Does an environmental Kuznets curve exist for biodiversity?

Does an environmental Kuznets
curve exist for biodiversity?

Simon Dietz[1]

Institut für Wirtschaftsforschung

Eidgenössische Technische Hochschule Zürich
Weinbergstrasse 35
CH-8006 Zürich

Postadresse:

Institut für Wirtschaftsforschung
ETH Zentrum   WEH
CH-8092 Zürich

Telefon: 0041-1-632 41 28
Telefax: 0041-1-632 12 18
email: library@wif.gess.ethz.ch

Within the past ten years, environmental economics has witnessed the development of a theory that has stirred up considerable controversy. The environmental Kuznets hypothesis predicts negative environmental impact first increases with rising income and then decreases. A wealth of indicators have been used to test this claim but one of the most important global environmental quantities, biodiversity, has remained untouched. The purpose of this study is thus to pose the question, is there an environmental Kuznets curve for biodiversity?

I. Background

The environmental Kuznets hypothesis is, in essence, empirical and predicts an inverted-U relationship between any environmental indicator and income per capita, plotted in two dimensions. In other words, damage to the environment first increases and then decreases with rising income.

As a purely empirical theory, the Kuznets hypothesis itself cannot explain this behaviour. Nevertheless, economists are more or less united on the factors supporting the income-driven relationship, if not on their relative importances. Ekins (1997) summarises the forces at play on three levels. On the first of these, „technique", „composition" and „scale" effects in the economy directly influence the level of environmental impact. These processes are steered by policy making (the second level) and policy in turn is underpinned by the wants and needs of the business community and the general public as a whole (the third).

Heavily disputed, however, are the results and the conclusions some researchers have been led to draw. In fact, evidence of the environmental Kuznets curve is thin - most authors only find an inverted-U shape in a fraction of cases and are hard-pressed to distinguish between linear, quadratic and cubic forms (Ekins 1997). In spite of this, some have been motivated to predict economic growth alone will cure our environmental problems[2] (e.g. Shafik and Bandyopadhyay 1992, cited in Ekins 1997).

---

[2] In fact, such conclusions ignore the point made by Ekins (1997) and others that policy changes underlie the relevant economic factors.

Studies have analysed a wide variety of environmental media, including air pollution (e.g. Selden and Song 1994), water pollution (e.g. Grossman and Krueger 1995, Shafik 1994) and deforestation (e.g. Cropper and Griffiths 1994, Koop and Tole 1999). Not among these, however, is biodiversity, one of the most significant global environmental quantities[3]. This study thus aims to add to the body of literature by asking, „is there an environmental Kuznets curve for biodiversity?"

In achieving this objective, a number of important questions must be answered which themselves generate a great deal of interest. Firstly, just how does one quantify biodiversity, given that the data must be suitable for testing the environmental Kuznets curve? This question is answered in part II. Secondly, would theory predict an environmental Kuznets curve for biodiversity, or are there some inherent barriers to yielding the inverted-U relationship? Part III discusses this point. Finally, in part IV, from an econometric viewpoint, how should the model be estimated?

II. Quantifying biodiversity

Quantifying biodiversity is the first major problem this study encounters. The reasons are twofold. Firstly, data on global biodiversity are extremely scarce and, in general, highly inaccurate. Secondly, whatever data do exist must be suitable for incorporation into the environmental Kuznets framework.

Before these points are elaborated upon, however, it is necessary to define what we interpret as biodiversity. Ecologists generally draw the distinction between three types of biodiversity. *Ecosystem diversity* relates to the variety of different ecosystems in the world (e.g. tundra, taiga), *species diversity* represents the variety of global species and *genetic diversity* pertains to the diversity of genetic information stored within each individual. In global discussions, it is commonly the second of these classes, species diversity, which receives attention and it is accordingly the focus of this study.

The environmental Kuznets curve has commonly been tested using either simple cross-sectional data or panel data - cross-sections through time. In general, panel data is favoured because it allows a very restrictive assumption manifest in cross-sectional analyses to be dropped; the effect on the

---

[3] An exhaustive discussion of the economic importance of biodiversity is given in Pearce et al (1992).

environment of income changes is assumed to be the same for all countries. In statistical terms, this means the regression coefficients are common to all groups in the cross-section[4]. We seek to avoid such an assumption in this case, because local conditions are likely to generate significant differences in the income/environment relationship between countries like, for example, Brazil and India. Therefore only panel data should be considered.

The question which now arises is, do time series cross-sections exist for species diversity? Essentially, the answer is no. The only data that explicitly deal with the number of species are cross-sections[5]. Additionally, and fundamentally important in any discussion of species diversity, we have only classified a fraction of the total number of species on earth. In fact, we do not know the true number even to the nearest order of magnitude (Wilson 1986). Estimates vary from 5 to 100 million. Therefore, we cannot hope to use a direct indicator of species number by nation through time.

How then have other researchers made dynamic estimates of global biodiversity? In almost all cases (Reid 1992), the species-area relationship has been applied. This relationship aims to estimate the number of species within a given area. Like the environmental Kuznets hypothesis, it is empirical and its biological significance must be inferred. This has led to the development of alternative theories behind the relationship and, correspondingly, alternative specifications of the model (Connor and McCoy 1979). Such problems need not, however, concern this study given that a biological justification is not required and that a specific form of equation has become established in the literature:

$$(1) \qquad S = cA^z$$

where S is the number of species, A is area and c and z are constants without meaning. The magnitude of z has been much discussed and most authors agree it lies somewhere between 0.15 and 0.35, (MacArthur and Wilson 1967), with 0.25 taken as a median. c, on the other hand, has apparently been little studied. Accordingly, studies estimating the change in biodiversity eliminate c by investigating the ratio of species in a given year (t) relative to a base year (0):

---

[4] See part 4 for more details. Koop and Tole (1999) also discuss the assumptions inherent in different model specifications.
[5] An example is the World Conservation Monitoring Centre's biodiversity database, available at www.wcmc.org.uk

4

$$S_O = cA_O{}^z$$

$$S_t = cA_t{}^z$$

(2) $S_t/S_O = A_t/A_O{}^z$

The simplicity of the relationship makes it very attractive to researchers of biodiversity dynamics. Unfortunately, it is understandably also subject to a series of assumptions, which cast doubt on the accuracy of results[6].

Given we now have a means to generate species numbers globally through time, the final issue to clear up is what should be taken to be A? In other words, what global ecosystem areas should be considered? It is obviously far too demanding to cover the whole of the earth's land area, both from the point of view of time and of the variations in the species-area relationship's z value with latitude. Such scope is also excessive, as the majority of the world's species is confined to the so-called tropical moist forests of the equatorial regions. Definitions of these forests vary but this study adopts Myers (1980) definition as „evergreen or partly evergreen forests, in areas receiving not less than 100mm of precipitation in any month for 2 out of three years, with mean annual temperature of $+24^oC$ and essentially frost free; in these forests some trees may be deciduous...“

Myers has carried out two global forest inventories based on this classification, but these obviously do not satisfy our requirement of having many points in time. The FAO's own "Forest resources assessment" (FAO 1990) is also limited to two points in time, 1980 and 1990. Thus we have encountered a major stumbling block to obtaining data for this study - the fact that reliable estimates of forest cover in the tropics are practically non-existent.

As each study defines tropical forests in a different way, inter-study comparisons cannot be made. Additionally, individual studies, particularly those carried out by the FAO, are subject to the limited reliability of estimates provided by each country[7]. The only source that will provide a sufficient time

---

[6] For a discussion of these assumptions and a survey of the species-area relationship's use in the literature, refer to Reid (1992).

[7] For a summary of the shortcomings of global (and particularly tropical data), readers are referred to Pearce and Brown (1994).

series is the FAO's Production Yearbook. Unfortunately, from the point of view of biodiversity, this source adopts a broad definition of forests as „all woody vegetations" (Koop and Tole 1999). Thus the „forests and woodlands" statistic covers forests within and outside Myers functional definition. In order to mitigate for this, this study confines itself to the around 40 countries identified by Myers as having „tracts of tropical moist forest that are appreciable in size or are significant for their ecological and biotic values" (Myers 1980)[8]. Unwanted interference from other forest types is thus minimised. Furthermore, Allen and Barnes (1985) demonstrate by rank correlation that the FAO Production Yearbook data are sufficiently similar to other studies of explicitly tropical forests to permit their use in assessing tropical deforestation.

III. Should we anticipate a Kuznetsian relationship?

To reiterate, the environmental Kuznets hypothesis predicts environmental damage first increases and then decreases with rising income. Graphically, the relationship should resemble an inverted-U shape. However, referring to the nature of the environmental indicator used in this study (equation 2), the relationship will be reversed and should resemble a U shape. Otherwise known as a parabola, the regression equation must therefore take the following quadratic form[9]:

$$y = x^2 - x$$

Yet can biodiversity really be expected to exhibit this behaviour? Consider the two halves of our Kuznets curve, the „falling limb" and the „rising limb". The „falling limb" indicates decreasing numbers of species. This is perfectly foreseeable given the well-documented history of extinctions based on habitat loss. We cannot, however, expect biodiversity to be replenished at the same rate. For example, the background species extinction rate has been put at 1 every 100 - 1000 years (Reid 1994). A creation rate of around the same order of magnitude is likely, so theory would not predict a „rising limb".

Instead, we might expect biodiversity to decrease and then *level off* with increasing income, since economic forces should fuel the drive for environmental improvement but biodiversity cannot

---

[8] The final number of countries studied is only 36 due to missing GDP data.
[9] The signs on the coefficients ensure the parabola is U shaped and not an inverted-U.

replenish itself at an equivalent rate.   Thus it would be revealing to estimate a hyperbolic curve as well:
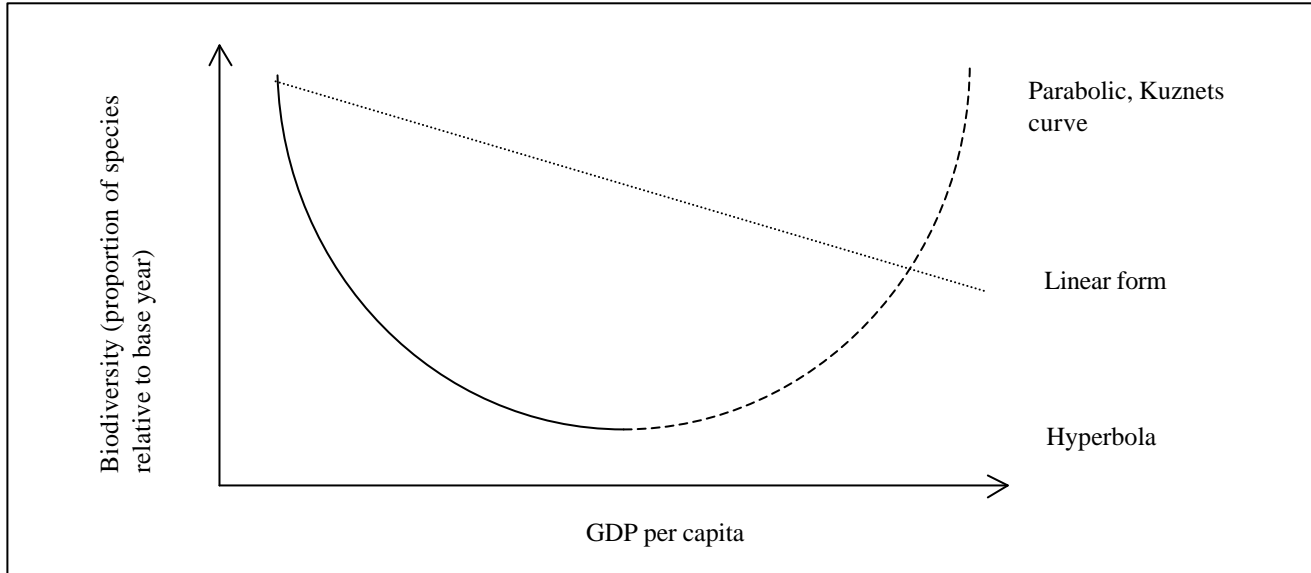
$$y = 1/x$$

Finally, inspection of the data for species diversity also indicates a close clustering of observations around the starting value, independent of income level. In view of this, a linear equation will also be tested:

$$y = -x$$

The conceptual forms of these three curves are presented in Figure 1.

Figure 1. Possible forms of the relationship

IV. Econometric considerations

Various techniques exist to estimate the above equations. The two basic methods available are *fixed* and *random effects.* They both build on the regression equation for simple ordinary least squares (OLS):

$$y(i,t) = a + \sum_{J=1}^{K} b(j)X(i,t,j) + e(i,t)$$

where y is the dependent variable, biodiversity, for country i at time t, α is the equation constant, equal for all countries, β represents the coefficient for independent variable x, where there are j = 1-k variables (each β also being common across all countries) and $\varepsilon(i,t)$ is a classical disturbance with

$$E[\varepsilon(i,t)] = 0,$$
$$Var[\varepsilon(i,t)] = \sigma^2(\varepsilon).$$

Fixed effects loosens up the assumption of commonality across countries by estimating a separate constant α for each country:

$$y(i,t) = a(i) + \sum_{J=1}^{K} b(j)X(i,t,j) + e(i,t)$$

Random effects works in a similar way but assumes that international heterogeneity is randomly (and normally) distributed. Each country now has a disturbance term $\mu(i)$:

$$y(i,t) = a + \sum_{J=1}^{K} b(j)X(i,t,j) + e(i,t) + m(i)$$

where

$$E[\mu(i)] = 0,$$
$$Var[\mu(i)] = \sigma^2(\mu),$$
$$Cov[\varepsilon(i,t), \mu(i)] = 0.$$

The difference between the two models, whether the vertical displacement of the regression equation should be parametric or random - has been debated in the literature (Greene 1997). Fixed effects

imply international differences are generated by country-specific factors not covered by our regressors. Random effects, on the other hand, imply national peculiarities are unimportant and differences should be assumed random. This is a responsible approach if the sample to be studied is part of a much larger population (Greene 1997). Yet here we have an almost complete set of countries containing significant tropical moist forests. Furthermore, intuition suggests environmental and economic factors should play a rather important role. We can test this empirically using the Hausman test[10] but, at this interjection, we would expect fixed effects to be favoured.

Koop and Tole (1999), however, charge both of these specifications with assuming too much commonality and recommend the additional freedom of the *random coefficients* method. In this case, each country possesses its own curve, drawn from a random distribution:

$$y(i,t) = \sum_{J=1}^{K} \boldsymbol{b}(i,j) X(i,t,j) + \boldsymbol{e}(i,t)$$

where

$$\beta(i) = \beta + v(i),$$
$$E[v(i)] = 0.$$

Thus we are no longer bound to the assumption that each country's curve possesses the same shape, yet we are not going so far as to say each country has its own curve which is unrelated to all others (this is the basis of the *fixed coefficients* model)[11].

However, the nature of the data in this study is such that the test for random coefficients is distorted[12]. Consequently we shall confine ourselves to comparing fixed and random effects empirically.

---

[10] The Hausman test is a purely empirical test used to differentiate between random effects and fixed effects. It assesses whether individual effects are correlated with the regressors. If so, random effects are inconsistent. If not, both random and fixed effects are consistent but random effects are more efficient.

[11] For a thorough exposition of all these methods, refer to Greene (1997).

[12] The test in question is a chi-squared test of the form

$$\boldsymbol{c}^2 = \sum_{i=1}^{n} [b_i - \tilde{b}]' \tilde{V}_i^{-1} [b_i - \tilde{b}]$$

To summarise, the variables and equations are as follows:

S(i,t) is the fraction of species in any year compared to the reference year 1970 (equation 2). This tends to cause the data to cluster around the reference mark 1.0, so all values are multiplied by a factor of 1000.

G(i,t) is income per capita. It is always converted to log form[13]. Data comes from the Penn World Table[14].

P(i,t) is population change[15], expressed as a percentage of the previous year. Data comes from the Penn World Table.

D(i,t) is population density, expressed as people per hectare. The variable is a combination of FAO Production Yearbook data on land area and Penn World Table population statistics.

T(t) is a linear time trend (T is simply the relevant year) used to correct the problem that both biodiversity and income display positive time dependency.

F(i,t) is forest area in hectares. The relative impact of deforestation in any country depends on absolute forest area and feeds back into future trends.

ε(i,t) is a classical error term.

The basic regression model for the quadratic equation is

(3) $S(i,t) = \alpha + \beta_1 \ln G(i,t) + \beta_2 (\ln G(i,t))^2 + \beta_3 P(i,t) + \beta_4 D(i,t) + \beta_5 T(t) + \beta_6 F(i,t) + \varepsilon(i,t).$

For the hyperbolic equation, it is

---

where b is the curve in country i, $\tilde{b}$ is the mean of all curves and V_i is a covariance matrix. In this asymptotically distributed data set, V_i is distorted and the test results correspondingly invalidated.

[13] Maddala (1992) recommends a semi-log form of this type when y is increasing more slowly than x.

[14] Summers and Heston (1991) explain the methodology and layout of mark 5 of the table. The most up-to-date version, 5.6, is available for download from the internet at http://datacentre2.chass.utoronto.ca/pwt/

(4) $S(i,t) = \alpha + \beta_1 1/\ln G(i,t) + \beta_3 P(i,t) + \beta_4 D(i,t) + \beta_5 T(t) + \beta_6 F(i,t) + \epsilon(i,t).$

The linear equation is

(5) $S(i,t) = \alpha + \beta_1 \ln G(i,t) + \beta_3 P(i,t) + \beta_4 D(i,t) + \beta_5 T(t) + \beta_6 F(i,t) + \epsilon(i,t).$

V. Results

Tables 1 to 4 report the empirical results of this study. Table 1 displays the outcome of tests between the basic model, fixed and random effects for each equation. Table 2 goes on to report results for the parabola using the model selected, Table 3 results for the hyperbola and Table 4 the linear equation. Econometric details are mostly confined to footnotes.

Table 1. Comparing estimators

	Parabolic equation	**Hyperbolic equation**	**Linear equation**
LM statistics[16]	3201.68**[17]	3389.13**	3491.40**
Hausman statistics	39.16**	28.97**	28.90**

LM test statistics outlined in Table 1 clearly favour both fixed and random effects over simple OLS[18]. The results of the Hausman test build on this foundation by demonstrating fixed effects to be the better estimator in all cases. Thus the prediction made in section IV, that national characteristics play a strong role in determining rates of deforestation and therefore change in biodiversity, is borne out.

---

[15] Population is often acknowledged to be a major determinant of the rate of tropical deforestation (Rudel 1994).
[16] "LM statistics" is the standard Lagrange Multiplier test for random effects over the basic model. It analyses whether the variance of $\mu(i)$ is equal to zero - i.e. whether random effects are constant, in which case simple OLS is valid.
[17] Statistically significant at the 1% level.

Table 2. Estimates of the parabola

Variable	Result for fixed effects
ln GDP	136.1742 (41.6000)**
$(\text{ln GDP})^2$	-8.2360 (2.7706)**
Year	-2.0329 (0.1102)**
Forest area	8.685E-04 (1.877E-04)**
Population change	77.5007 (41.7039)*
Population density	-7.1946 (3.1814)*
Kuznets curve	No
F-statistics	59.57**

Table 3. Estimates of the hyperbola

Variable	Result for fixed effects
1/lnGDP	-15871.5746 (3927.6100)**
Year	-2.1683 (0.0939)**
Forest area	0.8289 (0.0002)**
Population change	73.8347 (41.7691)*
Population density	-6.1428 (3.1445)*
F-statistics	62.95**

---

[18] This is not surprising in light of the previous argument for relaxing assumptions of commonality.

Table 4. Estimates of the linear equation

Variable	Result for fixed effects
lnGDP	12.3070 (3.1343)**
Year	-2.1741 (0.0940)**
Forest area	9.042E-04 (1.846E-03)**
Population change	69.0585 (41.8140)*
Population density	-5.1684 (3.1230)*
F-statistics	62.85**

>From Tables 2, 3 and 4 it can be seen that all variables are significant at the 1% confidence level, with the exception of population change and population density, which are only significant at the 5% level. Although they are admittedly significant, suspicions were raised and both terms were subject to a post-analysis Wald test, which revealed them to be jointly insignificant. Their presence does not, however, affect the results.

The category "F-statistics" tests whether all coefficients are jointly equal to zero - i.e. whether the model as a whole is significant. The results are significant at the 1% level for each of the three equations.

Of most significance to this study, however, is the result that the environmental Kuznets curve is not observed. From the coefficient values, we interpret the fixed effects model to fit an inverted-U shape, precisely the opposite of Kuznetsian behaviour in this study. In addition, the hyperbola and the linear equation are also significant. Therefore, we can conclude the environmental Kuznets curve does not apply here.

The significance of the linear equation is certainly an artefact of the data distribution, which displays strong concentration around the dependent variable's reference value, in spite of efforts to "blow it up". Disregarding this, and given that both the hyperbola and parabola offer a significant fit, we can conclude only the "falling limb" of the curve is confirmed by the data.

VI. Conclusions

From the discussion, the following conclusions regarding methodology can be drawn:

- The possibilities of generating accurate data for testing the environmental Kuznets hypothesis are extremely limited at present. The nature of the study tends to exclude the existing cross-sectional data sets that specifically handle species numbers, thus confining work to deriving figures from the area of habitat remaining. Not only is this method rough, it is also based on forest area, which is itself subject to a high degree of uncertainty.

- Fixed effects, whereby variation in the intercept is parameterised, are theoretically and empirically favoured over random effects, whereby group specific disturbances are random and have, in principle, nothing to do with the prevailing environmental, political and economic considerations in each country.

Finally, we can answer the question posed in the title - "does an environmental Kuznets curve exist for biodiversity?"

- No. The theoretical dynamics of species diversity would not predict a Kuznetsian-type development with income and this is confirmed by empirical results, which fail to yield the desired curve for the favoured model. Nevertheless, the rival theory of hyperbolic dynamics is no more persuasive, given the parabola, hyperbola and linear equation are all significant. We can infer the "falling limb" of the curve exists but our data do not permit any curve to be identified as best.

VII. Further Work

- The nature of the species-area relationship and the characteristics of tropical moist forests are presumed to vary significantly between the three relevant continents – South America, Africa and Asia (Reid 1992). It would therefore be worthwhile to perform the same analysis for each continent separately.

- Many studies of the environmental Kuznets curve (e.g. Allen and Barnes 1985) have considered cross-sections. This is the only type of data for which direct species counts exist. Therefore, it would be revealing to compare the results of this study and such a cross-sectional analysis, although the point made regarding the assumptions inherent in cross-sectional regression is still valid.

- The most important outcome of this study, that an environmental Kuznets curve does not exist for biodiversity, is accompanied by a lack of clarity regarding the correct type of curve. One way to illuminate this issue would be to test a non-parametric function, in which the shape of the curve is plotted as the calculation is made. The results of this study would point to the real species diversity/income trend underlying the data distribution here.

VIII. References

Allen, J.C. and Barnes, D.F., "The causes of deforestation in developing countries", *Annals of the Association of American Geographers* 75(2), pp.163-184, 1985.

Connor, E.F. and McCoy, E.D., "The statistics and biology of the species-area relationship", *The American Naturalist* 113(6), pp.791-833, 1979.

Cropper, M. and Griffiths, C., "The interaction of population growth and environmental quality", *Population Economics* 84(2), pp.250-254, 1994.

Ekins, P., "The Kuznets curve for the environment and economic growth: examining the evidence", *Environment and Planning A* 29, pp.805-830, 1997.

FAO, *Forest Resources Assessment 1990: Tropical Countries*, FAO, Rome, 1990.

FAO, *Production Yearbook*, FAO, Rome, Various Years.

Greene, W.H., *Econometric Analysis*, 3rd edn., Prentice Hall, Englewood Cliffs, NJ, 1997.

Grossman G. and Krueger A., "Econometric growth and the environment", *Quarterly Journal of Economics* 60, pp.353-377, 1995.

Koop, G. and Tole, L., "Is there an environmental Kuznets curve for deforestation?" *Journal of Development Economics* 58, pp.231-244, 1999.

MacArthur, R.H. and Wilson, E.O., *The Theory of Island Biogeography*, Princeton University Press, Princeton, NJ, 1967.

Maddala, G., *Introduction to Econometrics*, 2nd edn., Macmillan, New York, NY, 1992.

Myers, N., *Conservation of Tropical Moist Forests,* The National Research Council, National Academy of Sciences, Washington, DC, 1980.

Pearce, D.W., and Brown, K., "Saving the world's tropical forests", in Pearce, D.W. and Brown, K. (eds.), *The Causes of Tropical Deforestation*, UCL Press, London, pp.2-26, 1994.

Pearce, D.W., Moran D. and Fripp, E., *The Economic Value of Conserving Biological and Cultural Diversity*, Centre for Social and Economic Research on the Global Environment, UCL Press, London, 1992.

Reid, W.V., "How many species will there be? In Whitmore, T.C. and Sayer, J.A. (eds.) *Tropical Deforestation and Species Extinction*, Chapman and Hall, London, pp55-74, 1992.

Rudel, T. "Population, development and tropical deforestation: a cross-national study" in Pearce, D.W. and Brown, K. (eds.), *The Causes of Tropical Deforestation*, UCL Press, London, pp.96-105, 1994.

Selden, T. and Song, D., "Environmental quality and development: Is there a Kuznets curve for air pollution emissions?" *Journal of Environmental Economics and Management* 27, pp.147-162, 1994.

Shafik, N. "Economic development and environmental quality: an econometric analysis", *Oxford Economic Papers* 46, pp.757-773, 1994.

Summers, R. and Heston, A., "The Penn World Table (Mark 5): An expanded set of international comparisons, 1950-1988", *Quarterly Journal of Economics* 106, pp.327-369, 1991.

Wilson, E.O., "The current state of biological diversity", in Wilson E.O. (ed.) *Biodiversity*, National Academy Press, Washington, DC, pp.3-18, 1986.