Examining the Impact of Demographic Factors On Air Pollution

Matthew A. Cole
University of Birmingham

Eric Neumayer
London School of Economics

This study adds to the emerging literature examining empirically the link between population size, other demographic factors and pollution. We contribute by using more reliable estimation techniques and examine two air pollutants. By considering sulfur dioxide, we become the first study to explicitly examine the impact of demographic factors on a pollutant other than carbon dioxide at the cross-national level. We also take into account the urbanization rate and the average household size neglected by many prior cross-national econometric studies. For carbon dioxide emissions we find evidence that population increases are matched by proportional increases in emissions while a higher urbanization rate and lower average household size increase emissions. For sulfur dioxide emissions, we find a U-shaped relationship, with the population-emissions elasticity rising at higher population levels. Urbanization and average household size are not found to be significant determinants of sulfur dioxide emissions. For both pollutants, our results suggest that an increasing share of global emissions will be accounted for by developing countries. Implications for the environmental Kuznets curve literature are described and directions for further work identified.

KEY WORDS: carbon dioxide; demography and the environment; environmental Kuznets curve; IPAT; sulfur dioxide.

Please address correspondence to Eric Neumayer, Department of Geography and Environment, London School of Economics and Political Science, London WC2A 2AE, UK; e-mail: e.neumayer@lse.ac.uk

INTRODUCTION

This article contributes to the general debate on the link between population growth and the environment by analyzing the impact of demographic factors on two air pollutants. At the same time, it also contributes to a more focused debate on how population size and other demographic factors should be taken into account in future projections of air pollutant emissions. It is relevant to the large and still growing body of literature on the so-called environmental Kuznets curve (EKC), which posits that environmental pollution is first increasing and then decreasing with rising per capita income levels (see for example Cole, Rayner & Bates, 1997; Grossman & Krueger, 1995). In EKC studies, population is typically incorporated only as a scale variable.

Empirical studies which explicitly examine the link between population and pollution in a systematic quantitative manner are very few in number. Cramer (1998, 2002) and Cramer and Cheney (2000) examine the impact of population levels on air pollution in California and conclude that population is closely associated with some sources of emissions but not with others. Cramer's and Cramer and Cheney's focus on a single state in a developed country is interesting, but it also means that the global implications of their work are uncertain. Dietz and Rosa (1997) and York, Rosa and Dietz (2003a, b) focus on carbon dioxide emissions and energy use and, in the context of the Impact-Population-Affluence-Technology (IPAT) model, examine the roles played by population, affluence and technology.[1] They find that the elasticity of CO_2 emissions and energy are close to unity (i.e., a 1% increase in population leads to an approximately 1% increase in CO_2 emissions). They do not estimate how these elasticities may vary with population levels. All of these results are based on cross-sectional data for 1 year only. Finally, Shi (2003), again in the context of the IPAT model, uses a panel of cross-sectional *and* time series data. Shi finds population elasticities for CO_2 of between 1.41 and 1.65, depending on the model used, but does not examine how these may vary with different population levels. While a step in the right direction, Shi's study still estimates results for one pollutant only, CO_2, and also suffers from a potentially severe methodological problem: many of the variables used by Shi, particularly per capita income and CO_2 emissions, have a very strong upward trend over time. As such, they are not covariance stationary—condition required for non-biased and consistent regression results—thereby raising question marks over the validity of the estimated coefficients and elasticities.

Some ''environmental Kuznets curve'' (EKC) studies undertaken by economists have included population density as one of many determinants of pollution concentrations, but have tended to find mixed results (see for example, Grossman & Krueger, 1995; Hilton & Levinson 1998; Panayotou, 1997). None of these studies have investigated the population–pollution relationship further, or examined the wider impact of population levels (as opposed to spatial density) or other demographic factors on pollution.

The aim of this paper is to provide a detailed analysis of the impact of total population size and other demographic factors on air pollution emissions and to correct the weaknesses outlined above. We build on the papers by Dietz and Rosa (1997), Shi (2003) and York et al. (2003a, b) and improve on their studies in a number of ways. First, whereas these three studies examine only CO_2 and energy use, we extend the analysis to sulfur dioxide (SO_2) emissions, a pollutant with very different properties to CO_2 and hence potentially possessing a very different relationship with population. We also estimate results for CO_2 for means of comparison. Second, in contrast to Dietz and Rosa and York et al., we provide a cross-section *and* time-series panel data analysis. This allows us to capture changes over time and permits a more sophisticated research design controlling for latent country effects. Third, while Shi (2003) also uses a panel data approach we correct the methodological weakness with this study by ensuring that our variables are co-variance stationary by using a first-differenced estimator. Our estimated results are therefore consistent and free from bias. Fourth, we investigate the impact of a more comprehensive set of demographic factors on pollution including the age composition, the urbanization rate and the average household size. Many existing econometric studies neglect demographic factors other than total population size. Parikh and Shukla's (1995) analysis of the effect of the urbanization rate on energy use and greenhouse gas emissions in developing countries represents a notable exception in this regard.

METHODOLOGY

At the global level, the relationship between population and environmental impact is not easy to test in a way that leads to reliable and non-spurious estimates. This is because of the lack of data on environmental degradation covering a sufficiently large number of countries over a sufficiently large period of time. It is for this reason that studies have tended to focus on CO_2 emissions and energy use, for which cross-country and time-series data are available. The time dimension is necessary in order to avoid the problem of one-period cross-sectional regressions, which are likely to lead to spurious results if population size or growth or any other explanatory variable is correlated with unobserved or latent country effects.

Like Cramer (1998), Dietz and Rosa (1997), Shi (2003) and York et al. (2003a, b) we use a stochastic and non-tautological version of the famous IPAT model that originated from a dispute between Commoner, Corr & Stamler (1971) and Ehrlich and Holdren (1971):

$$I = f(P, A, T) \tag{1}$$

where *I* is environmental impact, *P* is population, *A* is affluence and *T* is technology. The IPAT model is most famous in its tautological or definitional identity formulation, which follows from Equation (2) if one defines *A* as consumption (*C*) per capita and *T* as pollution per unit of consumption:

$$I \equiv P \times A \times T \quad \text{if } A \equiv \frac{C}{P} \text{ and } T \equiv \frac{I}{C} \tag{2}$$

In such a formulation the model or its linearized version might be useful for accounting or decomposition purposes as in, for example, Bongaarts (1992), Commoner (1991, 1993), Holdren (1991) & Preston (1996), even though there is some dispute as to how this should be done (O'Neill & Chen, 2002). It has also been used in sensitivity analysis for projecting future CO_2 emissions (for example, O'Neill, MacKellar & Lutz, 2001). However, it is not useful for an empirical estimation of the population elasticity (i.e., the proportional change in pollution or environmental impact per given proportional change in population). For such estimation, we need to define the variables in non-tautological terms.

Our starting point for empirical estimation is therefore Equation (3), referred to as the stochastic IPAT model (STIRPAT) by Dietz and Rosa (1997);

$$I_i = aP_i^b A_i^c T_i^d e_i \tag{3}$$

where *a* is a constant, *b*, *c* and *d* are the exponents of *P*, *A* and *T*, respectively, that are to be estimated and *e* is the residual or error term. Subscript *i* denotes the cross-sectional units, namely countries in this paper.

If we now acknowledge the cross-sectional and time-series nature of our data, and express Equation (3) in logarithms so that it becomes additive, we have;

$$\ln I_{it} = a_i + k_t + b(\ln P_{it}) + c(\ln A_{it}) + d(\ln T_{it}) + e_{it} \tag{4}$$

where subscript *t* denotes the time period. Note that, with panel data, our constant, *a*, becomes country specific and can therefore capture country specific (time invariant) determinants of *I* other than *P, A* and *T*. Important examples for such determinants are climatic differences and geographical factors (Neumayer, 2002, 2004). Note also that we now have a time specific constant for each year, *k*, which captures effects which are common to all countries but which change over time, other than *P, A* and *T*.

Equation (4) provides our basic estimating equation, allowing a number of modifications and extensions to examine different aspects of the population–pollution relationship. Our estimation framework can be thought of as a modified version of the traditional EKC framework familiar to economists. The modifications are twofold. First, we do not use emissions per capita as our dependent variable, but use total emissions and include population size as a further explanatory variable. The traditional EKC framework implicitly assumes a population elasticity of one (i.e., population is only a scale factor), which is of course only one possibility and may conflict with reality. Instead, our aim is to estimate this elasticity. Equation (4) therefore represents the more general estimation framework compared to the traditional EKC methodology. Second, contrary to some EKC studies we do not estimate a reduced-form equation in which income is the only explanatory variable, but distinguish between various effects that are reminiscent of the distinction between scale, composition and technique in some EKC studies (e.g., Selden, Forest & Lockhart, 1999).

Note that the linear relationship between (logged) emissions and (logged) population implies that we estimate the direct effect of population on emissions only, but not the indirect effects that might work via the impact of population on either *A* or *T*. This has been critically noted by a number of participants in the ongoing discussion about the usefulness of IPAT (e.g., MacKellar, Lutz, Prinz & Goujon, 1995). Such complex interaction effects are beyond what can be achieved in this paper and are left for future research.

Cramer (2001) is concerned about potential feedbacks of pollution on population. Of course, the direct effect is likely to be small as mortality from pollution is very small (and zero at the moment with respect to CO_2 emissions). At the local level, there might be reason to be concerned about simultaneity bias as pollution might have an effect on net migration (Cramer, 2001, p. 23). However, at the cross-national level we see no reason to be concerned about this question.

When estimating Equation (4) our measures of *I* are carbon dioxide and sulfur dioxide emissions. In keeping with the previous IPAT literature, *A* is measured as per capita GDP. Technology, *T*, is a broad term which is intended to reflect technological, cultural, and institutional determinants of *I*, i.e., anything that could affect *I/C* (emissions per unit of consumption or production). In our standard model we use two measures of *T*, a country's energy intensity (total energy use per unit of GDP) and the share of manufacturing output in GDP. Energy intensity provides a measure of 'energy productivity' and as such should be directly related to the level and types of technology currently in place within a country. Similarly, manufacturing share provides a measure of the industrial structure of an economy, an obvious determinant of impact per unit of production. Energy intensity is partly determined by the sectoral structure of the economy, but we hope to cover the impact of technology *T* more comprehensively by including both variables in our estimations. Other aspects of 'technology' not captured by energy intensity and the manufacturing share will be picked up by the error term, *e*.

As concerns *P*, the most common approach is to simply use total population levels. However, as MacKellar et al. (1995) and others have pointed out, it is not a *priori* clear that only the individual, rather than, say, households or communities, is the relevant demographic unit. To this one can add that a whole range of other demographic factors beyond simple population levels might also impact on emissions. For example, the impact on emissions could differ across age groups, likely reflecting a number of issues, including consumption habits and patterns, work and leisure activities and attitudes to environmental issues (Tonn, Waidley & Petrich, 2001). One would expect that the economically active part of the population between the ages 14 and 64 has a higher impact on emissions than the retired above the age of 64 or the age group encompassing children and adolescents below the age of 14. A higher urbanization rate can also be expected to have a positive impact on emissions due to the typically more pollution intensive behavioral patterns of those in urban areas. For instance, in developing countries in particular, we would expect those in urban areas to utilize cars, motor cycles and buses to a greater extent than those in rural areas. Similarly, agricultural products are transported to the cities, often from places far away—see Parikh and Shukla (1995) who provide an early analysis of the effect of urbanization on energy use and greenhouse gas emissions in developing countries. We agree with O'Neill and Chen (2002, p. 60) who suggest that the effect of urbanization on emissions represents a promising but underdeveloped avenue of research. Lastly, household size can be expected to have an impact on emissions as households with greater size are likely to benefit from economies of scale in using space, energy use and transportation. Cramer (1998) analyses such an effect on local air pollution in California, but it is to be seen whether it holds in a cross-national setting as well.[2] For this reason, we include variables relating to the age structure of population, relating to the urban vs. rural settlement pattern and relating to the average household size in our estimations.

The great advantage of using panel data over a simple cross-sectional sample is that one can control for the country-specific fixed effects a_i. Failure to do so leads to biased estimates if these fixed or latent effects are correlated with the explanatory variables, as is likely to be the case. However, unfortunately the use of panel data also leads to more complications if some or all of the variables in the estimating equation follow a trend over time. Such trending typically implies what econometricians call non-stationarity. One implication of non-stationarity is that the estimated coefficients and their standard errors cannot be trusted. In formal terms, a variable is defined as stationary if its variance and its expected value do not depend on time and the covariance between the value of the variable at time t and at time $t + s$ does not depend on time. Only statistical inference with stationary variables provides valid results. In simple words, this is because if variables are non-stationary then any correlation between the explanatory and the dependent variable could be due to the trending in both variables that is caused by a third variable not included in the model. We tested for the non-stationarity of the variables in our model formally with the help of Levin, Lin and Chu's (2002) unit root test for panel data. For the dependent variables and several of the explanatory variables we could not reject the hypothesis of non-stationarity. Fortunately, it is often the case that if a variable X_t is non-stationary it is still what is called difference-stationary. This means that a transformation of the original variable called first differencing leads to a transformed variable $Y_t = (X_t - X_{t-1})$ that is stationary. A non-stationary variable that is difference-stationary is also said to be integrated of order one, or $I(1)$, whereas the first differenced transformation is integrated of order zero, or $I(0)$. We therefore took first differences of all the variables included in our regressions. We also tested the first-differenced variables and rejected the hypothesis of non-stationarity for all variables.

Our CO_2 sample covers 86 countries and 24 years (1975–1998), providing 2064 observations in total. For SO_2 our data cover 54 countries and 20 years (1971–1990), providing 1080 observations in total. However, because of the first differencing transformation of the variables we lose the first year of the data such that the sample comprises 1978 and 1026 observations, respectively. The number of observations was constrained by the availability of sulfur dioxide emissions, manufacturing share and energy intensity data. Appendix 1 provides a list of countries included in the samples. They make up approximately 82% of world population in the case of CO_2 emissions and approximately 72% of world population in the case of SO_2 emissions. Appendix 2 provides more information on variable definitions and the sources of all data. We computed variance inflation factors, which suggested no reason to be concerned about potential collinearity problems.

ECONOMETRIC ANALYSIS AND RESULTS

Estimation results for CO_2 emissions are provided in Table 1. Column I reports results for the basic model, in which total population size is the only demographic aspect looked at. The results are generally in line with expectations.[3] Since all variables are expressed in logarithms, the estimated coefficients can be interpreted as elasticities. Affluence (GDPpc) has the expected positive impact on emissions and its elasticity is just below one. The manufacturing share (MANFsh) is insignificant, but a higher energy intensity (ENERGYint) is associated with higher emissions. The estimated population elasticity, for example, is close to unity. This confirms the results reported by Dietz and Rosa (1997) and York et al. (2003).

As a next step, we investigate whether the emission elasticity of the population variable changes with population size. Our second model, for which results are reported in column II, therefore allows for a non-linear relationship between population and pollution emissions by including population squared (POP^2). The linear term becomes statistically insignificant, which suggests that the relationship is not quadratic.[4] We therefore only include the linear population term in the estimations that follow. Similarly, we have tested for non-linear relationships of the other explanatory variables, but have not found relevant evidence.[5]

In column III we additionally look at the age composition of population. We add two variables, namely the percentage of population that is below 14 and the percentage of population that is between 14 and 64 years old. Note that the share of elderly people above 64 years cannot be simultaneously added as the three shares add up to one and are therefore collinear. Adding the age composition hardly affects the population elasticity, which remains close to unity. Neither the affluence nor the technology variables are much affected in either this or consecutive estimations and are therefore not further discussed. A higher percentage of the age group between 14 and 64 years old has a positive impact on emissions that is only marginally significant, however. As expected, a higher share of very young people has no statistically significant impact on emissions.

TABLE 1

Estimation Results for CO_2 Emissions

	I	II	III	IV	V
GDPpc	0.877	0.876	0.867	0.863	0.892
	(5.90)***	(5.91)***	(5.79)***	(5.77)***	(5.29)***
MANFsh	0.020	0.013	0.019	0.013	−0.023
	(0.18)	(0.12)	(0.18)	(0.12)	(−0.14)
ENERGYint	0.346	0.351	0.342	0.346	0.543
	(3.58)***	(3.65)***	(3.56)***	(3.61)***	(4.02)***
POP	1.034	−3.054	1.078	0.922	0.980
	(3.27)***	(−1.42)	(3.51)***	(2.79)***	(3.87)***
$(POP)^2$		0.136			
		(2.06)**			
% POP <14			0.172	0.065	0.010
			(0.53)	(0.21)	(0.03)
% POP 15–64			0.995	0.871	0.425
			(1.72)*	(1.49)	(0.86)
% URBAN				0.663	0.700
				(1.90)*	(2.00)**
HOUSEHOLD SIZE					−0.499
					(−2.67)***
R^2	0.06	0.06	0.06	0.06	0.07
Observations	1978	1978	1978	1978	1707
No. of countries	86	86	86	86	86

All variables are held in logged form and estimated in first differences with ordinary least squares (OLS) and panel-corrected standard errors. Coefficients of year-specific time dummies and constant not reported. *t*-values in brackets.
* Significant at 0.1 level; ** at 0.05 level; *** at 0.01 level.

In column IV we take the urban versus rural settlement pattern into account in adding the share of urbanized population. The population elasticity is now slightly below one at 0.92. A higher rate of urbanization has the expected positive impact on emissions. The share of population in the economically active age groups now becomes marginally insignificant. The reason for this could be the high correlation between the two variables (partial correlation coefficient of 0.65).

Lastly, in column V we add the average household size to our model. Note that this variable is available for all countries in the sample, but not over the entire estimation period. Hence the number of observations is smaller in column V than in the other regressions. The population elasticity is again very close to unitary. The urbanization rate maintains its positive and statistically significant impact on emissions. A higher average household size is associated with lower emissions, as expected. Note that the population share of the economically active age groups now becomes more clearly insignificant. This suggests that its initial statistical significance might be entirely due to its correlation with the urbanization rate as pointed out above and its correlation with the average household size (partial correlation coefficient of −0.58). In other words, it would appear that the urbanization rate and average household size are the demographic factors that really matter. Interestingly, the coefficient size of the energy intensity variable rises once average household size is controlled for.

Table 2 reports results for a similar set of estimations, but with SO_2 emissions as the dependent variable. In column I it can be seen that affluence and energy intensity have the expected positive effects on SO_2 emissions, with the income elasticity being again close to unity. The simple linear population term is insignificant, however. In column II we investigate whether this is due to a non-linear effect of population size. The results suggest that this is indeed the case as both the linear and the squared population terms are statistically significant. This indicates that emissions experience a U-shaped relationship with population. Differentiating our estimated equation with respect to population and setting this equal to zero allows us to identify the turning point level of population: The estimated turning point is at around 5.4 million people. Thus, while we do find a U-shaped relationship between emissions and population, population generates an increase in emissions for all populations over 5.4 million (around a quarter of countries in our sample have a population below this threshold). The inclusion of a quadratic term in model (2) means we cannot interpret the estimated coefficients on POP and POP^2 as elasticities, as actual elasticities will in fact depend on the level of population. Elasticities can be calculated by partially differentiating our estimated equation with respect to population. If our equation to be estimated is as follows;

$$\ln I_{it} = a_i + k_t + b(\ln P_{it}) + c(\ln P_{it})^2 + d(\ln A_{it}) + f(\ln T_{it}) + e_{it} \qquad (5)$$

then the elasticity of I with respect to P, which we may call E_p, can be calculated as;

$$E_p = b + 2c(\ln P_{it}) \qquad (6)$$

Equation (6) therefore allows us to calculate elasticities for varying levels of population. The elasticity is −0.86 for countries with a population of one million, approximately the current population size of Swaziland. It is 0.31 at a population of 10 million (approximately the population of Portugal), 1.13

TABLE 2

Estimation Results for SO_2 Emissions

	I	II	III	IV	V
GDPpc	1.038 (4.91)***	1.012 (4.72)***	1.031 (5.27)***	1.034 (5.20)***	1.162 (10.42)***
MANFsh	0.043 (0.47)	0.031 (0.35)	0.030 (0.34)	0.031 (0.34)	−0.015 (0.16)
ENERGYint	0.845 (5.30)***	0.851 (5.34)***	0.856 (5.50)***	0.856 (5.50)***	0.925 (9.69)***
POP	0.501 (0.72)	−7.908 (−1.77)*	−8.495 (−1.82)*	−8.653 (−2.05)**	−6.114 (−1.89)*
$(POP)^2$		0.255 (1.99)**	0.274 (2.12)**	0.280 (2.51)**	0.225 (2.44)**
% POP <14			−0.351 (−0.21)	−0.344 (−0.21)	−0.329 (−0.81)
% POP 15–64			−1.232 (−0.42)	−1.198 (−0.40)	−1.441 (−1.60)
% URBAN				−0.137 (−0.18)	−0.441 (−1.02)
HOUSEHOLD SIZE					−0.257 (−0.88)
R^2	0.09	0.10	0.10	0.10	0.10
Observations	1026	1026	1026	1026	880
No. of countries	54	54	54	54	54

All variables are held in logged form and estimated in first differences with ordinary least squares (OLS) and panel-corrected standard errors. Coefficients of year-specific time dummies and constant not reported. *t*-values in brackets.
* significant at 0.1 level; ** at 0.05 level; *** at 0.01 level.

at a population of 50 million (approximately the population of Myanmar) and reaches 2.66 at a population of one billion (approximately India's current population size). Setting Equation (6) equal to zero and solving for P also provides the level of population at the turning point as referred to above. Our results therefore clearly suggest that the marginal impact of population on sulfur dioxide emissions is an increasing function of the level of population i.e., the greater the level of population, the greater the environmental impact of each additional unit of population.

As a next step, column III reports results from the model that examines the role played by the age structure of the population. Because we found total population size to have a non-linear effect on emissions, we retain the squared term in all estimations. It can be seen that the age structure of population has no statistically significant impact on emissions. Interestingly, the same is true for the urbanization rate added to the model in column IV and the average household size added in column V. We address this striking difference to our results for CO_2 emissions in the following section where we discuss the implications of our findings.

DISCUSSION AND CONCLUSION

The results reported above demonstrate that the link between population size and emissions of environmental pollutants is a complex one. There are clear differences between SO_2 and CO_2 emissions. In the case of CO_2, we found the elasticity of emissions with respect to population to be unity over the entire range of population sizes. We thus confirm the results of Dietz and Rosa (1997) and York et al. (2003a, b) who found a similar elasticity in their one-period cross-sectional sample and provide evidence against Shi's (2003) much higher estimate of between 1.41 and 1.65. One possible explanation is that our approach deals effectively with non-stationarity problems and therefore leads to more valid and reliable results. We have also controlled for urbanization, household size, and age structure, finding all to be statistically significant determinants of emissions (albeit weakly so in the last case).

In the case of SO_2, our results were quite different. We found the population-emissions elasticity to be actually negative for very small population sizes, but to rise rapidly as population increases. None of the other demographic variables were found to be statistically significant. What explains the contrast between the two pollutants? The most likely explanation is that SO_2 and CO_2 emissions differ in their sources. CO_2 emissions are generated by a great variety of economic and consumption activities that are influenced by demographic factors. SO_2 emissions, in contrast, mainly derive from stationary sources and from the production of electricity in particular. On the whole, more SO_2 emissions will be generated for more people, but other demographic factors will not affect emissions. However, deeper factors may also be at work. Holdren (1991) speculates that settlement patterns might change with higher population levels and that economies might have to resort to lower quality energy resources. He also suggests that 'where rates of population growth (. . .) are high, the attendant pressures can swamp the capacities of societies to plan and adapt in ways that could abate or reduce the environmental impacts of energy supply' (p. 249).

These results are relevant to the large body of literature on the EKC. First, in cases such as CO_2 where the population elasticity is close to one, the EKC approach of taking per capita emissions as the dependent variable, thus implicitly assuming a population elasticity of one, will not lead to biased results. In other cases such as SO_2, the EKC approach is likely to lead to biased results as it fails to take into account the fact that the population elasticity is non-constant in population size. Second, in cases like CO_2 where other demographic factors such as the urbanization rate and average household size are significant determinants of emissions, EKC studies will lead to inaccurate results since they typically fail to take such demographic factors into account.

Despite the differences identified, in the case of both pollutants, demographic trends suggest that a rising share of global emissions will be accounted for by developing countries. Continuing population growth is one reason, exacerbated in the case of SO_2 by the fact that developing countries are, on average, much larger than developed ones. In our sample, their mean population size is around 65.3 million people with a median at around 15.2 million, whereas for developed countries the mean is around 33 million and the median at around 9.7 million (see Table 3). In the case of CO_2 emissions, the impact of population growth will be reinforced by other demographic trends. The tendency towards rising urbanization is undisputed; in our sample, the mean urbanization rate for developing countries is around 56% whereas it is 78% in developed countries While less broadly recognized, average household size in developing countries is also likely to fall as young people move away earlier from their family home, marry at a later age and their parents increasingly live in separate homes. The average household size in developing countries is currently about 4.9, but only 2.6 in developed countries. O'Neill et al. (2001, p. 72) report projections that see the average household size in developing countries decreasing to between 2.4 and 3 over this century.

In contrast, in developed countries demographic factors will not change much in the future. Not only do they have low and sometimes zero (or even negative) current and projected population growth rates, but their urbanization rate will not increase as dramatically as in developing countries and their average household size is already very low.

In closing, we would suggest work in two areas. First, non-linearities in the relationship between population and pollution of the sort we identified in the case of SO_2 deserve systematic study. Second, more detailed attention is needed to account for the indirect effects that population growth and other demographic changes can have on the environment. In non-reported analysis we tested for interaction effects of the various demographic variables with our variables of affluence and technology. These interaction effects generally failed to assume statistical significance. This does not imply, however, that interaction and feedback effects are not important.

TABLE 3

Descriptive Information on Demographic Factors of Countries in Sample (1998 Unless Specified Otherwise)

Countries	Mean	Median
Population (million)		
All	56.7	13.2
Developed	33	9.7
Developing	65.3	15.2
Population growth in 1990s (%)		
All	1.77	1.78
Developed	0.63	0.53
Developing	2.19	2.24
Share of under 14 year olds (%)		
All	31.1	32.5
Developed	18.7	18.5
Developing	35.6	35.6
Share of 15–64 year old (%)		
All	61.3	61.8
Developed	66.7	67.9
Developing	59.3	60.2
Urbanization rate (%)		
All	61.7	62.3
Developed	77.9	77.1
Developing	55.9	55.1
Average household size		
All	4.3	4.3
Developed	2.6	2.6
Developing	4.9	4.8

Source: Authors' calculations from sources given in Appendix 2. Developed countries are United States, Canada, Western Europe, Japan, Australia and New Zealand; developing countries are all others in Appendix 1.

Rather, they might be more difficult to take into account adequately in modeling and estimation. Clearly, these questions deserve more systematic attention and the present authors would like to tackle these and related questions in future research.

ACKNOWLEDGEMENT

We would like to thank three anonymous reviewers and the editor for many helpful comments.

NOTES

1. Like the current paper, existing studies mainly look at air pollution for reasons of data availability. Taking a broader focus, York et al. (2003b) examine the impact of demographic factors on so-called ecological footprints. Ecological footprints supposedly measure the total land area hypothetically required to provide all the resources to and absorb all the pollution generated by a country's economy. We do not follow this path. For one reason, this can only be done in a cross-sectional analysis, in which one cannot control for many sources of estimation bias. More importantly, one of us has argued that the concept of ecological footprints does not represent a valid, reliable or methodologically sound indicator (see Neumayer 2003, pp. 172–177, for details).
2. Liu, Daily, Ehrlich and Luck (2003) show that the growing number of households puts pressure on endangered species in so-called biodiversity hotspots.
3. Note that the R^2 values are quite small. This is no reason for concern as R^2 is typically very small for models estimated in first differenced variables.
4. If the squared term is included in the estimations of columns III–V, then both the linear and the squared term are statistically insignificant, buttressing the conclusion that the relationship between population size and CO_2 emissions is linear rather than quadratic.
5. This result might be surprising with respect to the average income level since some EKC studies have found a non-linear relationship of per capita income and CO_2 emissions. Note, however, that the non-linear effect is likely to work through T and that these studies use reduced-form estimations in which emissions are regressed on income without controlling for T.

REFERENCES

Bongaarts, J. (1992). Population growth and global warming. *Population and Development Review, 18*(2), 299–319.

Commoner, B. (1991). Rapid population growth and environmental stress. *International Journal of Health Services, 21*(2), 199–227.

Commoner, B. (1993). Population, development, and the environment: trends and key issues in the developed countries. *International Journal of Health Services, 23*(3), 519–539.

Commoner, B., Corr, M., & Stamler, P. J. (1971). The causes of pollution. *Environment, 13*(3), 2–19.

Cole, M. A., Rayner, A. J., & Bates, J. M. (1997). The environmental Kuznets curve: An empirical analysis. *Environment and Development Economics, 2*(4), 401–416.

Cramer, J. C. (1998). Population growth and air quality in California. *Demography, 35*(1), 45–56.

Cramer, J. C. (2002). Population growth and local air pollution: methods, models, and results. In W. Lutz, A. Prskawetz & W. C. Sanderson (Eds.), *Population and environment* (A supplement to Vol. 28), *Population and development review,* pp. 22–52. New York: Population Council.

Cramer, J. C., & Cheney, R. P. (2000). Lost in the ozone: population growth and ozone in California. *Population and Environment, 21*(3), 315–337.

Dietz, T., & Rosa, E. A. (1997). Effects of population and affluence on CO_2 emissions. *Proceedings of the National Academy of Sciences, 94*(1), 175–179.

Ehrlich, P. R., & Holdren, J. P. (1971). Impact of population growth. *Science,* 171, 1212–1217.

Grossman, G. M., & Krueger, A.B. (1995). Economic growth and the environment. *Quarterly Journal of Economics, 110*(2), 353–377.

Hilton, F. G. H., & Levinson, A. (1998). Factoring the environmental Kuznets curve: evidence from automotive lead emissions. *Journal of Environmental Economics and Management, 35*(2), 126–141.

Holdren, J. P. (1991). Population and the energy problem. *Population and Environment, 12*(3), 231–255.
ITU (2002). *Telecommunications and other statistics on CD-Rom*. International Telecommunications Union.
Levin, A., Lin, C. F., & Chu, S. J. C. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics, 108*(1), 1–24.
Liu, J., Daily, G. C., Ehrlich, P. R., & Luck, G. W. (2003). Effects of household dynamics on resource consumption and biodiversity. *Nature, 421*, 530–533.
Neumayer, E. (2002). Can natural factors explain any cross-country differences in carbon dioxide emissions? *Energy Policy, 30*(1), 7–12.
Neumayer, E. (2003). *Weak versus strong sustainability: Exploring the limits of two opposing paradigms* (Second Revised Edition). Cheltenham and Northampton: Edward Elgar.
Neumayer, E. (2004). National carbon dioxide emissions: geography matters. *Area, 36*(1), 33–40.
MacKellar, F. L., Lutz, W., Prinz, C., & Goujon, A. (1995). Population, households, and CO_2 emissions. *Population and Development Review, 21*(4), 849–865.
O'Neill, B. C., & Chen, B. S. (2002). Demographic determinants of household energy use in the United States. In W. Lutz, A. Prskawetz & W. C. Sanderson (Eds.), *Population and environment* (A supplement to Vol. 28). *Population and development review,* pp. 53–88. New York: Population Council.
O'Neill, B. C., MacKellar, F. W., & Lutz, W. (2001). *Population and climate change*. Cambridge: Cambridge University Press.
Panayotou, T. (1997). Demystifying the environmental Kuznets curve: Turning a black box into a policy tool. *Environment and Development Economics, 2*(4), 465–484.
Pariakh, J., & Shukla, V. (1995). Urbanization, energy use and greenhouse effects in economic development. *Global Environmental Change, 5*(2), 87–103.
Preston, S. H. (1996). The effect of population growth on environmental quality. *Population Research and Policy Review, 15*(2), 95–108.
Selden, T. M., Forrest, A. S., & Lockhart, J. E. (1999). Analyzing the reductions in U.S. Air pollution emissions: 1970 to 1990. *Land Economics, 75*(1), 1–21.
Shi, A. (2003). The impact of population pressure on global Carbon dioxide emissions, 1975–1996: Evidence from pooled cross-country data. *Ecological Economics, 44*(1), 24–42.
Tonn, B. E., Waidley, G., & Petrich, C. (2001). The ageing US population and Environmental policy. *Journal of Environmental Planning and Management, 44*(6), 851–876.
World Bank (2002). *World development indicators on CD-Rom*. Washington, DC: World Bank.
York, R., Rosa, E. A., & Dietz, T. (2003a). STIRPAT, IPAT and ImPACT: analytic tools for unpacking the driving forces of environmental impacts. *Ecological Economics, 46*(3), 351–365.
York, R., Rosa, E. A., & Dietz, T. (2003b). Footprints on the earth: the environmental consequences of modernity. *American Sociological Review, 68*(2), 279–300.

APPENDIX 1

Countries included in the CO_2 estimation results: Algeria, Argentina, Australia, Austria, Bangladesh, Belgium, Benin, Bolivia, Brazil, Brunei, Cameroon, Canada, Chile, China, Colombia, Congo (Dem. Rep.), Congo (Rep.), Costa Rica, Côte d'Ivoire, Cyprus, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Finland, France, Gabon, Ghana, Greece, Guatemala, Haiti, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Luxembourg, Malaysia, Malta, Mexico, Morocco, Nepal, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Pakistan, Paraguay, Peru, Philippines, Portugal, Romania, Saudi Arabia, Senegal, Singapore, South Africa, South Korea, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syria, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, United Arab Emirates, United Kingdom, United States, Uruguay, Venezuela, Zambia, Zimbabwe.

Countries included in the SO_2 estimation results: Algeria, Argentina, Australia, Austria, Belgium, Bolivia, Brazil, Chile, China, Colombia, Denmark, Ecuador, Egypt, Finland, France, Ghana, Greece, Guatemala, Honduras, Hong Kong, Hungary, India, Indonesia, Ireland, Israel, Italy, Japan, Kenya, Luxembourg, Malaysia, Mexico, Morocco, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Peru, Philippines, Portugal, Saudi Arabia, South Korea, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syria, Thailand, Trinidad and Tobago, United Kingdom, United States, Uruguay, Venezuela.

APPENDIX 2

Variable	Source
Sulfur dioxide emissions	ASL and Associates http://www.asl-associates.com/sulfur1.htm
Carbon dioxide emissions Population Age structure of population Urbanization rate GDP per capita Energy intensity Manufacturing share of GDP	All from: World Bank (2002)
Average household size	ITU (2002) and World Bank (2002)