Columbia University

Department of Economics
Discussion Paper Series

Economic Growth and the Environment: What Can We Learn From Household Data?

Shubham Chaudhuri
Alexander S.P. Pfaff

Discussion Paper #:0102-51

Department of Economics
Columbia University
New York, NY 10027

April 2002

Economic growth and the environment: what can we learn from household data?[*]

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

February 2002

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 the household choice framework developed in Pfaff, Chaudhuri and Nye (2002a), which suggests that even if environmental quality is a normal good, non-monotonic environmental Engel curves can arise. 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 analysis by Grossman and Krueger (1995), and others, brought to the fore the possibility of non-monotonic relationships between income and environmental quality.[1] Initial economic growth degrades environment, but continued growth could reverse initial adverse effects. Not surprisingly, these 'Environmental Kuznets Curves' (or EKCs) have generated considerable debate. Others have questioned the robustness of these initial, aggregate findings, arguing that the relationship between environment and economic growth is sensitive to the pollutants studied and the data used.[2] Many have also 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 important to establish the validity and generality of these aggregate empirical findings. It is also crucial to obtain a better understanding of the channels through which the claimed effect might arise, as well as the role that policies can play in amplifying or in dampening this effect. This paper contributes to the latter, briefly illustrating a household-choice mechanism (relevant for private degradation[4] and for voting on the environmental regulations many consider a key to EKCs), and most importantly providing estimates of EKC-relevant income effects based on household-level choices that determine environmental quality.

To see the value of household-level analysis, consider that the models often used to demonstrate EKC results are planner models. In contrast to ours, these suggest neither new empirical paths nor explicit stories for how environmental regulations in fact arise. Certainly the neoclassical growth tradition has provided one way of thinking about how EKCs might arise.[5] Gruver 1976, for example, incorporates the choice of investment between productive and pollution-control capital. Under certain parameter configurations, the optimal growth path can feature accumulation of productive capital, in the initial stages of growth, until beyond a target stock of productive capital savings are shifted to pollution control.[6] Such models can be linked only to aggregate-level empirics of the sort now prevalent in the literature, however. Further, the dynamic representative agent framework lacks a realistic political-economic mechanism through which degradation might in reality be reversed, and will not easily explain regulatory choice given heterogeneous voters.

---

[1] See World Bank 1992, Selden & Song 1994, Shafik 1994 and Holtz-Eakin & Selden 1995.

[2] Special issues of Environment and Development Economics November 1997 and Ecological Economics May 1998, and Millimet & Stengos 1999, Harbaugh et al. 2000, Taskin & 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] Many forms of degradation of the environment feature components that are private, and we observe significant private provision of environmental abatement in the absence of regulations.

[5] 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).

[6] See, for instance, 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.

The point of departure for this paper, then, is our claim that whatever the EKC mechanism at work, it ultimately has to be linked back to a theoretically coherent and empirically grounded account of: i) how households' marginal valuations of the environment evolve with increases in household income; and ii) how those views get aggregated up and manifested through the political policy-making process and/or through markets. On these two questions, there has been surprisingly little research.

The difficulty with addressing the first question effectively empirically is that when households' environmental impacts effects are external, observed household choices will 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 indoor air quality in developing countries, and fuels choices. These choices can be expected to reflect pollution impacts because the indoor air quality degraded by fuels usage is a private good. While smoke does exit the household, with external impact, its effects on indoor air quality can be an enormous factor in household members' morbidity and mortality.[7] Consider that perhaps eighty percent of world exposure to particulates occurs indoors in developing countries (Smith 1993, p.551). Such exposure has severe health consequences.[8] 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.

Specifically, using household-level data from Pakistan, we estimate both disaggregated (natural gas, LPG, kerosene, wood, dung, and other biomass) and aggregated (modern and traditional) fuel-use Engel curves. We then translate these into effects of income on indoor air quality. A benefit of explicitly modeling the underlying mechanism is apparent here, concerning appropriate estimation. The natural first thought is that Tobit regression will adequately address the use/not and fuel-quantity decisions adequately. Our household-production model[9], however, makes

---

[7] Indoor air pollution is often the dominant source of exposure because of the close proximity and extended exposure of individuals to the sources. 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."

[8] 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.

[9] 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

clear that the decisions about whether to use a fuel, and how much to use, may react quite differently to income. For 'dirtier' fuels, higher income lowers the likelihood of using the fuel but, conditional on it being used, raises the quantity. This suggests the need for a "generalized Tobit" approach to the estimation.

We find that as incomes rise, households transition from 'traditional' fuels to 'modern', *cleaner* fuels, in keeping with previous work on "energy ladders". This finding accords well with intuition that increases in household income may permit improvements in indoor air quality. However, as per our model, such improvement may not be mononotic. In fact, in simulations based on our estimates, under plausible assumptions on the ratio of traditional-fuel to modern-fuel emissions we find an inverted-U relationship between emissions and income. In other words, increases in income appear 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.

Finally, given this evidence that household preferences may be consistent with EKC-like Engel curves for indoor air quality, we provide an illustration of how this household-level approach can address the second question above, specifically how households' views get aggregated up through the political policy-making process. When extended to a multi-agent setting with heterogeneous agents and external effects, our characterization of households can yield a non-monotonic aggregate relationship between income and environmental quality that mirrors the household's.

The remainder of the paper is as follows. Section 2 briefly summarizes the household-production model from Pfaff, Chaudhuri and Nye 2002a that seems appropriate for this setting. We present the basic characteristics approach used to model the demand for fuels, and describe how a non-monotonic Engel curve for indoor air quality might arise naturally even when households value indoor air quality and preferences are such that air quality would be a normal good if it could be directly purchased. Section 3 describes the data and presents descriptive statistics. The data come from the Pakistan Integrated Household Survey (PIHS), one of a number of micro consumption datasets produced at least in part by the World Bank within their Living Standards Measurement Surveys series of datasets. The PIHS includes an energy module that permits study of fuels purchases 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-usage Engel curves and simulations. 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

Here we summarize Pfaff, Chaudhuri and Nye 2002a, which provides derivations of the conditions for EKCs. 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

---

Pfaff, Chaudhuri and Nye 2002a, see Section 2 below and that paper for extended discussions.

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. 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, 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) fuel d and a "clean" fuel c. 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:

$$s(\vec{q}) = q_d + q_c \tag{2.1}$$

where $\vec{q} = (q_d, q_c)$ is the vector of quantities used of the dirty and clean fuels respectively. Without losing any of the basic intuitions, we can also assume that the degradation of indoor air quality a is fully linear in the marketed fuels. Thus, we assume both that the total emissions level e is linear in the purchased fuels:

$$e(\vec{q}) = \alpha q_d + \beta q_c \tag{2.2}$$

where $\alpha > \beta > 0$ (i.e., fuel d is indeed dirtier), and that indoor air quality is linear in total emissions, where A is the initial air quality endowment and $A > 0$:

$$a(e) = A - e \tag{2.3}$$

The household chooses the marketed $\vec{q}$ to maximize (2.4) subject to (2.5):

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

$$p_d q_d + p_c q_c = y \tag{2.5}$$

where y is household income and p_d and p_c are, respectively, the per-unit (of cooking services) prices of the dirty and clean goods (we also assume that $p_d < p_c$).

To provide a first intuition for the possibility of an EKC based on the fuels usage decisions, ignore that the fuels demand functions may not be differentiable at all incomes because of binding non-negativity constraints on fuels usage. Then we can represent the slope of the household Engel curve for indoor air quality a as:

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

At first glance, this suggests that a must fall with y; all fuels q degrade a, and fuels demands are surely normal. However, the fuels demands are *derived* demands, given household preferences for s and a. Thus, they need not be normal. In fact, within a characteristics/household production framework, inferior marketed goods can be quite common (Deaton and Muellbauer 1980, Lipsey and Rosenbluth 1971). If dirty fuels are inferior after a certain income, e.g., while clean fuels are normal, it is possible that the Engel curve for indoor air quality will be U-shaped. More generally, the household-level relationship between income and indoor air quality could take on any number of shapes, including a monotic rise in quality. This indeterminacy is an attractive property given the contested aggregate evidence.

The illustrative example depicted in Figure 1 helps to provide some intuition. 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 observed fuels usage? Again, an illustrative example makes the basic point (worked out in detail in Pfaff, Chaudhuri and Nye 2002a). Figure 2 shows that at low incomes, cleaner fuels may not be used at all. The likelihood of using cleaner fuels rises with income, and the later transitions make clear that the quantity used of the cleaner fuel will also increase with income. Note that when only the dirtier fuel is used, its quantity also increases with income. Yet as income continues to rise, the likelihood of using the dirtier fuel at all will fall.

3. Data

3.1. Source & Overview

The data for this study is the Pakistan Integrated Household Survey (PIHS) 1991.[10] The PIHS is a national survey, although it oversamples urban households.[11] It in-

---

[10] 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.

[11] 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

cludes individual and household-level data covering not only energy consumption but also housing conditions, education, health, employment characteristics, self-employment activity, consumption, migration, fertility, credit and savings. Thus it permits us to attain our empirical goal of studying both the propensity to use traditional versus modern fuels at all and, conditional on useage, the quantity consumed and how it moves with household income (controlling for other relevant variables).

For each of the 4,800 households the PIHS indicates whether a 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.[12] In addition, several questions permit calculation of the quantities of these fuels used for cooking (see Appendix A for details). The PIHS also contains information on the following characteristics of a family: household income, household size, number of adults in the household, number of children in the household, household head's age and education, number of members of the household that have suffered respiratory problems. In addition, we constructed variables which may affect indoor air pollution exposure conditional on fuel useage, including: whether the home is used as a place of work, whether the family employs a servant to cook meals; where the cook cooks, the number of rooms, whether windows are present, and whether the home has a chimney.

3.2. Descriptive Statistics

Table 1 presents household and fuel-use statistics, and other relevant information. 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 income). 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.[13]

The top half of Table 2 presents statistics on use of fuels for cooking, by income quartile (expanding on the right-most columns of Table 1). The "modern" fuels (kerosene, LPG, natural gas) rise in use from poorer to richer quartiles, though variation exists: kerosene usage rises at earlier incomes, while LPG and natural gas rise more in the upper quartiles. The "traditional" fuels (dung, firewood, other biomass) fall in use from the poorer to the richer quartiles, though again variation exists: the use of dung drops earlier, while biomass and firewood use drop only later.

---

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.

[12] 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.

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

The bottom half of Table 2 presents information on the quantities used in cooking. The quantities follow much the same pattern as the use decisions for the modern fuels. For traditional fuels, in contrast, the quantities give a different picture than the binary use decisions, holding steady or rising as income rises.

This seems surprising, but note that quantities blend probability of use with quantity conditional upon use. If for whatever reason a household continues to use traditional fuels as it gets richer, it will use increasing amounts of traditional fuels, as its consumption of fuels services will rise with income. That can explains these descriptive statistics, and also suggests a "generalized Tobit" approach to estimation, in particular for the traditional fuels where the effect of income on the probability of use is of the opposite sign from that for quantity conditional on use.

Table 1 suggests one reason why some richer households use traditional fuels, a lack of access to modern fuels. Over 80% of rural households use wood for cooking; this must include some richer households. It makes sense that some rural areas do not have access to some fuels. This motivates an additional issue in the estimation, controlling for access to fuels in attempting to estimate the effect of income.

Table 3 shows that overall less than half of all households in our sample use only one fuel for cooking. Note that a non-negligible fraction of households even uses three fuels for cooking. The bottom half of the table creates modern and traditional groupings and reconsiders the use decisions in that light. These aggregated statistics present much the same message as in Table 2: the use for cooking of groupings involving only modern fuels rises from the poorest to the richest quartiles, the mix of traditional fuels and kerosene first rises in the middle quartiles and then drops, and the use of groupings involving only traditional fuels drops.

4. Econometric Issues, Estimation and Simulations

While our theory starts with a household's preferences for non-marketed goods and derives the demand for marketed fuels, our empirical strategy proceeds in the opposite direction. We begin by estimating Engel curves for the fuels and work back, through simulations, to the revealed preferences for air quality and cooking/heating services. Though fairly straightforward in principle, the practical implementation of this 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 below, next present our empirical specification and the results of our estimation, and finally provide simulations of the income - indoor air quality relationship.

4.1. Econometric Issues

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.[14] 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 this 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. Fuel collection and household production

A feature of a number of developing agrarian economies is that many commodities that a household consumes are produced by the household or instead are "collected". Table 1 reflects this—32% of the sample report collecting firewood and 21% report collecting dung. The proportions are higher in rural areas. Almost all other biomass is either collected or obtained as the byproduct of agricultural cultivation. Dung is also obtained as the byproduct of animal husbandry activities.

That some 'marketed' fuels are not always purchased raises conceptual issues on two fronts. First, without an observed market transaction the effective price of the fuel has to be *imputed*. There are several ways this might be done, none of which are entirely satisfactory. A common approach is to assign the average market price of the fuel in the community. With home-produced fuels, especially those obtained as the byproducts of other activities—e.g. crop byproducts from cultivation or dung from animal husbandry—this seems reasonable. However, for fuel collection for consumption, valuation at market prices potentially overstates the 'price' faced by the household. The opportunity cost of the time collecting is presumably lower than the market price—it would not otherwise make sense to be collecting.

An alternative approach is to price the collected fuel in terms of the opportunity costs of the time spent in collection. Determining the appropriate opportunity cost can, however, be difficult. A natural option might be to value the time spent collecting fuel by a household member at the average wage rate for that member's demographic category (i.e., males, females, children). But this seems inappropriate in a setting where underemployment is likely to be significant.

Again, we have no clean way of resolving this issue. The measures of income that we use include the imputed value of the time spent collecting firewood, and may

---

[14] 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.

therefore overestimate the incomes of poorer households. Given that, to avoid any spurious correlation that could be generated by incorrectly imputing values both for income and for our dependent variable, e.g. for the expenditures on collected fuels, we estimate the fuel-use Engel curves with quantities as the dependent variable.

4.1.3. Controlling for prices

The conceptual difficulties in appropriately defining fuel prices are compounded by the fact that for many households expenditures on modern fuels are simply not reported. While we use quantities as our dependent variable for the Engel curves, we cannot ignore the fact that we do not observe prices, since omitting prices from the Engel curves may introduce biases in the estimated income coefficients.

Whether it would depends on whether prices and incomes are correlated. A positive correlation between price and income might arise from unobserved quality variation. If higher income households use higher quality fuels that cost more, they may cut down on the quantity consumed. A negative correlation might arise because "access costs" are lower in richer areas. In either case, with incomplete controls for prices income would pick up some of the price effect, biasing its coefficient.

We use area and month fixed effects to control for price variation. Areas are defined as mentioned above (see Section 4.1.1) when discussing the issue of access. The month fixed effects are controlling for the month in which the household was interviewed, and this may partially control for seasonal price variations.

4.1.4. Substitutability across and aggregation of fuels

An interesting feature of the fuel choice behavior in the PIHS is that for most fuels, at a *given level of income* we observe both households with significant consumption and households with zero consumption. For example, while low income household are less likely to use natural gas, some low income households use a lot of wood and no dung, while others use a lot of dung and no wood. This reflects the fact that wood and dung are relatively good substitutes within the traditional fuels.

Substitutability *within categories* of fuels raises the issue of the appropriate level of aggregation of the fuels for estimation. In estimating Engel curves for each fuel, we ignore information about the other fuels used by the households. For instance, in the wood-use equation, the Tobit estimation procedure tries to fit a regression line that accomodates both high-income households who use no wood as well as low-income households who use no wood, without recognizing that in the former case the households use natural gas instead, whereas in the latter, the households use dung. What this does is raise the estimate of the variance of the disturbance term in the Tobit regression, and potentially biases the income coefficient.

To address this problem, we aggregate the fuels into two categories: modern (i.e., natural gas, LPG, and kerosene), which in this context represents the "clean" fuel, and traditional (wood, dung, and other biomass), here considered the "dirty" fuel. We exclude coal and charcoal, as they are little used, and electricity, as it is almost never used for cooking. We use conversion factors from HESS (1993) to convert our varied fuel-quantity measures into a common energy unit, megajoules, and estimate Engel curves for traditional-fuel and modern-fuel megajoules. We have estimated the disaggregated-fuels Tobit regressions, and are happy to provide them. However, given the aggregation issue, income effects in those results that match the aggregated-fuels Tobits, and the value of further exploration including of the Tobit - "generalized Tobit" specification choice, we focus on the results for the aggregated-fuels categories.

4.2. Estimation: fuel-use Engel curves

4.2.1. Specification

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

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

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.[15] 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. As seen above, and as in theory (Pfaff, Chaudhuri and Nye 2002a), 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.

Cragg 1971 considers the excessive restrictiveness of the Tobit specification, and proposes one approach to "generalized Tobit" (a term used in varied ways). We implement his approach, which involves separating the probit (use/nonuse) and

---

[15] 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.

truncated (quantity conditional on use) regressions. Previewing Table 4, for traditional fuels a likelihood ratio test comparing Cragg with Tobit always soundly rejects the Tobit. The gains from this additional flexibility are not surprising, given the clear theoretical predictions and the descriptive statistics for traditional fuels.

4.2.2. Aggregated fuel-use Engel curves

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 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.[16]

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

---

[16] 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.

households will, in general, use more fuel does not 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 energy) to cook for a household of five than it does to cook for a single individual.[17] 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.[18] 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.[19] 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

---

[17] 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.

[18] 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.

[19] 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.

traditional and more likely to use modern fuels, as we find in Table 4.[20]

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.

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

---

[20] 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.

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 and to alleviate any concerns that results are being driven by rural-urban differences in levels of access and prices.

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.[21] 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.[22] 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

---

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

[22] 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.

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_h^t and q_h^m 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, $\widetilde{p}_h$, takes the form:

$$\widetilde{p}_h = \left(\rho q_h^t\right)^\theta + \left(q_h^m\right)^\theta \tag{4.2}$$

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.[23] The higher the value of ρ, the more polluting 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:

$$\widetilde{p}_h = \alpha + \beta_1 y_h + \beta_2 (y_h)^2 + \beta_3 (y_h)^3 + \beta_4 (y_h)^4 + \beta_5 d_h + \Sigma_m \gamma_m M_h^m + \Sigma_k \delta_k A_h^k + \epsilon_h \tag{4.3}$$

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.[24] 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

---

[23] 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.

[24] 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.

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.[25] In Figure 5, this ratio is assumed to be 100, while in Figure 6, it 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 of 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.

5.1. Household preferences and abatement technologies

Consider an economy with a large number of households with varying income levels. Household preferences are described by the utility function $U(c, q)$ where c is the household's consumption of a composite good, and q is the level of environmental

---

[25] 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.

quality the household enjoys. We make the standard assumptions that:

$$\begin{aligned}\text{(i) } U_c &> 0 \quad \text{(ii) } U_{cc} < 0 \quad \text{(iii) } U_q > 0 \quad \text{(iv) } U_{qq} < 0 \quad \text{(v) } U_{cq} > 0 \\ &\text{(vi) } U_{qq}U_{cc} - U_{qq}^2 > 0 \quad \text{(vii) } \lim_{c \to 0} U_c(c,q) = +\infty \qquad (5.1)\end{aligned}$$

Now suppose that the environmental quality that a household enjoys depends not just on its own consumption level and abatement expenditures—as above—but also on the consumption and abatement of other households.[26] Let C denote overall consumption in the economy, and A denote aggregate abatement expenditures. The environmental quality enjoyed by a household h is then $q_h = q(A, C)$. We assume that:

$$\begin{aligned}\text{(i) } q_A &> 0 \quad \text{(ii) } q_{AA} < 0 \quad \text{(iii) } q_C < 0 \quad \text{(iv) } q_{CC} < 0 \quad \text{(v) } q_{AC} > 0 \\ &\text{(vi) } q_{AA}q_{CC} - q_{AC}^2 > 0 \quad \text{(vii) } q(0,0) = Q > 0 \qquad (5.2)\end{aligned}$$

If the number of households is large enough so that each individual household ignores its effects on environmental quality, it is clear that no single household will choose to independently incur any abatement expenditures. 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 simple voting mechanisms can coordinate individual decisions. We consider a simple 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.

5.2. Rising preferred tax rates and the median voter theorem

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

$$\max_{0 \le t \le 1} V(t; y) = U((1-t)y, q(tY, (1-t)Y) \qquad (5.3)$$

where y is the household's income, t is the proportional tax rate, and Y is aggregate income. The first-order condition for this maximization problem is:

$$yU_c \ge U_q Y[q_A - q_C] \qquad (5.4)$$

which holds with equality if $t > 0$.. Let $t^*(y; Y, Q) \in [0, 1]$ denote the preferred tax rate of a household with income y. The assumptions (5.1) and (5.2) ensure:[27]

$$t^*(y; Y, Q) = 0 \qquad \text{for all } y \le \widehat{y}(Y, Q) \qquad (5.5)$$

where $\widehat{y}(Y, Q)$ is implicitly defined by:

$$\widehat{y}U_c(\widehat{y}, q(0, Y)) = U_q(\widehat{y}, q(0, Y))Y[q_A(0, Y) - q_C(0, Y)] \qquad (5.6)$$

---

[26] In this more general context we adopt a more general representation of the abatement technology and abatement expenditures than we did in Section 2. Fuel-switching represents a particular form of abatement, with the "abatement expenditures" being the increased cost of obtaining services implied by the switch from cheaper and dirtier traditional fuels to cleaner but more expensive modern fuels.

[27] The proof is straightforward. Details are provided in Pfaff, Chaudhuri and Nye 2002b.

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. Furthermore, for $y > \widehat{y}(Y, Q)$, we have that:

$$\frac{\partial t^*(y; Y, Q)}{\partial y} = \frac{U_c[\frac{(1-t)yU_{cc}}{U_c} + 1] - U_{qc}(1-t)Y[q_A - q_C]}{U_{qq}Y^2[q_A - q_C]^2 - 2yYU_{cq}[q_A - q_C] + y^2U_{cc} + U_qY^2[q_{AA} - 2q_{AC} + q_{CC}]} \tag{5.7}$$

The denominator of (5.7) is negative given the assumed concavity of $U(c, q)$ and (5.2). Standard assumptions about preferences do not, however, pin down the sign of the numerator. But if we make the additional assumption that:

$$\frac{-cU_{cc}}{U_c} > 1 \qquad \text{for all } c \tag{5.8}$$

the numerator will also be negative, and the preferred tax rate rises (weakly) monotonically with household income. This additional assumption essentially ensures that preferences for consumption are sufficiently elastic.[28] 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.

Since preferences are single-peaked, and the preferred tax rises (weakly) monotonically with income, the median voter theorem applies. The tax rate which will emerge from the simple majority voting procedure is the tax rate preferred by the median voter, in this case the household with the median level of household income.

5.3. Comparative statics

Let y_m denote the median household income. The prevailing tax rate will be given by $t^*(y_m; Y, Q)$. Note that Y and Q affect the prevailing tax rate through two channels: the threshold level of income below which a tax rate of zero will be preferred; and the magnitude of the preferred tax rate when the tax rate is positive.

The framework presented here generates a rich set of comparative static effects, and in particular, suggests that the link between economic growth and the environment is multifaceted, varying with the specifics of the growth process and the particular environmental amenity of interest. That in turn has obvious implications for empirical analyses using aggregate data. For instance, we can trace through the partial effect—everything else held constant—of an increase in the endowment of environmental quality on the tax rate, the level of abatement expenditures and the ultimate level of environmental quality. Such an exercise is useful in thinking about how policy might differ across environmental amenities with differing endowments. To see this, write:

$$q(A, C) = q(t^*(y_m; Y, Q)Y, (1 - t^*(y_m; Y, Q))Y)$$

from which we get:

$$\frac{\partial q(t^*(y_m; Y, Q)Y, (1 - t^*(y_m; Y, Q))Y)}{\partial Q} = [q_A - q_C]Y \frac{\partial t^*(y_m; Y, Q)}{\partial Q}$$

---

[28] Stokey (1998) requires such an assumption for an EKC in a representative agent framework, given a specific abatement technology. Pfaff, Chaudhuri and Nye 2002b do not require this.

Or 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. Here the relevant expression would be:

$$\frac{\partial q(t^*(y_m;Y,Q)Y,(1-t^*(y_m;Y,Q))Y)}{\partial Y} = [q_A t^* + q_C(1-t^*)] + [q_A - q_C]Y\frac{\partial t^*(y_m;Y,Q)}{\partial Y}$$

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. And in that case the relevant expression would be:

$$\frac{\partial q(t^*(y_m;Y,Q)Y,(1-t^*(y_m;Y,Q))Y)}{\partial y_m} = [q_A - q_C]Y\frac{\partial t^*(y_m;Y,Q)}{\partial y_m}$$

And finally we might consider the impact of income growth when it is equiproportionately distributed, implying that the ratio of y_m to Y remains unchanged.

5.4. An illustrative example

To convey the sorts of results this framework can yield, we briefly work through this last comparative static, drawing upon the detailed example that appears in Pfaff, Chaudhuri and Nye 2002b. We assume that environmental quality is given by:

$$q = Q + \gamma A - \delta C \qquad \gamma, \delta > 0 \tag{5.9}$$

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 equiproportionately distributed, or in other words, that the ratio N remains unchanged over time. Household preferences are as described above.

As in (5.6), $\widehat{y}(Q,Y)$ is the threshold level of income below which a household will prefer a tax rate of zero. Our assumptions about preferences imply that as aggregate income falls to zero, $\widehat{y}(Q,Y)$ goes to infinity. As income goes to infinity, $\widehat{y}$ goes to zero, while in between $\widehat{y}$ falls monotonically with aggregate income. Thus, there exists a y_l such that for $y_m < y_l$, a tax rate of zero is preferred ($t^*(y_m;Y,Q) = 0$). Thus, there are no tax-financed public abatement expenditures, and the effect of increasing income is to raise consumption levels and lower environmental quality.

To characterize the impact of increases in y_m beyond the threshold y_l, we begin by noting that $\Delta \equiv \gamma t - \delta(1-t)$ represents the direct impact of an increase in aggregate income Y on environmental quality, given an initial tax rate of t. We show that if $\Delta < 0$, as will be the case at low values of t, the preferred tax rate for a household with income y rises with aggregate income Y. Even if their own incomes are unchanged, households recognize that aggregate incomes have increased and that, at the 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.

From $t = 0$ at the threshold y_l, the preferred tax rate will rise with increases in income until income reaches an upper threshold y_h defined implicitly by:

$$t^*(y_h, Ny_h, Q) = \frac{\delta}{\gamma + \delta} \tag{5.10}$$

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 (5.10) as long as y_m remains above y_h.

Next we consider the impact of increasing median income on environmental quality. *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 $Y = Ny_m$. This has a direct impact Δ plus an indirect impact realized through the change in the preferred tax rate induced by the increase in Y. When y_m is below the upper threshold y_h, the net impact is ambiguous. However, it is easily verified that once y_m crosses the upper-threshold y_h, 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_h where $0 < y_l < y_h$, such that environmental quality falls with income when median income is below y_l and rises when median income is above y_h.

6. Conclusion

Indoor air quality is a major health issue in developing countries, one driven by households' fuels choices. This paper empirically characterized fuels choices, for insights into the behavioral linkages between household preferences, levels of income, and environmental degradation. First we showed in a simple household-production model why, if indoor air quality is a normal (or a luxury) good, the *possibility* of "household-level EKCs" (i.e., non-monotonic Engel curves for indoor air quality) arises quite 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). Using plausible assumptions regarding the emissions implied by fuel use, we found that 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 of particular interest given the rarity of cases in which households significantly internalize pollution impacts, revealing their valuation of environmental quality, here indoor air quality which is so significant for health.

The EKC ("environmental Kuznets curve") literature presents cross-national statistical analyses linking income and ambient air quality. Such analyses are unable to distinguish any particular mechanism through which such a result might come about, although explanations (such as Seldon and Song's "regulatory J curve") have been offered in terms of political-economic mechanisms and sectoral adjustment patterns. A natural question is what can be taken from this paper that is of relevance to that literature. Our first response is to concede that we have not performed the explicit empirical aggregation which would directly tie our household-level work to the aggregate phenomena. However, by presenting both empirical evidence of an analogous household-level phenomenon and a household-level microtheoretic explanation distinct from exogenous government intervention or sectoral adjustment, we have provided perspective on the microfoundations for environmental Kuznets curves more generally, including an additional perspective on the fact that many empirical investigations have not turned up such EKC results. And analysis of household behaviors provides relevant perspective on the EKC mechanisms *either alongside or underlying* political-economic mechanisms. To demonstrate that such household behavior may underlie government intervention, we explicitly modeled how aggregate-level EKCs could arise from voting, given the type of household preferences that we observe to be revealed in fuels choices.

A direction for research arises from the fact that in the fuel-type regressions, we ignored the fact that the decisions for each type are made simultaneously and the fuel types are substitutes (some individual fuels are close substitutes, the fact which drove our use of aggregated fuel types). One could consider econometrically the integration, facing a menu of substitutes, of the binary (extensive margin) use decision with the continuous (intensive margin) decision about how much of a fuel to use conditional on use. Another research issue is that our forecasts of fuel choice, conditional on income growth, take as given the set of fuels used for cooking. More generally, we take as fixed the mix of products involved in the cooking process, despite the assumption (implicit in such forecasts) of economic development underlying the income growth. We could in future research formally consider the possibility of shifts in the mix of cooking-related products along the development path.

Data Appendix

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

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

[5] 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.

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

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

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

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

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

[11] 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.

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

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

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

[15] 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.

[16] 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.

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

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

[19] 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.

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

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

[22] Kahn, M.E.(1998), “A Household Level Environmental Kuznets Curve,” *Economics Letters*

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

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

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

[26] ________(1966b), “Change and innovation in the technology of consumption,” *American Economic Review*, Vol. 56, pp. 14-23.

[27] 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.

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

[29] Pfaff, A.S.P. (2001, revision requested). “Regional Resource Benefits of Urban 'UnSprawl' ”. *Land Use Policy.*

[30] 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.*

[31] 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.

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

[33] 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.

[34] 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.

[35] 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.

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

[37] 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.

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

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

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

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

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

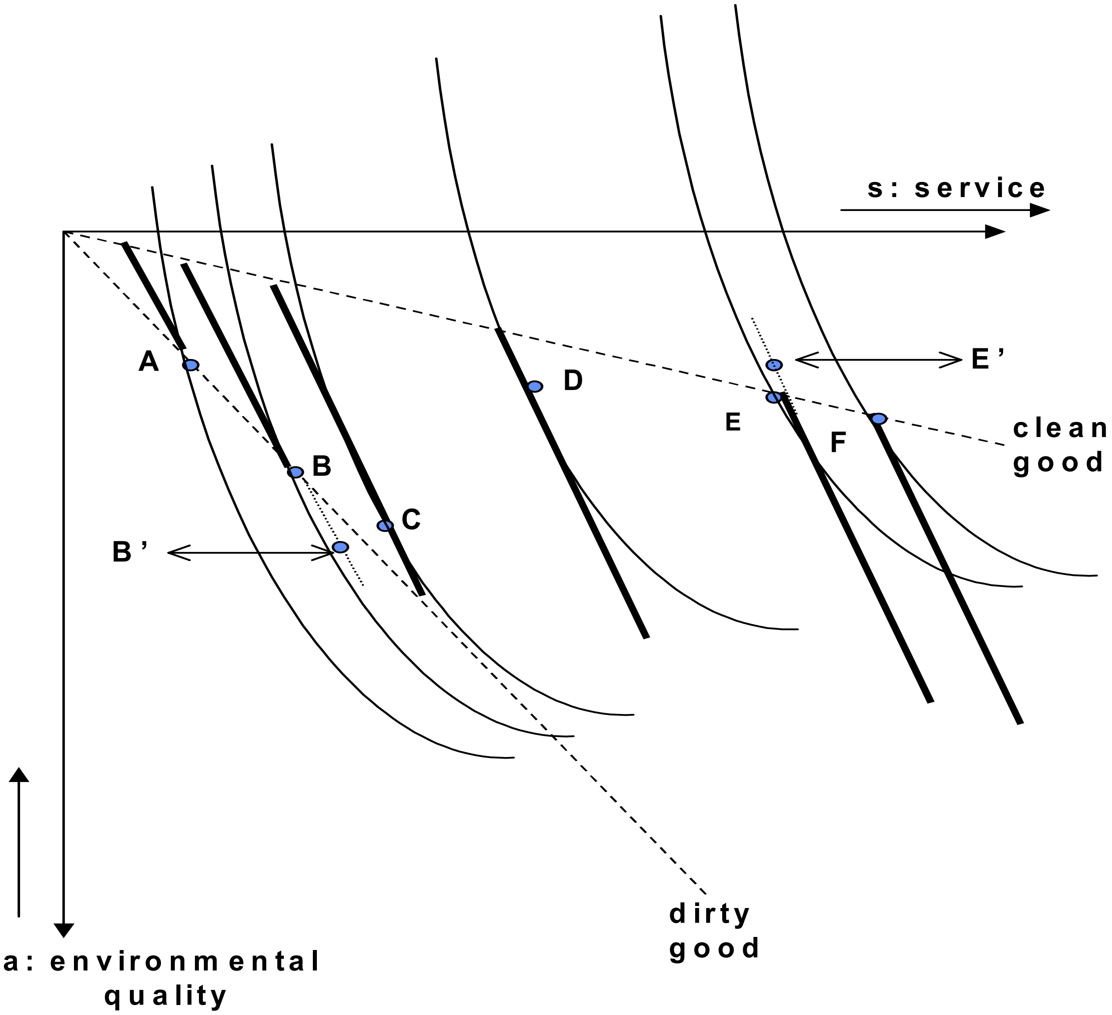

Figure 1

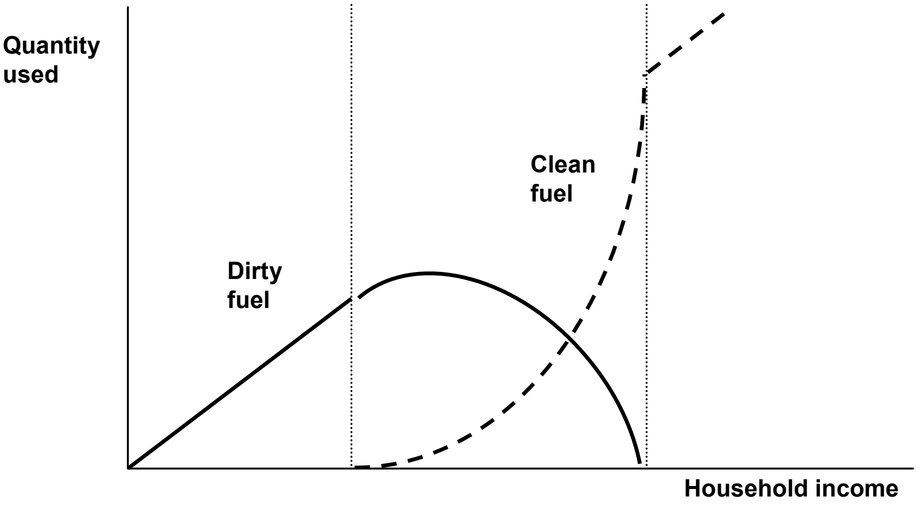

Figure 2

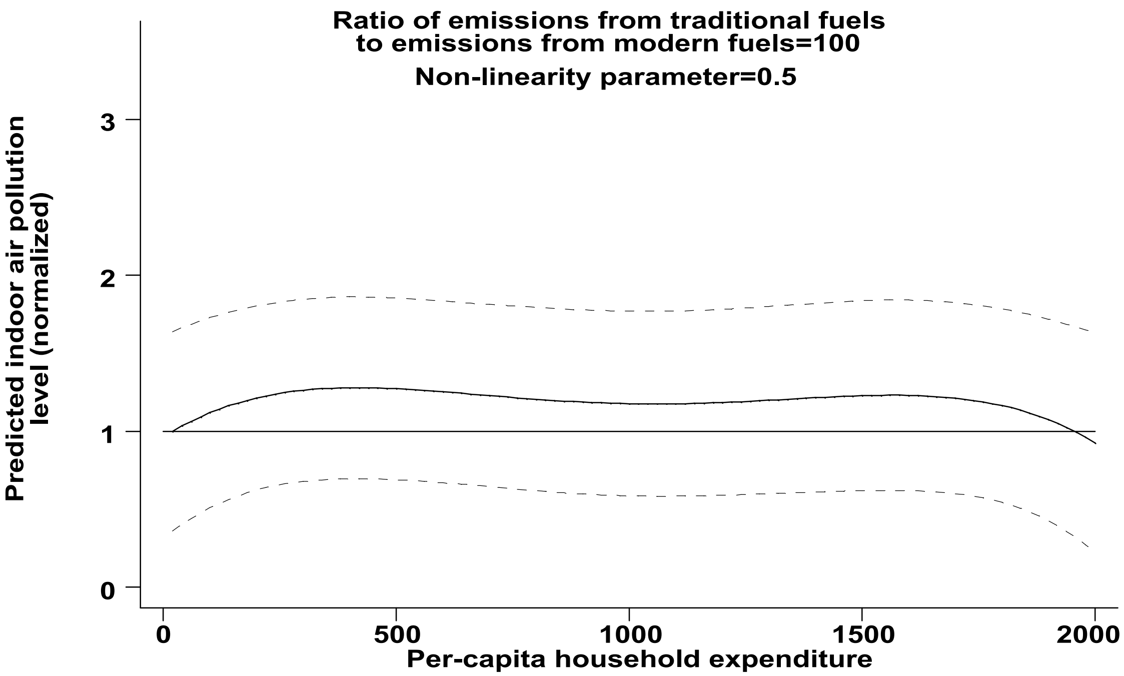

Figure 3

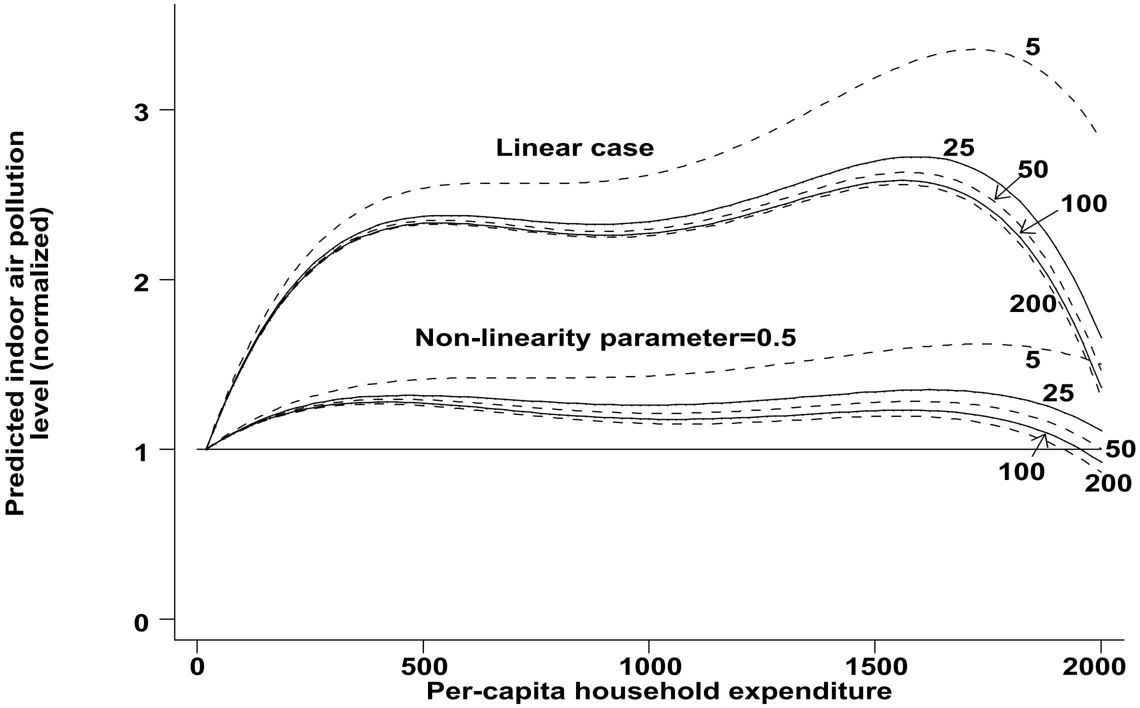

Figure 4
Predicted indoor air pollution levels under alternative assumptions regarding the ratio of emissions from traditional fuels to emissions from modern fuels: non-linearity parameter = 1, non-linearity parameter = 0.5

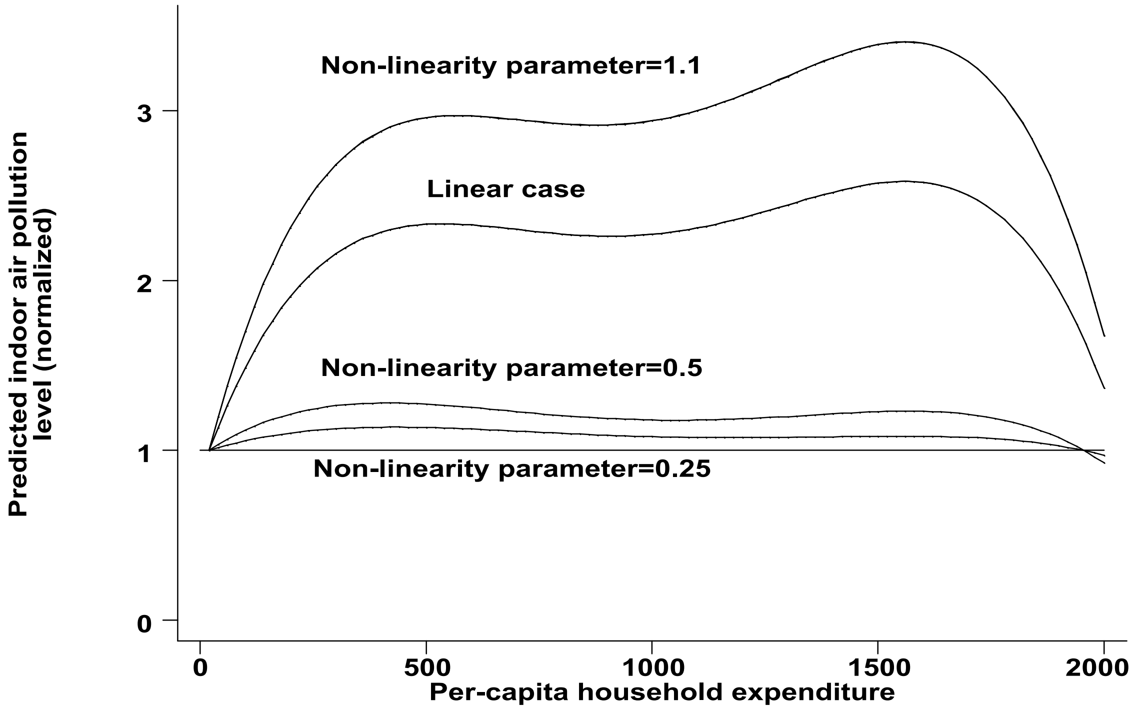

Figure 5
Predicted indoor air pollution levels under different assumptions
about the degree and direction of non-linearity in the emissions function:
emissions ratio = 100

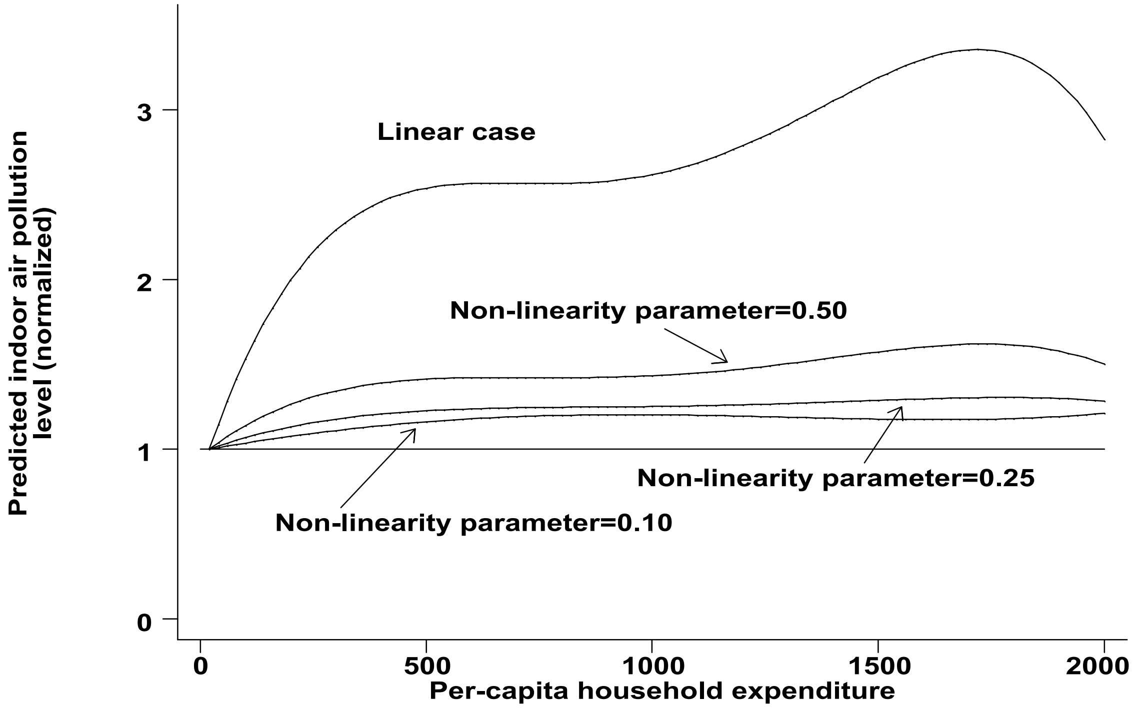

Figure 6
Predicted indoor air pollution levels under different assumptions
about the degree and direction of non-linearity in the emissions function:
emissions ratio = 5

Table 1
Summary statistics

	Sample				
Means	**Full**	**Rural**	**Urban**	**Poor**	**Rich**
Household income	4608	3295	5971	1434	7777
Per-capita income	728	447	1019	177	1139
Household expenditure	4957	4122	5821	3935	5978
Per-capita expenditure	770	582	964	486	876
Food expenditure	1967	2064	1865	1886	2048
Fuel expenditure	203	124	284	160	245
Household size	7.5	7.6	7.3	8.1	6.8
No. of adults	3.7	3.7	3.7	3.6	3.8
No. of kids	3.8	4.0	3.7	4.6	3.1
Education of head	6.2	6.3	6.2	6.2	6.3
Age of head	45.9	45.7	46.1	45.5	46.3
No. of rooms	2.5	2.3	2.6	2.3	2.7
Proportion of households that use:					
electricity (for cooking)	0.76(0.02)	0.58(0.01)	0.95(0.02)	0.67(0.01)	0.85(0.03)
gas (for cooking)	0.18(0.18)	0.01(0.01)	0.36(0.36)	0.07(0.07)	0.28(0.28)
LPG (for cooking)	0.08(0.08)	0.04(0.04)	0.12(0.11)	0.05(0.05)	0.11(0.10)
kerosene (for cooking)	0.73(0.17)	0.90(0.08)	0.56(0.26)	0.83(0.13)	0.64(0.21)
wood (for cooking)	0.70(0.64)	0.91(0.83)	0.49(0.45)	0.82(0.75)	0.59(0.54)
dung (for cooking)	0.45(0.41)	0.67(0.62)	0.22(0.19)	0.58(0.54)	0.32(0.28)
charcoal (for cooking)	0.08(0.01)	0.09(0.01)	0.01(0.01)	0.09(0.01)	0.07(0.01)
coal (for cooking)	0.01(0.00)	0.01(0.00)	0.00(0.00)	0.01(0.00)	0.01(0.00)
biomass (for cooking)	0.23(0.22)	0.36(0.34)	0.10(0.09)	0.30(0.29)	0.16(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
No. of households	4650	2366	2284	2323	2332

Notes: expenditure and income are per month; households in the ‘poor’ (‘rich’) subsample are those with per-capita expenditures below (above) the median for the full sample.

Table 2
Fuel use by expenditure percentiles: disaggregated

	Poorest quartile	Middle quartiles	Richest quartile
Fuel use for cooking: proportion of households			
natural gas	0.06	0.14	0.37
LPG	0.05	0.07	0.13
kerosene	0.08	0.16	0.19
wood	0.81	0.72	0.45
dung	0.60	0.45	0.22
Other biomass	0.32	0.24	0.12
Fuel use for cooking: averages per capita per month			
natural gas (hours)	0.97	2.82	11.77
LPG (hours)	0.68	0.93	2.38
Kerosene (hours)	1.66	4.05	6.68
Wood (kgs.)	30.04	28.76	32.53
Dung (kgs.)	18.61	20.66	23.24
Other biomass (kgs.)	18.88	23.85	25.91

Notes: averages reported are unconditional means, i.e., include households with zero consumption.

Table 3
Multiple fuel use by expenditure percentiles: aggregated

	Expenditure quartile		
Proportion of households that use:	Poorest quartile	Middle quartiles	Richest quartile
one fuel	0.32	0.41	0.64
two fuels	0.46	0.41	0.26
three fuels	0.21	0.16	0.09
four fuels	0.01	0.01	0.01
five fuels	0.00	0.00	0.00
only natural gas or LPG	0.07	0.15	0.41
only natural gas or LPG, and kerosene	0.01	0.01	0.03
only kerosene	0.02	0.05	0.08
only traditional fuels and kerosene	0.05	0.09	0.07
only traditional fuels	0.82	0.65	0.35
both modern fuels and traditional fuels	0.03	0.04	0.05

Table 4
Distinguishing the fuel-choice and fuel-use decisions

BTUs From "Traditional" Fuels (dung, biomass, fuelwood) *

Method	Tobit	Cragg's Probit	Cragg's Truncated[a]	Tobit	Cragg's Probit	Cragg's Truncated[a]
Specification	linear	linear	linear	quartic	quartic	Quartic
Expenditure/pc	*-.057 (1.94)*	*-.001 (11.8)*	*.450 (6.83)*	*1.07 (1.80)*	*-.004 (1.90)*	*2.64 (1.89)*
(Expendit./pc)^2	---	---	---	-.002 (1.48)	4E^{-6} (0.9)	-.003 (1.20)
(Expendit./pc)^3	---	---	---	1E^{-6} (1.3)	-2E^{-9} (0.6)	2E^{-6} (1.1)
(Expendit./pc)^4	---	---	---	-3E^{-10} (1.3)	3E^{-13} (0.4)	-6E^{-10} (1.1)
HouseholdSize	-43.3 (15.9)	-.013 (1.51)	-137 (16.6)	-41.9 (15.3)	-.018 (2.00)	-120 (14.7)
Constant	1243 (17.9)	3.05 (9.15)	1348 (9.63)	1015 (8.74)	3.80 (7.61)	804 (3.03)
MonthEffects [b]	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000
AreasEffects [b]	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000
Fit/Comparison	R$^2_{OLS}$=34%	R$^2_{OLS}$=47%	R$^2_{OLS}$=25%	R$^2_{OLS}$=34%	R$^2_{OLS}$=47%	R$^2_{OLS}$=26%
# observations	3876	3876	2888	3876	3876	2888

BTUs From "Modern" Fuels (kerosene, LPG, natural gas) *

Method	Tobit	Cragg's Probit	Cragg's Truncated[c]	Tobit	Cragg's Probit	Cragg's Truncated[c]
Specification	linear	linear	linear	quartic	quartic	Quartic
Expenditure/pc	*3.02 (14.5)*	*.001 (15.2)*	*5.38 (6.33)*	*20.0 (3.75)*	*.004 (2.24)*	*13.7 (2.12)*
(Expendit./pc)^2	---	---	---	-.023 (2.38)	-2E^{-6} (0.5)	-.024 (2.10)
(Expendit./pc)^3	---	---	---	1E^{-5} (1.8)	-6E^{-10} (0.2)	2E^{-4} (2.1)
(Expendit./pc)^4	---	---	---	-2E^{-9} (1.4)	4E^{-13} (0.6)	-4E^{-9} (2.0)
HouseholdSize	58.5 (2.81)	.046 (6.72)	-75.5 (0.75)	70.9 (3.37)	.052 (7.56)	-82.9 (3.47)
Constant	-7009 (10)	-2.53 (12.8)	-17250 (4)	-10751 (9)	-3.39 (9.46)	-1273 (0.9)
MonthEffects [b]	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000
AreasEffects [b]	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000	p = .0000
Fit	R$^2_{OLS}$=29%	R$^2_{OLS}$=47%	R$^2_{OLS}$=26%	R$^2_{OLS}$=30%	R$^2_{OLS}$=47%	R$^2_{OLS}$=26%
Observations	4331	4331	1562	4331	4331	1562

*: in parentheses next to the coefficients are t statistics, or ratios of the coefficients to the coefficients' standard errors

[a]: likelihood-ratio tests comparing these Cragg runs to the single Tobit regression clearly reject the Tobit (p = .0000)

[b]: we report p values for joint tests of significance for 11 month dummies and the area (spatial-fixed-effect) dummies

[c]: for linear, not all convergence criteria satisfied; for quartic, report OLS coefficients and tests (MLE did not converge); in all our robustness checks for modern fuels, β^{ExpPC} is positive and significant for both the use and quantity decisions

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 (14.245)	0.001409 (9.992)
Per-capita household expenditure squared	4.60E-07 (10.147)	-2.95E-07 (6.179)
Per-capita household expenditure cubed	-3.43E-11 (8.266)	2.01E-11 (4.806)
Per-capita household expenditure to the fourth	5.24E-16 (7.775)	-2.99E-16 (4.453)
No. of adult males in the household	-0.03428 (1.499)	-0.00847 (0.324)
No. of adult females in the household	-0.09513 (3.396)	0.078999 (2.453)
No. of boys (below age 15) in the household	-0.01279 (0.783)	-0.02274 (1.154)
No. of girls (below age 15) in the household	-0.05845 (3.720)	0.014425 (9.992)
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 (10.491)	0.001316 (8.625)
Per-capita household expenditure squared	4.19E-07 (6.626)	-2.46E-07 (4.756)
Per-capita household expenditure cubed	-2.96E-11 (5.035)	1.54E-11 (3.437)
Per-capita household expenditure to the fourth	4.44E-16 (4.642)	-2.23E-16 (3.096)
No. of adult males in the household	-0.04045 (1.366)	0.012029 (0.425)
No. of adult females in the household	-0.13914 (3.712)	0.119538 (3.365)
No. of boys (below age 15) in the household	-0.01271 (0.589)	-0.02198 (1.027)
No. of girls (below age 15) in the household	-0.04218 (1.992)	0.007585 (0.359)
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.