OECD Statistics Working Papers 2013/01

More than the Sum of their Parts: Valuing Environmental Quality by Combining Life Satisfaction Surveys and GIS Data

Jérôme Silva,

Zachary S. Brown

https://dx.doi.org/10.1787/5k4840hfpwkb-en

OECD

Unclassified **STD/DOC(2013)1**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **10-Apr-2013**

English - Or. English

STATISTICS DIRECTORATE

STD/DOC(2013)1
Unclassified

Cancels & replaces the same document of 22 March 2013

MORE THAN THE SUM OF THEIR PARTS: VALUING ENVIRONMENTAL QUALITY BY COMBINING LIFE SATISFACTION SURVEYS AND GIS DATA

STATISTICS DIRECTORATE

WORKING PAPER No 50

This paper has been prepared by Jérôme Silva and Zachary Brown from the OECD Environment Directorate.

For further information please contact jerome.silva@oecd.org and zachary.brown@oecd.org

JT03337885

Complete document available on OLIS in its original format
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

English - Or. English

OECD STATISTICS WORKING PAPER SERIES

The OECD Statistics Working Paper Series - managed by the OECD Statistics Directorate - is designed to make available in a timely fashion and to a wider readership selected studies prepared by OECD staff or by outside consultants working on OECD projects. The papers included are of a technical, methodological or statistical policy nature and relate to statistical work relevant to the Organisation. The Working Papers are generally available only in their original language - English or French - with a summary in the other.

Comments on the papers are welcome and should be communicated to the authors or to the OECD Statistics Directorate, 2 rue André Pascal, 75775 Paris Cedex 16, France.

The opinions expressed in these papers are the sole responsibility of the authors and do not necessarily reflect those of the OECD or of the governments of its Member countries.

www.oecd.org/std/publicationsdocuments/workingpapers/

© OECD 2013
Applications for permission to reproduce or translate all or part of this material should be made to: OECD Publications, 2 rue André-Pascal, 75775 Paris, Cedex 16, France; e-mail: rights@oecd.org

MORE THAN THE SUM OF THEIR PARTS: VALUING ENVIRONMENTAL QUALITY BY COMBINING LIFE SATISFACTION SURVEYS AND GIS DATA

Jérôme Silva and Zachary Brown
OECD
March 2013

ABSTRACT

While environmental economics studies using stated life satisfaction data have been gaining attention, much of this body of work remains exploratory. In this study we contribute to this emerging body of research by combining OECD survey data from four European countries on life satisfaction and perceptions of environmental quality with independent (i.e. mechanical) measurements of air quality and urbanity, from the European Environment Agency, to provide a broad picture of the environmental determinants of life satisfaction, and monetary valuation of air quality improvements. We also estimate that the value of a 1% reduction in air pollution (measured as mean annual PM_{10} concentrations) is worth the same on average as a 0.71% increase in per capita income. We find that environments which respondents perceive as noisy and lacking in access to green space have a significantly detrimental impact on life satisfaction. However, controlling for these negative factors (air, noise, and lack of green space), we also find a large positive residual impact of urban environments on life satisfaction. The use of independent, GIS-based measures of urbanity (proportion of urban surface area around households), as opposed to survey-based stated perceptions of urbanity, increases the precision of estimated air quality impacts on life satisfaction. Taken as a whole, our analysis highlights the need for conducting LS-based environmental assessment and valuation exercises using a broad array of independent data sources, in order both to obtain unbiased regression estimates and to facilitate interpretation of these estimates.

Keywords: life satisfaction, environmental valuation
JEL codes: C21, H23, H41, Q51, Q53, R20

RÉSUMÉ

Alors que les études sur l'économie de l'environnement qui font appel à des données sur le bien-être suscitent un intérêt grandissant, ces travaux conservent pour beaucoup d'entre eux un caractère exploratoire. Dans cette étude, nous apportons une contribution à ce domaine de recherche émergent en combinant des données issues d'enquêtes menées par l'OCDE dans quatre pays européens sur la satisfaction à l'égard de la vie et la qualité perçue de l'environnement, avec des mesures indépendantes (physiques) de la qualité de l'air et du caractère urbain provenant de l'Agence européenne pour l'environnement (EEA), dans le but de dresser un tableau général des déterminants environnementaux de la satisfaction à l'égard de la vie et de produire une évaluation monétaire des améliorations de la qualité de l'air. Nous estimons également qu'une réduction de 1 % de la pollution de l'air (mesurée comme la concentration annuelle moyenne de PM_{10}) a la même valeur en moyenne qu'une hausse de 0.71 % du revenu par habitant. Nous constatons que les milieux perçus par les répondants comme bruyants et manquant de possibilités d'accès à des espaces verts ont un effet négatif sensible sur le bien-être. Cependant, si nous neutralisons l'effet de ces facteurs négatifs (air, bruit et manque d'espaces verts), nous observons aussi un fort impact résiduel positif des milieux urbains sur la satisfaction à l'égard de la vie. Le fait de recourir à des systèmes d'information géographique pour obtenir des mesures indépendantes du caractère urbain (en l'occurrence, la proportion de surfaces urbanisées autour du foyer), au lieu de s'en remettre aux appréciations sur ce point des répondants aux enquêtes, permet des estimations plus précises de l'impact de la qualité de l'air sur la satisfaction à l'égard de la vie. Dans l'ensemble, notre analyse met en lumière la nécessité de faire appel à un large éventail de sources de données indépendantes pour conduire des évaluations environnementales fondées sur la satisfaction à l'égard de la vie, afin d'obtenir des estimations par régression sans biais et de faciliter l'interprétation de ces estimations.

TABLE OF CONTENTS

INTRODUCTION 6
DATA SOURCES AND SUMMARY STATISTICS 8
Household survey data 8
Airbase 9
Corine Land Cover 10
Merging the datasets 10
DEFINITION OF VARIABLES AND SUMMARY STATISTICS 12
MODEL AND ANALYSIS 19
Estimating the marginal value of air pollution reductions 20
CONCLUSIONS 21
REFERENCES 22

Tables

Table 1: Geocoding respondent postal codes. 9
Table 2: Definitions of the variables 14
Table 3: Summary statistics for the sample 15
Table 4: Life satisfaction regression results 17
Table 5: Estimated elasticity between per capita income and PM10 18

Figures

Figure 1: Life satisfaction elicitation instrument 9
Figure 2: OECD Household Survey Respondent locations and EEA air stations locations across Europe10
Figure 3: Illustration of Corine Land Cover Classification 11
Figure 4: Histograms of household PM_10 exposure by country 13
Figure 5: Mean 2009 PM_10 by stated satisfaction with local air quality 13
Figure 6: Proportion of urban areas within 5km of household, by households' subjective 16
Figure 7: Correlation between urban surface area and air pollution 16

INTRODUCTION

1. The use of life satisfaction (LS) data for valuing environmental goods and the costs of pollution is now an established empirical method in environmental economics, but analysis using this method has yet to directly inform policy, e.g in benefit-cost analysis (BCA). The theory underlying the use of LS measurements for valuation is simple and robust, the basic idea being to compare the partial correlations between LS, environmental quality, and income to obtain a marginal value of environmental quality in monetary terms. By using observable – rather than hypothetical – levels of environmental quality and pollution, while at the same time avoiding assumptions about perfect market efficiency, LS valuation approaches can complement existing stated preference (SP) and revealed preference (RP) valuation methods (Welsch, 2009). Furthermore, research into the validity of self-reported LS measurements (e.g. elicited from surveys) has established that – when appropriately collected – these data can provide a valid psychological measure, i.e. one which is stable, responsive to changes in life circumstances, and corresponds well to observable signs of individuals' happiness (Diener, 2006; Diener, 1999; OECD, 2013).

2. While LS-based valuation studies have been gaining attention, the body of empirical work in this area is still exploratory, and it remains unknown how robust the resulting valuation estimates are. Early econometric approaches to conduct valuation with LS data used aggregate data on average nation-level happiness, environmental quality, and income (Welsch, 2002; Welsch, 2007; Welsch, 2008). Those studies using individual-level data either focus on relatively small geographic areas (MacKerron, 2009; Van Praag, 2005), do not employ independent measurements of environmental quality (Rehdanz, 2008), or use aggregate measures of environmental quality (Frey, 2009; Israel, 2003; Silva, 2012) – e.g. average air pollution across an entire city, national-level water pollution indicators. The reasons for this apparent data tradeoff are understandable: It is rarely the case that, for a large geographic area, we have records of surveyed households' locations which could be matched to independent, micro-level indicators of environmental quality. The present study is in the advantageous position of having such data.

3. But why exactly is there a role for analyses that use such detailed information? One reason is that using aggregated data can result in imprecise estimates, by excluding systematic individual-level variation. This imprecision, while not necessarily a problem for the validity of LS data, can in practice hinder the use of these data for policy: If various aspects of environmental quality are not found to have any statistically significant effect on happiness – or to the contrary found to have an implausibly large effect – then policymakers may view LS data as 'strange,' and intractable for policy analysis, even if the statistical imprecision itself does not imply the absence or implausibility of an effect. Furthermore, existing analyses using aggregate LS data assume that survey responses are cardinal rather than ordinal, i.e. the difference between a happiness rating of ten and nine means the same as the difference between a rating of one and zero {see discussion by Welsch, 2009).

4. In contrast, using individual-data from a limited geographic area or with only survey self-reports poses its own set of econometric problems. One problem is external validity: Without expansion of geographic scope or study replication, we do not know if, for example, an LS estimate of the marginal value of nitrogen oxide reductions in London is relevant for other regions. Furthermore, by often focusing only on cities (which tend to have lower air quality than rural areas), 'objective' measurements of air quality may exhibit less variation – and hence reduce statistical power – compared to data that include a range of landcover types.

5. On the other hand, using fully disaggregated data that come entirely from one survey instrument can both bias the estimates and make interpretation of results difficult. Bias can arise because measurement error associated with the environmental quality variable is more likely to be correlated with the life satisfaction data when both sets of measurements come from the same data source; such correlated measurement error can cause attenuation bias in regression analysis (e.g. Duncan, 1985). Furthermore, because surveys normally elicit only perceptions of environmental quality (e.g. "is your air good or bad?"), it can be difficult to interpret the results based only on these data, since it is not clear how these perceptions relate to 'reality.'

6. The ideal LS-based valuation study would therefore fulfill the following conditions: Such a study would use individual-level, geocoded data involving variegated geographic areas and external measurements of environmental quaulity, for reasons discussed above. It would include a rich set of control variables, to account for the likely correlation between environmental quality and many other factors, including urban/rural aspects, household wealth, health status, surrounding noise levels, access to green space, etc. An ideal study would also observe the same individuals at multiple points in time (i.e. the study would use panel data).

7. The present paper contributes to building a more robust LS-based environmental valuation literature by being the first (as far as we are aware) to use an individual-level dataset combining representative samples of households for four countries (France, the Netherlands, Spain, and Sweden). While the present data are not a panel, they do provide a unique combination of measurements. The data permit spatial scope and precision in the econometric analysis, and at the same time allow us to control for a wide range of household-specific and location-specific characteristics. The LS data, income measurements, and a large set of self-reported control variables come from a websurvey administered to approximately 1,000 households in each of the countries. The measurements for air pollution – the aspect of environmental quality which we focus on here – come from GIS databases maintained by the European Environment Agency (EEA). To control for potential effects of built environments on LS, we also utilize GIS-based landcover data provided by the EEA.

8. We find that local air pollution unambiguously decreases self-reported LS. In our preferred econometric model, the impact of decreasing average annual particulate matter concentrations by 1% is equivalent to increasing per capita income by 0.71%. We also identify significant, independent LS effects of noise and access to green space. However, controlling for these negative factors (air, noise, and lack of green space), we also find a large positive residual impact of urban environments on life satisfaction. More broadly, our analysis highlights the necessity of conducting LS-based valuation exercises using a broad array of data sources, both to obtain unbiased regression estimates and to facilitate interpretation of these estimates. In our regression analysis, the clearest picture emerges from LS regression models when independently assessed air quality measures are combined with satellite-based indicators of urbanity *and* with subjectively assessed noise levels and access to green space. Including only a subset of these variables in the regression – or using subjective, surveyed-based measures of environmental quality in place of a mechanically measured variable, when the latter is available – results in estimates that are less precise and which are more difficult to interpret for economic analysis. In particular, the use of independent, GIS-based measures of urbanity (proportion of urban surface area around households), as opposed to survey-based stated perceptions of urbanity, increases the precision of estimated air quality impacts on life satisfaction.

DATA SOURCES AND SUMMARY STATISTICS

9. This analysis combines three, merged sets of data from different sources. The life satisfaction data come from a 2011 wave of an OECD household survey on environmental attitudes and behaviours. The air quality data come from the AirBase database created by the European Environmental Agency (EEA). The landcover data come from the Corine Land Cover (CLC) database, also created and maintained by the EEA.

Household survey data

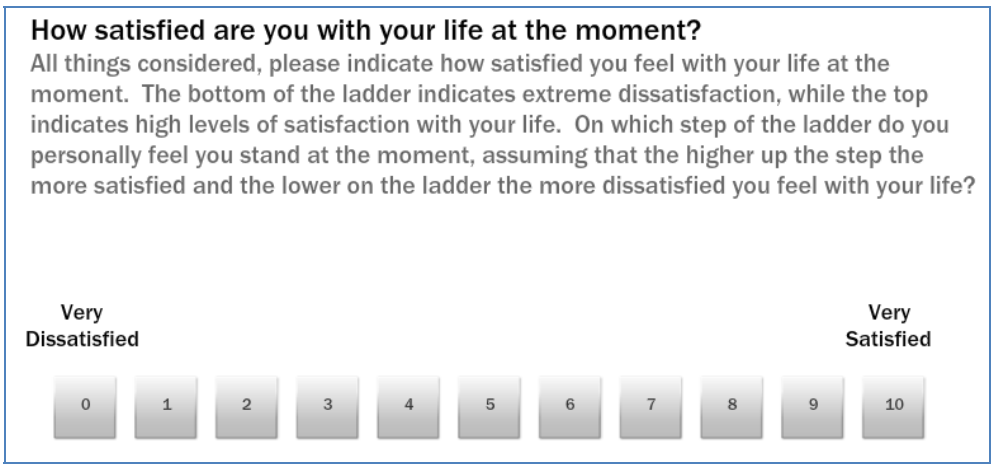
10. The life satisfaction data and a range of control variables come from a 2011 web-based survey of 11 OECD countries, five of which—France, Netherlands, Spain, Sweden, and Switzerland—are in the European Union and therefore are covered by the EEA databases. The household questionnaire for this survey was developed by OECD, with input from government officials in the participating countries. A pilot survey was first conducted with 500 respondents in three of the countries, and the questionnaire was revised based on results from the pilot. The finalised questionnaire was then translated into Dutch, French, German, Hebrew, Korean, Japanese, Spanish and Swedish. Translations complied with the International Standard ISO 20252: All translations were carried out and double-checked by native speakers.[1]

11. For each country, nationally respresentative groups of approximately 1,000 households were sampled. Global Market Insight (GMI), a web survey firm, conducted the survey. To obtain nationally representative samples, GMI maintains panels of respondents who receive periodic requests to participate in various web-based surveys. GMI adheres to the ESOMAR 26 guidelines for web-based panel surveys (see Callegaro, 2008). Respondents from GMI's panels were invited to participate in the OECD survey using a quota sampling methodology. Using national-level statistical and census data, quotas targets were set for gender, age, household income quintile, and geographic distribution within each country. The completion rate for the survey was greater than 75% in each of the countries, and all quotas were achieved to within 20% of their targets, with the exception of the highest income quintiles for some countries.

12. Potential respondents who started the questionnaire were asked whether they met the screening criteria presented above (living in non-institutional settings and influential in household financial decisions). If they did not meet the criteria, they were thanked for their time and screened out of the sample. The overall drop-out rate was 21%. This varied from 13% in Korea to 35% in Chile. Median response times by country varied from 28 minutes (Korea) to 41 minutes (Chile), with the medians in all countries being less than 35 minutes except Chile. In cleaning the data, it was observed that a small fraction of respondents were "speeders"—i.e. these individuals progressed through the survey at such a rapid pace that the validity of their responses is suspect. To address this issue, respondents who completed the questionnaire in less than half of the median response time for their country were removed from the dataset. Readers who are interested in reviewing the English version of the survey instrument can find it at http://qsurvey.gmisurveys.com/dc/index.html?p=PWuYTGX. Data on life satisfaction were elicited using the web-based instrument depicted in **Figure 1**.

[1] See www.oecd.org/environment/households for details on implementation and sampling.

Figure 1: Life satisfaction elicitation instrument



Geocoding the survey data

13. Crucially, data are available on the households' locations down to the level of postal code. These postal codes were manually geocoded into latitude and longitude using the web-based GPS Visualizer tool (Schneider, 2012). This procedure permitted the geocoding of over 85% of respondents in 4 of 5 countries, but was much less successful for Switzerland due to the low fraction of respondents in that country who reported their postal code (**Table 1**). We therefore exclude the Swiss sample from this analysis, and focus only on the French, Dutch, Spanish, and Swedish samples. A map of households' geocoded coordinates is displayed in **Figure 2**.

Table 1: Geocoding respondent postal codes.

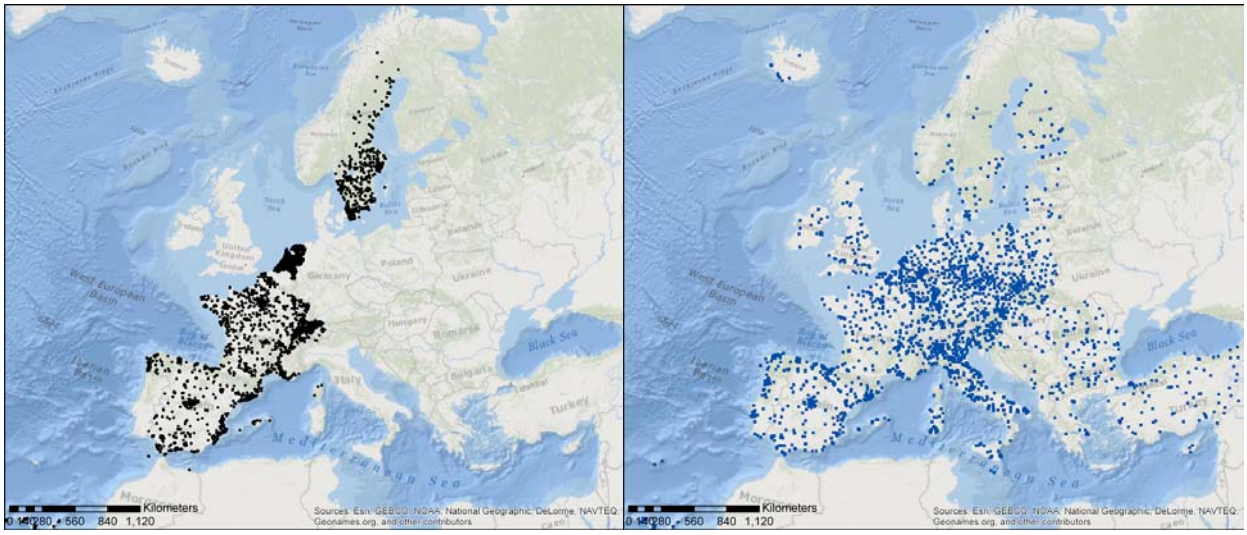
	France	Netherlands	Spain	Sweden	Switzerland
Total respondents	1,227	1,301	1,101	1,012	1,089
Postal codes available	1,222	1,295	1,091	888	372
Geocoded respondents	1,170	1,286	1,089	882	372
Percent geocoded	95%	99%	99%	87%	34%
Within 20k of monitoring station	75%	64%	84%	45%	--

Airbase

14. The air quality measurements come from AirBase. AirBase is an open-access air quality database system created by the European Environment Agency (EEA). The air quality database consists of multi-annual time series of air quality measurement data and their statistics for a representative selection of monitoring stations and for a number of pollutants. It also contains meta-information on the monitoring networks, their stations and their measurements. AirBase includes measurement data for a wide array of air pollutants variables, including PM_{10}, $PM_{2.5}$, NO_X, SO_2, CO, as well as lead. We focus here on the PM_{10} data, because particulate matter is widely used as a measure of air pollution in both environmental health and environmental economics research: There is robust evidence of its negative health impacts, and is regarded as a type of pollution that can be easily noticed by households (e.g. Deguen, 2008).

Figure 2: OECD Household Survey Respondent locations (for European Countries only, black) and EEA air stations locations across Europe (blue)

Household locations derived from respondent postal codes



Corine Land Cover

15. Indicators of urbanity were obtained from the corine land cover (CLC) database maintained by the EEA. This database that provides consistent information on land cover across Europe using a nomenclature of 44 standard classes, with 3 levels of detail. The satellite images are processed to form a map composed of 100m x 100m contiguous cells, which are each coded using the 44 classes. For the purpose of this study, the 44 classes were aggregated into two categories:

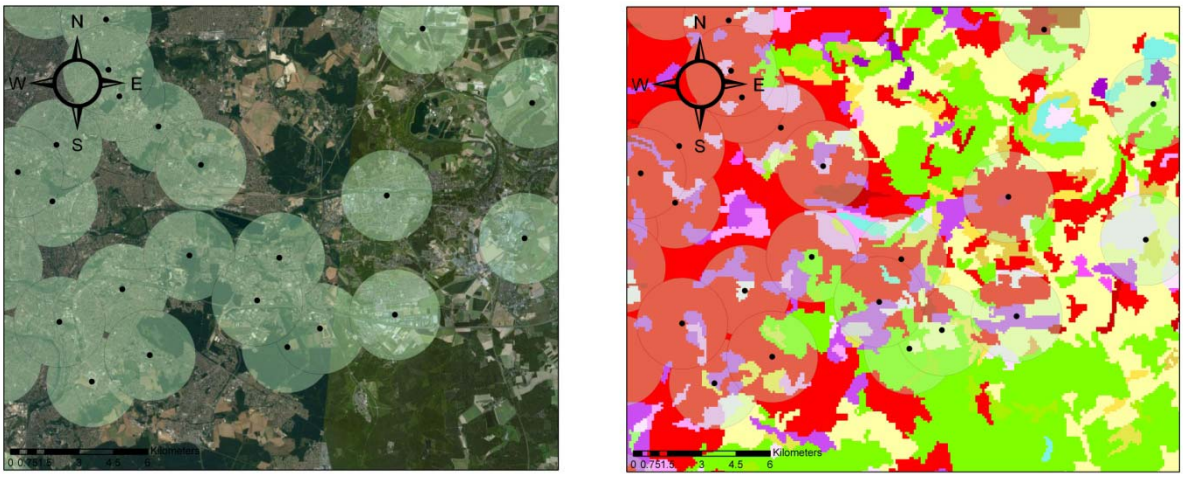
- *Urban areas* include all artificial surfaces excluding artificial, non agricultural vegetated areas (i.e. parks).

- *Other areas:* This includes (a) "green" areas (i.e. agricultural areas, forest and semi-natural areas, and artificial non agricultural areas), and (b) water areas, corresponding to wetlands and water bodies (inland and marine waters).

Merging the datasets

16. Using the estimated latitudes and longitudes, the household survey data were matched with the locations of each household's five nearest air pollution monitoring stations in AirBase and with the cells in CLC database. The matching between households with their nearest stations was done by computing the distances between the household and monitoring station coordinates via the haversine formula (Renfree, 2010). Those households estimated to be more than 20km from the nearest air monitoring station are excluded from the analysis (see **Table 1**, last row). In total 3,012 observations remain after executing these merging and cleaning procedures, and this is the sample we focus on for the rest of the analysis.

Figure 3: Illustration of Corine Land Cover Classification (detail of the East of Ile-de-France Region)



Left panel: Satellite image. Right Panel: Corine Land Cover classification, red and violet areas represent constructed surfaces while green and yellow areas represent heterogeneous agricultural areas. Circles show 5km radius around households.

DEFINITION OF VARIABLES AND SUMMARY STATISTICS

17. **Figure 4** shows country-level histograms of mean annual household PM_{10} exposure. The WHO guidelines specify a maximum for average annual PM_{10} concentration of 20 µg/m3. Our data imply that this limit is exceeded for 85% of the respondents. It is not an uncommon to find large populations, even in relatively high-income areas, with annual PM_{10} concentrations that far exceed recommended thresholds (Pandey, 2006). In the Netherlands the variance in the distribution of PM_{10} is much lower than the other countries. This is likely because of the relatively small and uniform land area over which these data are spread; the distribution of PM_{10} concentrations as measured by air monitoring stations is likely to vary less than in the other countries.

18. **Figure 5** shows how households' estimated PM_{10} exposure varies by survey respondents' stated level of dis/satisfaction with their local air quality. The correlation between PM_{10} and subjective perceptions of air quality is clear, though there are some non-linear abnormalities in Spain and Sweden.

19. **Figure 6** shows the proportion of land around the household classified as urban according to the CLC data, plotted by respondents' stated perceptions of their local surroundings. In particular, the figure shows that the respondents' perceived level of urbanity matches up fairly well with our preferred CLC measure of urban area (URBANPROP50). However, there are some interesting differences between countries, particularly with regard to what respondents classify as a 'surburb' or a 'small village/town.' What Dutch respondents classify as a village/town – in terms of GIS-based urban area – is more likely to be classified as a city/suburb by respondents in the other countries. **Figure 7** provides a scatterplot of the air pollution and urbanity data, showing the correlations between these variables in each of the four countries analysed.

20. **Table 2** provides a description of all variables used in the analysis, and **Table 3** provides summary statistics for the sample. Each country's sample is roughly representative of its population (a systematic comparison between the sample and estimated population statistics can be found by following the web-link in footnote 1). The summary statistics we focus on concern the dependent and independent variables of interest, i.e. LS, perceptions of environmental quality (EQ) and local surrounding, and the AirBase and CLC data.

21. The LS data show the characteristic smooth, unimodal and skewed distribution seen in other studies using such data (e.g. MacKerron, 2009). In France and Sweden the mode occurs at LS=6, while in Spain and the Netherlands the mode is located at LS=7. Assuming there is no systematic difference in the interpretation or response style with the LS scale used in the questionnaire, the Dutch appear to clearly be the 'happiest' respondents in our sample. Looking at the perceptions of local EQ, we see that respondents are most frequently dissatisfied with their local air quality and noise levels, and that access to green space is less of a problem in general.

Figure 4: Histograms of household PM_10 exposure by country

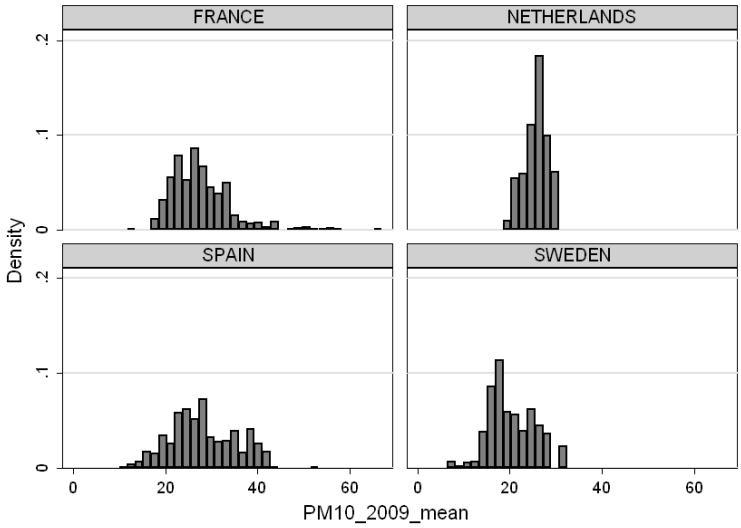


Figure 5: Mean 2009 PM_{10} by stated satisfaction with local air quality

The data on stated satisfaction with local air quality in this figure are a disaggregated version of the AIRBAD variable in Table 2

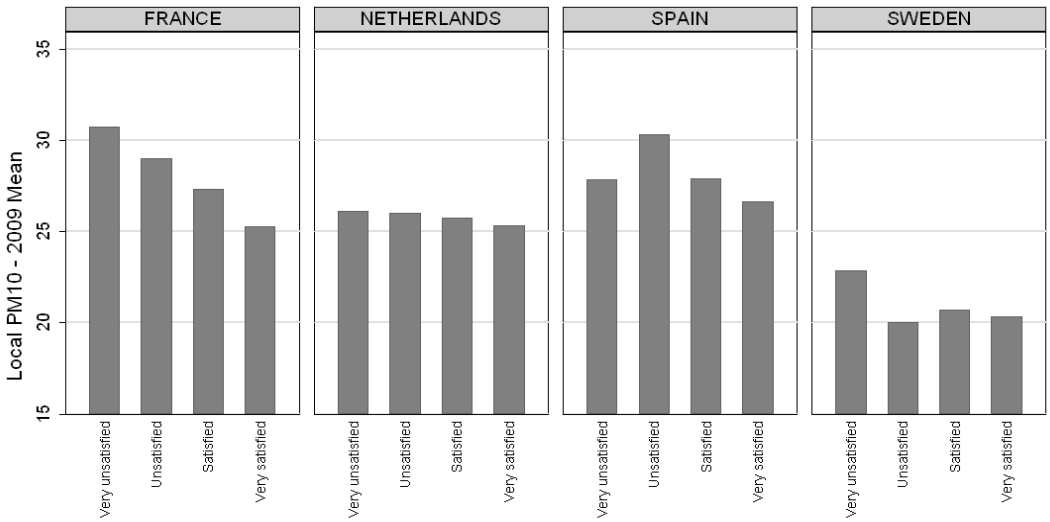


Table 2: Definitions of the variables

VARIABLES	*Description*
AGE	Age of respondent, in years.
FEMALE	Binary variable taking a value of 1 if respondent is female, and 0 otherwise.
MULTI_OCCUP	Binary variable taking a value of 1 if respondent lives in a household with more than one person, and 0 otherwise.
MARRIED	Binary variable taking a value of 1 if respondent is married, and 0 otherwise
YNG_CHILD	Binary variable taking a value of 1 if young children are present in the household, 0 otherwise.
INCOME	Per capita annual income, in 2007 euros.
POST2ND_EDUC	Number of years of post-secondary education that the respondent completed.
UNEMPLOYED	Binary variable taking a value of 1 if respondent is unemployed and 0 otherwise.
DISABLED	Binary variable taking a value of 1 if respondent is disabled, and 0 otherwise
MED_CNDTNS	Number of medical conditions stated to be present in the household, divided by the number of household members.[1]
HOME_RENT	Binary variable taking a value of 1 if the household rents their home, and 0 otherwise.
HOME_SIZE	Size of residence, in square meters (m^2).
URBANPROP50	*From Corine Land Cover database*. Estimated proportion of land within a 5km radius of household's location that is classified as 'Urban.'
PM10MEAN2009	*From AirBase*. Mean annual PM_{10} (micrograms per cubic meter, $\mu g/m^3$), distance-weighted average from five closest air monitoring stations.[2]
LIFE_SATISFACTION	Response to the question "How satisfied are you with your life at the moent?" using an 11 point scale from zero (Very Dissatisfied) to ten (Very Satisfied).
NOPARKS	Binary variable taking a value of 1 if respondent states "Very dissatisfied" or "Dissatisfied" (on a 5 point scale) in response to the question "How satisfied are you with your access to green spaces (e.g. parks, forests)?"
AIRBAD	Binary variable taking a value of 1 if respondent states "Very dissatisfied" or "Dissatisfied" (on a 5 point scale) in response to the question "How satisfied are you with your local air quality?"
NOISY	Binary variable taking a value of 1 if respondent states "Very dissatisfied" or "Dissatisfied" (on a 5 point scale) in response to the question "How satisfied are you with your local noise levels?"

[1] This information was maintained by the survey firm for all potential respondents in their webpanels; it consists of a binary record indicating presence/absence *in the household* for 83 different medical conditions, including everything from allergies to brain cancer. Because larger household size increases the likelihood of having a given condition, the total number of conditions was divided by household size.

[2] The formula is $\bar{P}_i = \sum_{h=1}^{5} e^{-d_h^i} P_h^i / \sum_{h=1}^{5} e^{-d_h^i}$, where $\bar{P}_i$ is the mean PM_{10} for household i, d_h^i is the distance (in km) between household i and the hth nearest station, and P_h^i is the mean PM_{10} reading for the hth nearest station.

Table 3: Summary statistics for the sample

	-Pooled Sample-	*France*	*Netherlands*	*Spain*	*Sweden*
VARIABLES	*mean (std. dev.)*	*mean (std. dev.)*	*mean (std. dev.)*	*mean (std. dev.)*	*mean (std. dev.)*
Observations	3,012	880	821	914	397
Percent of pooled sample	100%	29%	27%	30%	13%
AGE (years)	44	43	46	42	43
	(14)	(14)	(13)	(13)	(14)
FEMALE	49%	51%	45%	50%	51%
MULTI_OCCUP	86%	87%	84%	92%	76%
MARRIED	57%	56%	61%	56%	51%
YNG_CHILD	15%	14%	12%	17%	17%
INCOME (2007 euros)	€ 16,652	€ 17,511	€ 18,500	€ 11,977	€ 21,690
	(€ 9,699)	(€ 9,858)	(€ 9,759)	(€ 7,168)	(€ 9,965)
POST2ND_EDUC	3.2	2.7	3.9	3.4	2.7
	(2.4)	(2.2)	(2.5)	(2.5)	(2.3)
UNEMPLOYED	9%	5%	4%	15%	11%
DISABLED	4%	3%	8%	2%	3%
MED_CNDTNS	1.7	1.6	1.4	2.2	1.8
(# per person)	(2.1)	(1.8)	(1.7)	(2.4)	(2.2)
HOME_RENT	32%	40%	30%	20%	42%
HOME_SIZE (m^2)	107	98	125	106	96
	(53)	(43)	(63)	(52)	(42)
URBANPROP50	0.364	0.422	0.325	0.347	0.356
	(0.247)	(0.291)	(0.189)	(0.247)	(0.218)
PM10MEAN2009 ($\mu g/m^3$)	26.34	27.58	25.67	28.18	20.72
	(5.380)	(4.742)	(1.902)	(5.887)	(6.242)
LIFE_SATISFACTION					
0 (Very dissatisfied)	1%	1%	0%	1%	3%
1	2%	3%	1%	2%	4%
2	4%	5%	1%	4%	6%
3	4%	6%	3%	3%	6%
4	11%	14%	6%	11%	11%
5	15%	18%	15%	14%	15%
6	26%	27%	30%	23%	22%
7	24%	17%	34%	24%	21%
8	7%	4%	8%	8%	8%
10 (Very satisfied)	5%	3%	3%	8%	5%
NOPARKS	17%	16%	10%	29%	9%
AIRBAD	29%	35%	20%	36%	20%
NOISY	29%	34%	18%	38%	23%

Figure 6: Proportion of urban areas within 5km of household, by households' subjective, stated type of neighbourhood

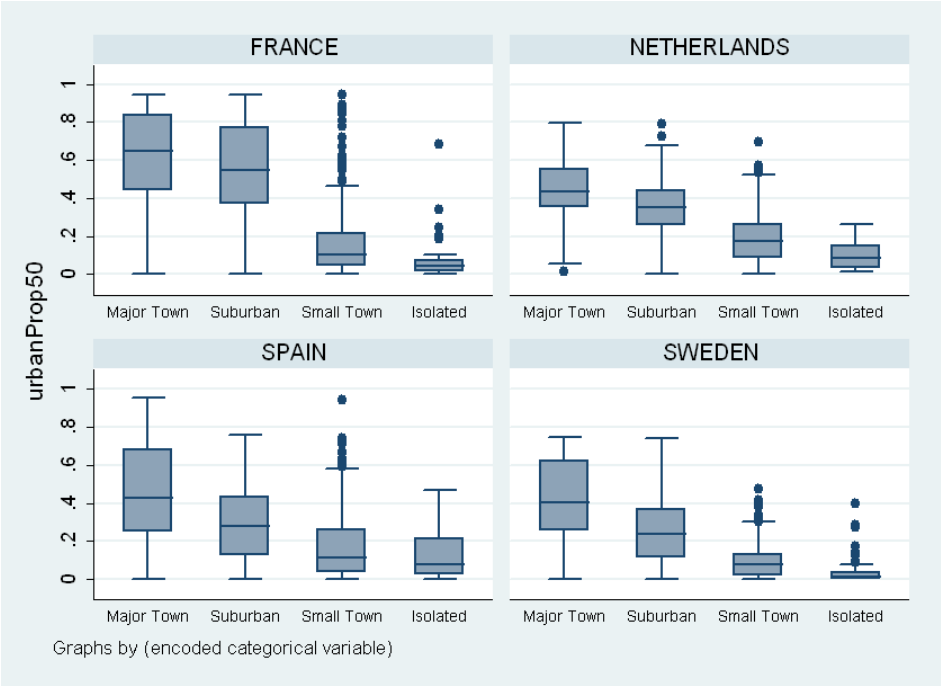


Figure 7: Correlation between urban surface area and air pollution
The correlations corresponding to the above plots are 0.4 in France, 0.3 in the Netherlands, 0.3 in Spain and 0.2 in Sweden

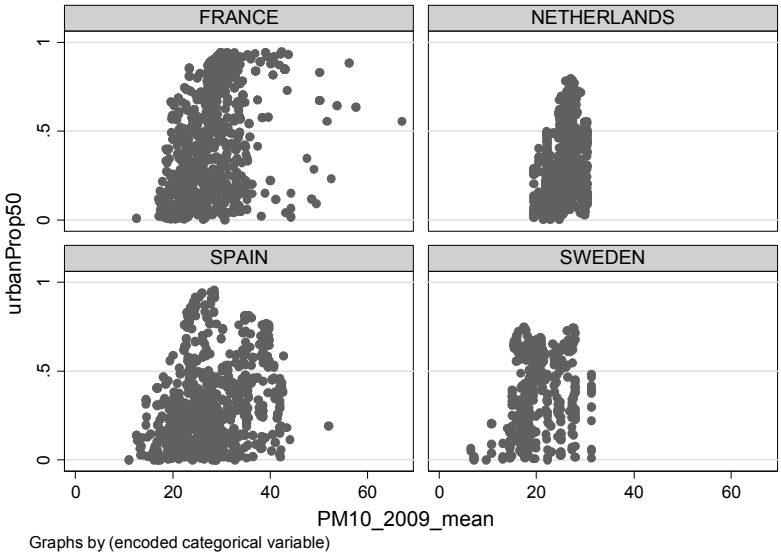


Table 4: Life satisfaction regression results

VARIABLES	Ordered probit				OLS
	Model 0	Model 1	Model 2	Model 3	
FEMALE	0.0484	0.0437	0.0370	0.0310	0.0567
	(0.0381)	(0.0380)	(0.0379)	(0.0380)	(0.0658)
AGE	-0.0524***	-0.0520***	-0.0537***	-0.0539***	-0.0916***
	(0.00991)	(0.00994)	(0.00993)	(0.00988)	(0.0171)
AGE^2	0.000588***	0.000584***	0.000599***	0.000599***	0.00100***
	(0.000111)	(0.000111)	(0.000111)	(0.000111)	(0.000190)
MULTI_OCCUP	0.287***	0.287***	0.266***	0.262***	0.458***
	(0.0700)	(0.0697)	(0.0693)	(0.0698)	(0.123)
MARRIED	0.126**	0.124**	0.128***	0.136***	0.232***
	(0.0498)	(0.0497)	(0.0497)	(0.0495)	(0.0863)
YNG_CHILD	0.0784	0.0799	0.0835	0.0853	0.118
	(0.0609)	(0.0608)	(0.0607)	(0.0604)	(0.104)
POST2ND_EDUC (years)	0.00445	0.00328	0.00253	0.000973	0.00758
	(0.00672)	(0.00670)	(0.00673)	(0.00673)	(0.0116)
UNEMPLOYED	-0.358***	-0.361***	-0.360***	-0.373***	-0.745***
	(0.0816)	(0.0816)	(0.0814)	(0.0807)	(0.153)
DISABLED	-0.205**	-0.210**	-0.211**	-0.223**	-0.384**
	(0.103)	(0.102)	(0.102)	(0.102)	(0.194)
MED_CDTNS	-0.0592***	-0.0597***	-0.0591***	-0.0635***	-0.113***
	(0.0106)	(0.0106)	(0.0106)	(0.0107)	(0.0195)
HOME_RENT	-0.221**	-0.233**	-0.234**	-0.249**	-0.515***
	(0.105)	(0.105)	(0.105)	(0.104)	(0.183)
(HOME_RENT=0) × HOME_SIZE (m^2)	0.00179***	0.00178***	0.00181***	0.00195***	0.00249***
	(0.000442)	(0.000441)	(0.000442)	(0.000438)	(0.000692)
(HOME_RENT=1) × HOME_SIZE (m^2)	0.00265***	0.00272***	0.00281***	0.00309***	0.00504***
	(0.00100)	(0.000998)	(0.00100)	(0.000986)	(0.00170)
Log(INCOME)	0.391***	0.391***	0.382***	0.393***	0.674***
	(0.0435)	(0.0434)	(0.0434)	(0.0436)	(0.0752)
CITY	0.104**	0.103**			
	(0.0477)	(0.0478)			
SUBURB	-0.103**	-0.0957*			
	(0.0502)	(0.0505)			
AIRBAD	-0.0918*				
	(0.0483)				
NOISY	-0.214***	-0.240***	-0.243***		-0.420***
	(0.0484)	(0.0460)	(0.0458)		(0.0810)
NOPARKS	-0.138**	-0.155***	-0.151***		-0.279***
	(0.0578)	(0.0569)	(0.0565)		(0.100)
Log(PM10MEAN2009)		-0.210*	-0.270**	-0.300***	-0.492**
		(0.110)	(0.113)	(0.113)	(0.198)
URBANPROP50			0.212**	0.101	0.369**
			(0.0854)	(0.0845)	(0.149)

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4, continued: Life satisfaction regression results

VARIABLES	Ordered probit				OLS
	Model 0	Model 1	Model 2	Model 3	
France (reference point)	--	--	--	--	--
NETHERLANDS	0.355***	0.352***	0.382***	0.410***	0.681***
	(0.0491)	(0.0492)	(0.0494)	(0.0491)	(0.0837)
SPAIN	0.542***	0.549***	0.598***	0.566***	0.983***
	(0.0568)	(0.0567)	(0.0556)	(0.0552)	(0.0954)
SWEDEN	0.0597	0.00276	0.0120	0.0328	-0.0638
	(0.0684)	(0.0784)	(0.0767)	(0.0767)	(0.139)
Constant					2.821***
					(1.053)
Observations	3,012	3,012	3,012	3,012	3,012
Parameters	35	35	34	32	21
R-squared					0.174
Pseudo R-squared	0.0460	0.0460	0.0453	0.0412	
Log-likelihood	-5702	-5702	-5706	-5731	

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Estimated elasticity between per capita income and PM_{10}

Model:	Ordered Probit			OLS
	Model 1	Model 2	Model 3	
Point estimate	0.54%	0.71%	0.76%	0.73%
95% confidence interval	(-0.03% - 1.10%)	(0.10% - 1.30%)	(0.17% - 1.40%)	(0.13% - 1.33%)

Point estimates and confidence intervals obtained from applying the Delta method to the ratio of corresponding regression coefficients for Log(INCOME) and Log(PM10MEAN2009) in **Table 4**.

MODEL AND ANALYSIS

22. The life satisfaction data are analysed by treating life satisfaction as a latent variable which allegedly depends on INCOME, air quality (measured by PM10MEAN2009), local noise levels (NOISY), access to green space (NOPARKS), other factors related to urbanity (captured by URBANPROP50), on a set of control variables $\boldsymbol{X}$ and a random residual ε:

$$l = \beta_0 + \beta_{INC} \log \mathrm{INCOME} + \beta_1 \log \mathrm{PM10MEAN2009} + \beta_2 \mathrm{NOISY} + \beta_3 \mathrm{NOPARKS} + \beta_4 \mathrm{URBANPROP50} + \boldsymbol{\beta_X X} + \varepsilon \quad (1)$$

where the β's are the quantitative effects of these determinants on life satisfaction.

23. Following usual procedure, we treat the observed life satisfaction survey data as an ordinal indicator of the latent variable l, and thus apply usual ordinal probit statistical models in the analysis. We estimate four different ordered probit regression models of LS (on the same set of observations), as well as an OLS model for comparison. In Model 0 only survey data are included (i.e. this regression ignores the EEA data on air pollution and land cover). In Model 1 we introduce the EEA PM_{10} measure of air pollution in place of the subjective AIRBAD variable from the survey. In Model 2 we replace the subjective perceptions of urbanity (the CITY and SUBURB dummy variables) with the satellite-based EEA data, i.e. the URBANPROP50 variable. Model 3 is a robustness check, excluding the additional attitudinal variables NOISY and NOPARKS from the regression. The OLS model uses the same covariates as Model 2.

24. **Table 4** presents the estimation results. Before turning to the air quality and land cover estimates, we examine the validity of the regression results in terms of how well they compare to other LS studies: In line with previous studies (e.g. MacKerron, 2009; Silva, 2012; OECD, 2012), we find that respondent age has a nonlinear relationship with LS, decreasing in early adulthood until the mid 40s, and then increasing. Being married has a significantly positive effect on LS, whereas unemployment and being disabled has a significantly negative effect. Our measure of health status – the average number of health conditions per household member (MED_CDTNS) – has a significantly negative effect on LS. Other findings from these regressions are that respondents in multi-person households are happier, that homeowners are generally happier than renters, and that happier people tend to live in larger residences – an effect which is enhanced by home-ownership. It should be noted that the effects relating to home-ownership and residence size could be related to household wealth, for which we have no other proxy aside from income.

25. Turning to the key variables of interest – and how their estimated effects vary across models – we can draw a couple of initial conclusions: First, the key variables necessary to perform the valuation exercise (air pollution and income) are significant and of the expected sign. Second, the various indicators of urbanity clearly are capturing some aspect of life satisfaction, but how and to what extent is a topic we will return to below. Third, the other (subjective) measures of environmental quality – perceived noise levels and access to green space – are significant and of the expected sign in all the models in which they appear. Finally, when we include the GIS-based measure of urbanity in place of the subjective measure (Model 2 compared to Model 1), the precision of the estimated LS effect of PM_{10} increases: The p-value decreases from 0.056 to 0.017.

26. The fact that different measures of urbanity appear to increase the statistical precision of other estimates should lead us to think harder about what aspects of life satisfaction these variables are capturing: When using the subjective measures, living in a city (relative to a small village or isolated dwelling) evidently has a positive net effect on LS, controlling for other factors including noise and air quality. But living in a suburb evidently incurs a LS penalty. It is difficult, however, to read too much into these effects because different people may classify the same place as being a city or a suburb (e.g. **Figure 6**) in a way that it is linked to their perception of how that place relates to their life satisfaction.

27. When we include a GIS-based measure of urbanity in place of the survey-based measure, a more urban surrounding appears to be associated with greater LS, again as long as we control for other factors such as noise and lack of access to green space (Model 2 in **Table 4**). It is clearly important to control for as many *specific* factors in urban environments as possible: When we drop the subjective NOISEBAD and NOPARKS variables (Model 3 in **Table 4**), then our URBANPROP50 variable no longer has any significant link with LS. It is likely that the 'good' and 'bad' aspects of urbanity cancel each other out here, since neither one of them are independently controlled in this model. In this way, measures of urbanity may really be viewed as a type of fixed effect in LS models, in that they absorb many aspects implicated in LS but which we may not observe (e.g. better public services, employment opportunities and access to markets and cultural centres, but joined with crowdedness and a higher cost of living). Having general measures of urbanity can therefore be very useful when controlling for many of these aspects, for example when attempting to estimate the specific impacts of air quality on LS.

Estimating the marginal value of air pollution reductions

28. Economic analysis and valuation using LS data treats these data as proxy measurements for indirect utility V (Welsch, 2009). Standard theory in environmental economics, in turn, supposes that indirect utility of an individual is a function F such that:

$$V = F(\boldsymbol{Q}, M, \boldsymbol{X}, v) \tag{2}$$

where $\boldsymbol{Q}$ is a vector of attributes of environmental quality (including air quality), M is income, $\boldsymbol{X}$ is a vector of other relevant and observed factors which determine indirect utility, and v captures the unobserved factors determining V.

29. A marginal value of environmental quality improvements can then be obtained by estimating the slope of the indifference curves between a component q of the vector $\boldsymbol{Q}$ and money M:

$$MV_q = -\frac{\partial V / \partial q}{\partial V / \partial M} \tag{3}$$

Note that this marginal value only includes the direct effect of q on indirect utility (e.g. it excludes potential effects of environmental quality on income and on health). The elasticity between environmental quality and income along a utility indifference curve is therefore $\epsilon_q \equiv MV_q \cdot (q/M)$. If we take indirect utility V to be equivalent to life satisfaction l in the regression equation at the beginning of the section, then this elasticity is constant and is simply the ratio of the coefficients, i.e. $\epsilon_Q = \beta_1/\beta_{INC}$.

30. Computing the ratio of coefficients between PM_{10} and income from Table 4, we therefore find that a 1% decrease in mean annual PM_{10} concentrations is equivalent to a 0.71% gain in per capita income on average, in our preferred specification (Model 2). The calculation for all of the estimated models is presented in Table 5. This value is in the range found in other studies. In their air quality valuation study of London area residents, and using a very similar methods and data to those used here, MacKerron and Mourato (2009) estimate a 5.3% elasticity between mean annual NO_2 concentration and income (the

authors concluded this result to be implausibly large). Welsch (2002) finds an elasticity of 0.52% with respect to Nitrogen concentrations, focusing only on the direct effects of air pollution on LS (as we do here). But Welsch's analysis is based on national pollution trends – i.e. pollution data at country level – and national average level of satisfaction in 10 countries, which is difficult to compare with the individual model in the present paper. In an OECD study, Silva, Keulenaer et al. (2012) find an elasticity of 1.5% between PM_{10} and income using individual-level survey data combined with city-level PM_{10} data spanning 66 cities around the globe. Despite having some key differences in design (they do not contain rural areas in their data sample), their study is probably most closely comparable to the present one; our valuation estimate is less than half of theirs.

CONCLUSIONS

31. In this study we have combined survey data from four European countries on life satisfaction and perceptions of environmental quality with independent measurements of air quality and urbanity to provide a broad picture of the environmental determinants of life satisfaction, and monetary valuation of air quality improvements. We conclude that environments that respondents perceive as noisy, polluted and lacking in access to green space have a significantly detrimental impact on life satisfaction. But controlling for these aspects of environmental quality, we find also that living in an urban environment tends to correspond to higher levels of life satisfaction. In monetary terms, we also estimate that the value of a 1% reduction in air pollution (measured as mean annual PM_{10} concentrations) is worth the same on average as a 0.71% increase in per capita income.

32. There are some limitations of the present analysis, which provide the basis for future research: As with air quality, it would be ideal to have 'objective' (e.g. GIS-based) measures of households' access to green space as well as some independent measures of noise levels. This would allow not only the valuation of these environmental attributes, but also allow us to circumvent remaining attenuation bias resulting from the endogeneity between LS and these other perception-based, subjective measures of quality.

33. It would also be enlightening to examine the *indirect* effects of air pollution on life satisfaction. For example, the effects of air pollution on income and health and the subsequent effects on life satisfaction should in theory be accounted for in a full valuation exercise. Welsch (2007) conducts an LS analysis of air pollution, with a valuation exercise which accounts for the general equilibrium effects of air pollution, i.e. the notion that air pollution is an "input" into the production of material wellbeing and therefore conveys some material benefits in addition to costs. However, analysing the indirect effects associated with health and income would require more detailed information on health status and on exogenous factors determining wealth/income (e.g. education level of parents).

34. Nevertheless, the unique combination of data – at a highly disaggregated level, but at the same time covering a broad geographic area and tied to independent measurements of environmental quality – provides one of the most detailed pictures yet of how environmental quality is related to life satisfaction. The picture that emerges unambiguously identifies a range of environmental factors which determine people's life satisfaction, and which should in principle be considered in welfare-based policy analysis.

REFERENCES

Callegaro, M. and C. DiSogra (2008). "Computing Response Metrics for Online Panels." Public Opinion Quarterly **72**(5): 1008-1032.

Deguen, S., G. Pédrono, et al. (2008). "Association Between Pollution and Public Perception of Air Quality-SEQAP, a Risk Perception Study in France." Epidemiology **19**(6): S216 210.1097/1001.ede.0000340147.0000345793.0000340140f.

Diener, E. (2006). "Guidelines for national indicators of subjective well-being and ill-being." Applied Research in Quality of Life **1**(2): 151-157.

Diener, E., E. M. Suh, et al. (1999). "Subjective well-being: Three decades of progress." Psychological Bulletin **125**(2): 276-302.

Duncan, G. J. and D. H. Hill (1985). "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data." Journal of Labor Economics **3**(4): 508-532.

Frey, B. S., S. Luechinger, et al. (2009). "The life satisfaction approach to valuing public goods: The case of terrorism." Public Choice **138**(3): 317-345.

Israel, D. and A. Levinson (2003). "Examining the relationship between household satisfaction and pollution." Manuscript, Indiana State University.

MacKerron, G. and S. Mourato (2009). "Life satisfaction and air quality in London." Ecological Economics **68**(5): 1441-1453.

OECD (2012). "Exploring determinants of subjective wellbeing in OECD countries: Evidence from the World Values Survey." OECD Statistics Directorate Working Papers **46**.

OECD (2013). OECD Guidelines on Measuring Subjective Well-being, OECD Publishing.

Pandey, K. D., D. Wheeler, et al. (2006). Ambient Particulate Matter Concentrations in Residential and Pollution Hotspot Areas of World Cities: New Estimates Based on the Global Model of Ambient Particulates (GMAPS). World Bank.

Rehdanz, K. and D. Maddison (2008). "Local environmental quality and life-satisfaction in Germany." Ecological Economics **64**(4): 787-797.

Renfree, J. (2010). Distance calculation using Haversine formula [Matlab implementation].

Schneider, A. (2012). GPS Visualizer.

Silva, J., F. d. Keulenaer, et al. (2012). "Environmental Quality and Life Satisfaction." OECD Environment Working Papers **44**.

Van Praag, B. M. S. and B. E. Baarsma (2005). "Using Happiness Surveys to Value Intangibles: The Case of Airport Noise." The Economic Journal **115**(500): 224-246.

Welsch, H. (2002). "Preferences over Prosperity and Pollution: Environmental Valuation based on Happiness Surveys." Kyklos **55**(4): 473-494.

Welsch, H. (2007). "Environmental welfare analysis: A life satisfaction approach." Ecological Economics **62**(3): 544-551.

Welsch, H. (2008). "The social costs of civil conflict: evidence from surveys of happiness." Kyklos **61**(2): 320-340.

Welsch, H. and J. Kühling (2009). "Using happiness data for environmental valuation: issues and applications." Journal of Economic Surveys **23**(2): 385-406.