kluwer Academic Publishers.

[39] Pfaff, A.S.P., S. Chaudhuri and H.L.M. Nye (2002a, accepted with minor revisions). "Why might one expect Environmental Kuznets Curves? examining the desirability and feasibility of substitution". *Environmental and Resource Economics.*

[40] Pfaff, A.S.P., S. Chaudhuri and H.L.M. Nye (2002b, *submitted*)."Endowments, Preferences, Abatement and Voting: microfoundations of Environmental Kuznets Curves". *mimeo*, Department of Economics, Columbia University.

[41] Plourde, C.G. (1972). "A Model of Waste Accumulation and Disposal". *Canadian Journal of Economics* 5(1):119-125.

[42] Pope, C.A. (1989). "Respiratory disease associated with community air pollution and a steel mill, Utah Valley." *American Journal of Public Health* 79:623-628.

[43] Saint-Paul, G. (1995). "Discussion", in Goldin, Ian and L. Alan Winters, eds., *The Economics of Sustainable Development.* Cambridge University Press for the OECD and Centre for Economic Policy Research. pp.47-50.

[44] Seldon and Song (1995), "Neoclassical growth, the J curve for abatement and the inverted U curve for pollution," *Journal of Environmental Economics and Management,* 29(2), pp. 162-168.

[45] Seldon and Song (1994), "Environmental quality and development: is there a U for air pollution emissions?," *Journal of Environmental Economics and Management,* 27(2):147-162.

[46] Shafik, N. (1994). "Economic development and environmental quality: an econometric analysis," *Oxford Economic Papers*, v.46.

[47] Smith, Kirk (1987). *Biofuels, Air Pollution, and Health: A Global Review.* Plenum Press, New York, 452p.

[48] Smith, Kirk (1988). "Air Pollution: Assessing Total Exposure in Developing Countries". *Environment* 30(10):16-35.

[49] Smith, Kirk (1993). "Fuel Combustion, Air Pollution Exposure, and Health: The Situation in Developing Countries". *Annual Reveiw of Energy and Environment* 18:529-566.

[50] Smith, Kirk *et al* (1994). "Air Pollution and the Energy Ladder in Asian Cities". *Energy* 19:587-600.

[51] Spengler, J. and Martin Cohen. "Emissions from Indoor Combustion Sources", in *Indoor Air and Human Health,* R. Gammage,, ed. S. Kaye's Lewis Publishers. 1985.

[52] Stephens, J.K. (1976). "A Relatively Optimistic Analysis of Growth and Pollution in a Neoclassical Framework". *Journal of Environmental Economics and Management* 3:85-96.

[53] Stokey, Nancy L. (1998). "Are There Limits to Growth?". *International Economic Review* 39(1):1-31.

[54] Tahvonen, O. and J. Kuuluvainen (1993). "Economic Growth, Pollution, and Renewable Resources". *Journal of Environmental Economics and Management* 24:101–118.

[55] Taskin, F. and O. Zaim (2000). "Searching for a Kuznets Curve in Environmental Efficiency using Kernel Estimation". *Economics Letters* 68:217-223.

[56] Thomas, Duncan (1990). "Intra-household resource allocation: an inferential approach," *Journal of Human Resources*, 24(4), Fall, pp.635-664.

[57] Wilson, R and J. Spengler, eds.(1996). *Particles in Our Air: Concentrations and Health Effects.* Harvard University Press

[58] World Bank (1992). *World Development Report 1992: Development and the Environment.* Oxford University Press for the World Bank, Oxford, 308p.

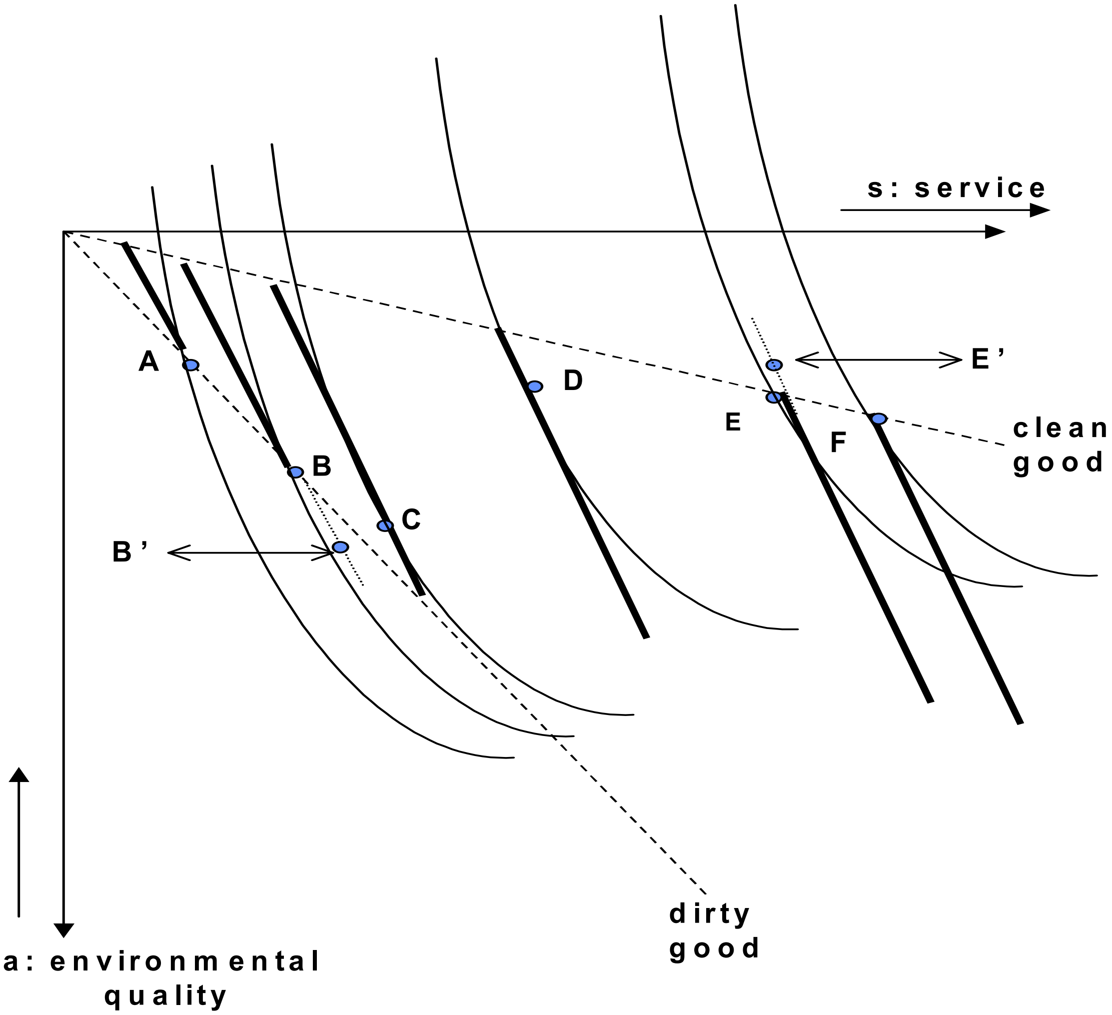

Figure 1

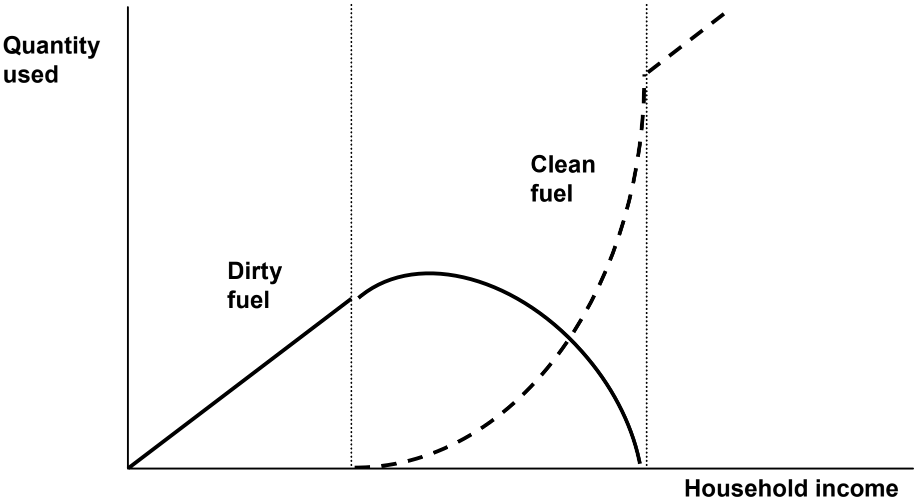

Figure 2

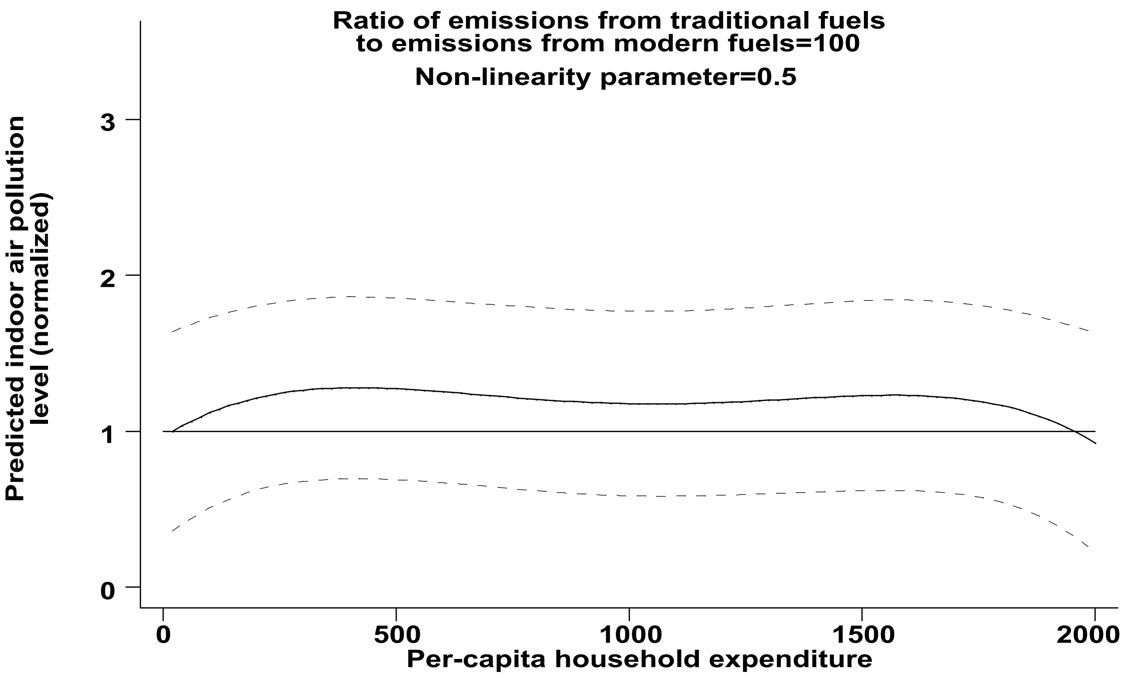

Figure 3

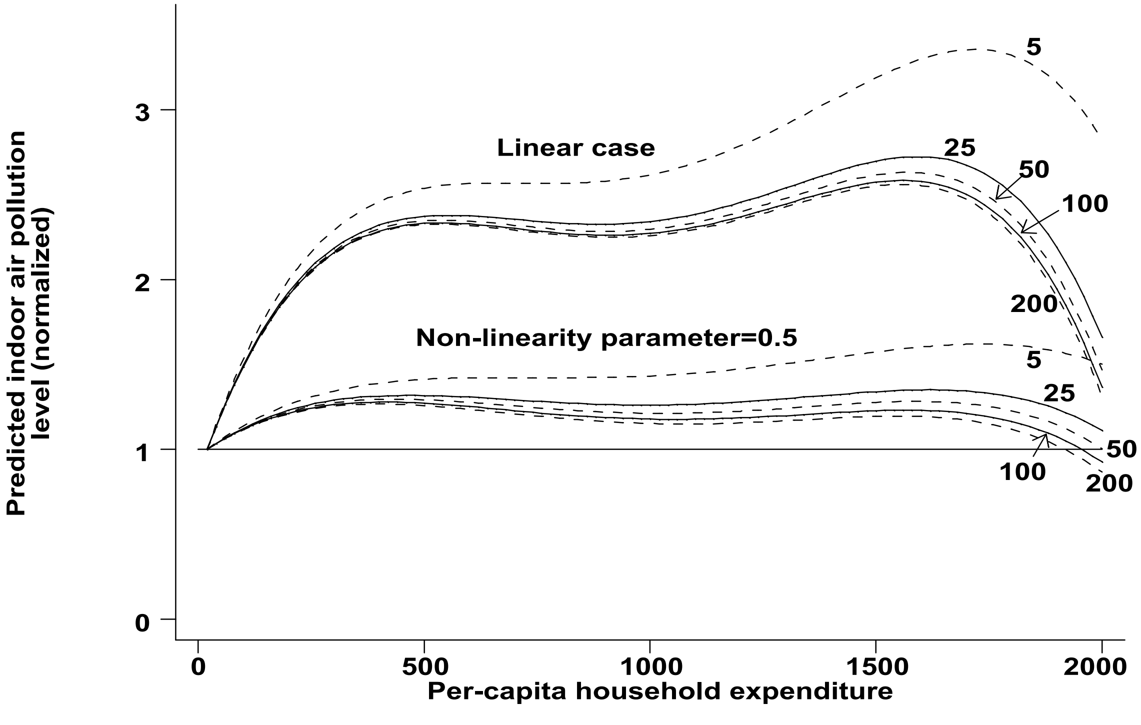

Figure 4
Predicted indoor air pollution levels under alternative assumptions regarding the ratio of emissions from traditional fuels to emissions from modern fuels: non-linearity parameter = 1, non-linearity parameter = 0.5

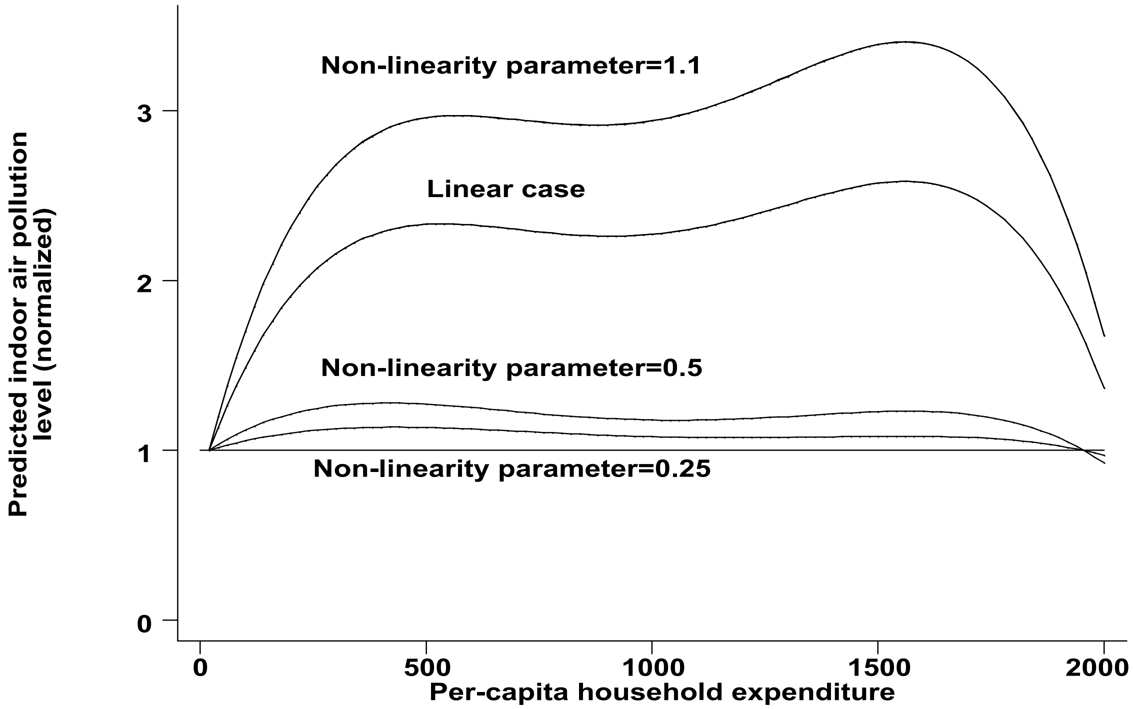

Figure 5
Predicted indoor air pollution levels under different assumptions
about the degree and direction of non-linearity in the emissions function:
emissions ratio = 100

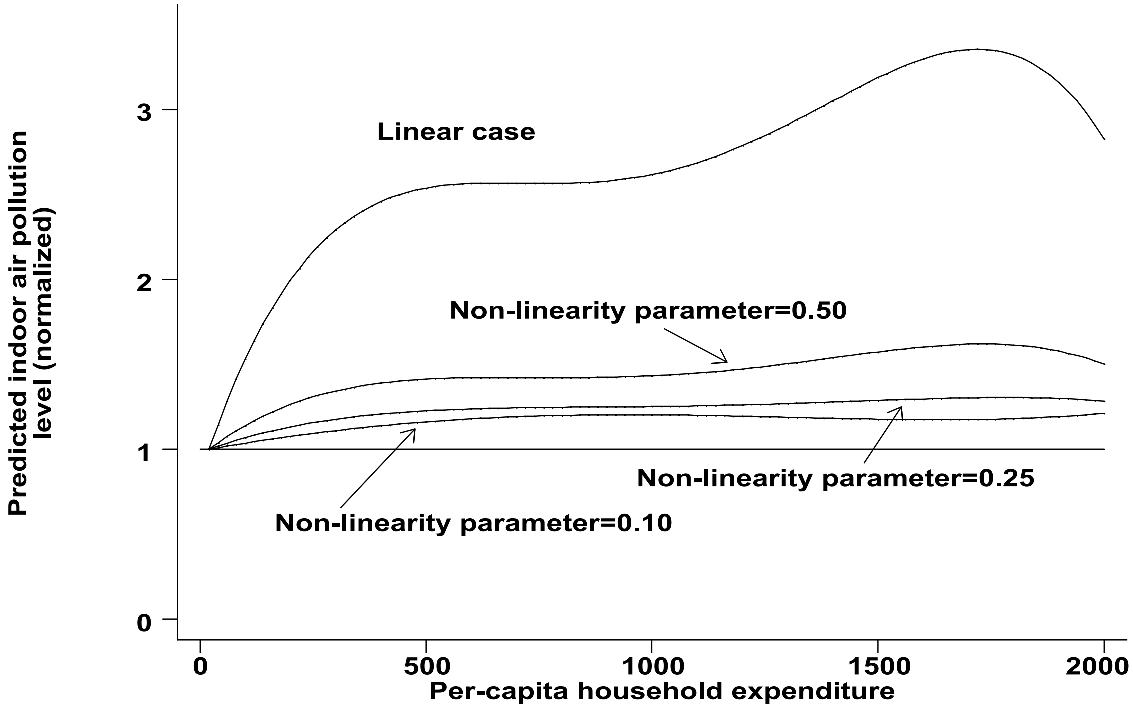

Figure 6
Predicted indoor air pollution levels under different assumptions
about the degree and direction of non-linearity in the emissions function:
emissions ratio = 5

Table 1
Summary statistics

		Sample				
Means		Full	Rural	Urban	Poor	Rich
Household income: monthly (Rs.)		4608	3295	5971	1434	7777
Per-capita income: monthly (Rs.)		728	447	1019	177	1139
Household expenditure: monthly (Rs.)		4957	4122	5821	3935	5978
Per-capita expenditure: monthly (Rs.)		770	582	964	486	876
Household food expenditure: monthly (Rs.)		1967	2064	1865	1886	2048
Household fuel expenditure: monthly (Rs.)		203	124	284	160	245
Household size		7.5	7.6	7.3	8.1	6.8
Number of adults		3.7	3.7	3.7	3.6	3.8
Number of children		3.8	4.0	3.7	4.6	3.1
Education of household head		6.2	6.3	6.2	6.2	6.3
Age of household head		45.9	45.7	46.1	45.5	46.3
Number of rooms		2.5	2.3	2.6	2.3	2.7
Proportion of households that use:						
Electricity	…at all	0.76	0.58	0.95	0.67	0.85
	…for cooking	0.02	0.01	0.02	0.01	0.03
Natural gas	…at all	0.18	0.01	0.36	0.07	0.28
	…for cooking	0.18	0.01	0.36	0.07	0.28
LPG cylinder	…at all	0.08	0.04	0.12	0.05	0.11
	…for cooking	0.08	0.04	0.12	0.05	0.11
Kerosene	…at all	0.73	0.90	0.56	0.83	0.64
	…for cooking	0.17	0.08	0.26	0.13	0.21
Wood	…at all	0.70	0.91	0.49	0.82	0.59
	…for cooking	0.64	0.83	0.45	0.75	0.54
Dung	…at all	0.45	0.67	0.22	0.58	0.32
	…for cooking	0.41	0.62	0.19	0.54	0.28
Charcoal	…at all	0.08	0.09	0.01	0.09	0.07
	…for cooking	0.01	0.01	0.01	0.01	0.01
Coal	…at all	0.01	0.01	0.00	0.01	0.01
	…for cooking	0.00	0.00	0.00	0.00	0.00
Biomass	…at all	0.23	0.36	0.10	0.30	0.16
	…for cooking	0.22	0.34	0.09	0.29	0.15
Proportion of households that:						
…collect wood		0.32	0.57	0.06	0.43	0.21
…collect dung		0.21	0.35	0.06	0.29	0.13
…report smoke irritation		0.69	0.93	0.46	0.82	0.56
…have no windows		0.48	0.61	0.35	0.60	0.36
…have a chimney		0.16	0.22	0.10	0.20	0.14
…have a servant		0.01	0.02	0.01	0.02	0.01
Number of households		4650	2366	2284	2323	2332

Notes: households are classified as poor (rich) if their per-capita expenditure was below (above) the median for the full sample.

Table 2
Fuel use (for cooking), by expenditure quartiles

Proportion of households	Full sample	Poorest quartile	26th to 50th percentile	51st to 75th percentile	Richest quartile
Wood	0.68	0.79	0.75	0.68	0.48
Dung	0.43	0.57	0.49	0.41	0.25
Other biomass	0.23	0.32	0.27	0.19	0.13
Kerosene	0.15	0.07	0.14	0.19	0.18
LPG cylinders	0.08	0.02	0.05	0.10	0.14
Natural gas	0.18	0.07	0.11	0.18	0.36
Usage levels: conditional means					
Wood: kilograms	29.8	23.3	28.1	31.6	40.5
Dung: kilograms	20.4	16.7	19.6	22.1	27.2
Other biomass: kilograms	22.7	18.5	21.2	22.4	36.6
Kerosene: hours	27.9	19.3	21.9	30.6	33.1
LPG cylinders: hours	16.1	8.9	13.5	14.2	19.1
Natural gas: hours	25.7	17.5	22.8	21.5	30.2

Notes: means of quantity used are conditional means, i.e., only including households with positive levels of use.

Table 3
Fuel use (for cooking), by aggregated fuel categories and expenditure deciles

Expenditure Decile	Proportion of households that use:			Usage levels (BTUs): conditional mean		Usage levels (BTUs): unconditional means	
	Traditional	Modern	Both	Traditional	Modern	Traditional	Modern
Poorest	0.95	0.09	0.04	502	1604	480	141
2nd	0.89	0.18	0.08	578	1469	517	271
3rd	0.88	0.20	0.08	588	1686	518	344
4th	0.85	0.28	0.13	613	2197	521	624
5th	0.79	0.33	0.12	718	2013	567	674
6th	0.77	0.41	0.18	679	1909	520	787
7th	0.71	0.44	0.15	720	2062	510	905
8th	0.71	0.50	0.21	794	2025	564	1013
9th	0.59	0.60	0.18	890	2551	521	1522
Richest	0.35	0.75	0.10	1027	5230	364	3914
Full sample	0.75	0.38	0.13	678	2688	508	1019

Notes: (1) we classify natural gas, LPG and kerosene as modern fuels and wood, dung, and biomass as traditional fuels. (2) usage levels are measured in BTUs and are per-capita per month. (3) conditional means are means conditional on positive levels of use. (4) unconditional means include households with zero use.

Table 4
Fuel use Engel curves: distinguishing the fuel-choice and fuel-use decisions

Traditional fuels

Dependent variable:	Linear specification[a]			Quartic specification[a]		
BTUs from traditional fuels (wood, dung, biomass)		Generalized Tobit: Cragg (1971)			Generalized Tobit: Cragg (1971)	
	Tobit[b]	Probit	Truncated	Tobit[b]	Probit	Truncated
Per-capita expenditure	-.057	-.001	.450	1.07	-.004	2.64
	(1.94)	(11.8)	(6.83)	(1.80)	(1.90)	(1.89)
Per-capita expenditure: squared				-.002	4.0E-06	-.003
				(1.48)	(0.90)	(1.20)
Per-capita expenditure: cubed				1.0E-06	-2.0E-09	2.0E-06
				(1.3)	(0.60)	(1.10)
Per-capita expenditure: quartic				-3.0E-10	3.0E-13	-6.0E-10
				(1.30)	(0.40)	(1.10)
Household size	-43.3	-.013	-137	-41.9	-.018	-120
	(15.9)	(1.51)	(16.6)	(15.30)	(2.00)	(14.7)
Constant	1243	3.05	1348	1015	3.80	804
	(17.9)	(9.15)	(9.63)	(8.74)	(7.61)	(3.03)
R-squared	0.34	0.47	0.25	0.34	0.47	0.26
Number of observations	3876	3876	2888	3876	3876	2888

Modern fuels

Dependent variable:	Linear specification[a]			Quartic specification[a]		
BTUs from modern fuels (natural gas, LPG, kerosene)		Generalized Tobit: Cragg (1971)			Generalized Tobit: Cragg (1971)	
	Tobit[b]	Probit	Truncated	Tobit[b]	Probit	Truncated
Per-capita expenditure	3.02	.001	5.38	20.0	.004	13.7
	(14.5)	(15.2)	(6.33)	(3.75)	(2.24)	(2.12)
Per-capita expenditure: squared				-.023	-2.0E-06	.024
				(2.38)	(0.50)	(2.10)
Per-capita expenditure: cubed				1.0E-05	-6.0E-10	2.0E-04
				(1.80)	(0.20)	(2.10)
Per-capita expenditure: quartic				-2.0E-09	4.0E-13	-4.0E-09
				(1.40)	(0.60)	(2.00)
Household size	58.5	.046	-75.5	70.9	.052	-82.9
	(2.81)	(6.72)	(0.75)	(3.37)	(7.56)	(3.47)
Constant	-7009	-2.53	-17250	-10751	-3.39	-1273
	(10.0)	(12.8)	(4.00)	(9.00)	(9.46)	(0.90)
R-squared	0.29	0.47	0.25	0.34	0.47	0.26
Number of observations	4331	4331	1562	4331	4331	1562

Notes: absolute values of t-statistics are reported in parentheses below the coefficient estimates. In the simple generalization of the Tobit model suggested by Cragg (1971) the probability of a limit observation (the fuel choice decision) is estimated as a Probit, and a separate truncated regression model is estimated for the non-limit observations (the fuel quantity conditional on use decision). (a) specification includes a full set of month and spatial fixed-effects. Month effects are jointly highly significant (p-value=0.0000) as are the spatial effects; (b) likelihood ratio tests comparing the generalized Tobit results to the simple Tobit results clearly reject the Tobit (p-value=0.0000).

Table 5

Household composition and fuel-choice: Probit regressions of fuel-choice

Full sample

Variable	**Traditional fuel use?**	**Modern fuel use?**
Per-capita household expenditure	-0.00184	0.001409
	(14.2)	(9.99)
Per-capita household expenditure squared	4.60E-07	-2.95E-07
	(10.1)	(6.18)
Per-capita household expenditure cubed	-3.43E-11	2.01E-11
	(8.27)	(4.81)
Per-capita household expenditure to the fourth	5.24E-16	-2.99E-16
	(7.785)	(4.45)
No. of adult males in the household	-0.03428	-0.00847
	(1.50)	(0.32)
No. of adult females in the household	-0.09513	0.078999
	(3.40)	(2.45)
No. of boys (below age 15) in the household	-0.01279	-0.02274
	(0.78)	(1.15)
No. of girls (below age 15) in the household	-0.05845	0.014425
	(3.72)	(9.99)
Month effects (p-value from chi-square test of joint significance)	(0.000)	(0.000)
Area effects (p-value from chi-square test of joint significance)	(0.000)	(0.000)
No. of observations	4562	4106

Sub-sample of urban households

Variable	**Traditional fuel use?**	**Modern fuel use?**
Per-capita household expenditure	-0.00185	0.001316
	(10.49)	(8.62)
Per-capita household expenditure squared	4.19E-07	-2.46E-07
	(6.63)	(4.76)
Per-capita household expenditure cubed	-2.96E-11	1.54E-11
	(5.03)	(3.44)
Per-capita household expenditure to the fourth	4.44E-16	-2.23E-16
	(4.64)	(3.10)
No. of adult males in the household	-0.04045	0.012029
	(1.37)	(0.42)
No. of adult females in the household	-0.13914	0.119538
	(3.71)	(3.37)
No. of boys (below age 15) in the household	-0.01271	-0.02198
	(0.59)	(1.03)
No. of girls (below age 15) in the household	-0.04218	0.007585
	(1.99)	(0.36)
Month effects (p-value from chi-square test of joint significance)	(0.000)	(0.000)
Area effects (p-value from chi-square test of joint significance)	(0.000)	(0.000)
No. of observations	2267	2267

Note: absolute value of t-statistics reported in parentheses.