**SCHOOL OF
SOCIAL SCIENCES
(DIVISION OF ECONOMICS)**

NANYANG TECHNOLOGICAL UNIVERSITY

HE2023 Housing Economics

Measuring Willingness to Pay for Air Quality Using Hedonic Regression

Written by:

Ivan Wong Ze Xi (U1840701J)

Nigel Yee (U1840714B)

Pang Zheng Jie (U1840542J)

Poon Wei En Eliza (U1931810H)

Table of Contents

1. Introduction

Air pollutants like PM10 and NOx are widely known to give rise to negative externalities, due to their potential hazards to health as documented by existing research. As such, all governments should and already do place an emphasis on mitigating the effects of such pollution. However, policy making for such environmental issues are extremely challenging. A prominent issue lies in the estimation of the value of a non-market entity, the lack of a market renders it difficult to estimate the price, or rather the willingness to pay (WTP) to avoid pollution. Eliciting the WTP is important in creating or updating policies as it facilitates decisions such as the scale, extent or reach of the policy. Hedonic regressions are often used to model non-market entities on the prices of goods, by breaking it down into its relevant constituent characteristics. This paper thus utilises the hedonic housing price model to obtain estimates of households' WTP in London to avoid pollution, which could serve as a useful indicator of whether the London government is taking adequate action to combat the effects of pollution.

2. Literature Review

Currie (2015) explored how industrial plants could affect housing markets and infant health. The research concluded that the toxic air emissions affected air quality only within 1 mile of the plant, with plant openings leading to a 11% decline in housing values within 0.5 mile.

Chay and Greenstone (2005) investigated how nonattainment status from the Clean Air Act (CAA) can be used as an instrumental variable (IV) to isolate changes in Total Suspended Particles (TSP). With the IV in place, they estimated that elasticity of housing values with respect to particulate concentrations ranged between -0.20 to -0.35.

Smith and Huang (1995) reconstructed 86 estimates for the marginal WTP for reducing total suspended particulates. Freeman's (1974) proposed on using point estimates of the MWTP together with linearized MWTP functions to measure the value of air quality improvements. Smith and Huang (1995) focused on developing point estimates of the MWTP for reducing air pollution using two approaches: A Minimum Absolute Deviation (MAD) estimator and an OLS estimator (Huber consistent). They found that the mean MWTP is nearly 5 times greater than the median value, suggesting that outliers are important influencers to any summary statistics for these estimates.

Luechinger (2009) estimated the full WTP for clean air through the incorporation of the life satisfaction approach into the hedonic price model. With the suggestions from previous studies that the hedonic model underestimates the benefits of clean air to a large extent, the author initiated the usage of this standard non-market valuation technique in the construction of the complementing model. In addition, the author explored the usage of instrumental variables through the exploitation of the natural experiment created by the mandated scrubber installation at power plants, with wind directions dividing counties into treatment and control groups.

3. Methodology

The validity of hedonic price regression depends on the inclusion of relevant variables into the model. Control variables should be at least correlated with house prices as well as the 2 different kinds of air pollutants ($lnNOx$ and $lnPM10$). Exclusion of these highly correlated variables will lead to Omitted Variable Bias (OVB). The correlation matrix above indicates the statistical relationship between the variables. We relied on both the statistics and the theoretical understanding of the relationship between variables to determine relevant variables for our model.

Here, we shall illustrate the importance of including the 3 types of control measures (housing, locational and neighbourhood characteristics). If variables that are highly correlated with each other are included, the model will suffer from multicollinearity. From table 1 in Appendix, pairwise correlation matrix verified that the two different air pollutants (*lnNOx* and *lnPM10*). are highly correlated with each other (0.7019) and hence, should be estimated independently in each separate regression models.

In addition, *freehold* and *thamesriv_dist* have high correlation with other variables. As the effect of both variables can be explained by other factors, we decided to not include freehold and *thamesriv_dist*. We choose terrace variable over freehold as the former is an observable feature. Even though *dist_to_cbd* and *bus_distnear* are highly correlated to other variables, we decided to keep them in the model as they play a significant role in the level of pollution from a particular distance. Excluding them could lead to potential omitted variable biasness in β_1, our estimate for WTP for pollution.

3.1 Pooled OLS

In summary, our model specification for Pooled OLS will include all 3 types of controls: Housing characteristics (HC_{it}), Location characteristics (Dist_{it}), and Neighborhood characteristics (NC_{jt}) in order to obtain an accurate measure of the WTP for air pollution:

$$lnPrice_{ijt} = \beta_0 + \beta_1\, lnPollution_{ijt} + \beta_{HC}\, HC_{it} + \beta_{dist}\text{Dist}_{it} + \beta_{NC}NC_{jt} + \varepsilon_{ijt}$$

$lnPrice_{ijt}$ represents the natural logarithm of housing price of house *i*, situated in neighborhood *j*, at time period *t* stated on the transfer deed and all the transactions that took place in London. $lnPollution$ denotes the natural logarithm of the pollution metric, either NOx, or PM10, and ε_{ijt} denotes the random error term. NOx comprises of different oxides of nitrogen and its main sources are from automobile and industrial emissions. PM10 is a particulate matter with a diameter of less than 10 micrometers and its main sources are from combustion, dust from construction, and landfills and agriculture. From the correlation matrix in the Appendix, the correlation coefficient between the two pollutants is 0.7019. This positive relationship indicates that we should run two separate regression models.

We are interested in estimating β_1 that captures the WTP for air pollution. Different variants of the regression models and its results are shown in Table 2.

For the individual housing characteristics in our model, we employed dummy variables:

$$lnPrice_{ijt} = \beta_0 + \beta_1\ lnPollution_{ijt} + \beta_2\ Detached_i\ + \beta_3 Semid_i\ + \beta_4\ Terrace_i\ +$$
$$\beta_5\ Newbuild_i + \varepsilon_{ijt}$$

where $Detached_i,\ Semid_i, Terrace_i$ are equal to one if the house is either a detached house, semidetached house, or a terrace and zero if otherwise(base group is flat). $Newbuild_i$ equals one if the house is newly constructed project and zero if otherwise.

We conducted a regression with housing characteristics control as shown in Table 2 (1) and Table 2 (4). The air pollution estimate is positive which suggest that a 1% increase in NOx and PM10 would indicate a 0.1626% and 0.1261% increase in housing value respectively. This is counter intuitive as people should be paying a premium to stay away from areas with high air pollution. In fact, such a naïve model will undoubtedly result in a strong presence of OVB. We proceed to include locational characteristics.

$$lnPrice_{ijt} = \beta_0 + \beta_1\ lnPollution_{ijt} + \ldots + \beta_6 Tube\_Distnear_i\ +$$
$$\beta_7 Bus\_Dist\_Near_{it} + \beta_8 Nearest\_Park\_Dist_{it} + \beta_9\ Heritage\_Count\_200m_{it} +$$
$$\beta_{10} Dist\_To\_CBD_{it} + \beta_{11} Bus\_StopCount\_200m_{it}\ \varepsilon_{ijt}$$

We conducted a regression to control both housing and location characteristics as shown in Table 2 (2) and Table 2 (5). The new air pollution estimates suggested that a 1% increase in NOx and PM10 would indicate a 0.0604% increase and a 0.156% decrease in housing value respectively. The decrease in the positive estimate for WTP for air pollution in the NOx model indicates a correct direction of our method, and the negative estimate for WTP for air pollution in the PM10 model provides us the correct interpretation. However, we are fully aware that the estimates are still biased as we have yet to include the neighborhood characteristics controls. We proceed to include neighborhood characteristics in a bid to further reduce OVB.

$$lnPrice_{ijt} = \beta_0 + \beta_1\, lnPollution_{ijt} + \ldots + \beta_{12}\, percent\_noedu_{jt} + \beta_{13}\, percent\_minorities_{jt} + \beta_{14}\, percent\_single\_parent_{jt} + \beta_{15}\, grossannual\_pay_{jt} + \beta_{16}\, hoursworked_{jt} + \beta_{17}\, jobdensity_{jt} + \beta_{18}\, unemployment_{jt} + \beta_{19}\, popdensity_{jt} + \varepsilon_{ijt}$$

In this regression, we further control for the neighborhood characteristics as shown in Table 2 (3) and Table 2 (6). The new air pollution estimates suggested that a 1% increase in NOx and PM10 would indicate a 0.0143% decrease and a 0.0608% decrease in housing value respectively. The signs of the estimates for WTP for air pollution are negative in their respective models. We then proceed to control for fixed effects to further reduce biasness.

3.2 Fixed Effects (FE) Model

Next, we attempt to control locational fixed effects at 2 different levels: MSOA and LAD level. LAD stands for Local Area District, which consists of multiple Middle Super Output Areas (MSOA). Recognising that there are district areas (i.e., OA and LSOA) present in the dataset, it is noteworthy to understand the mobility nature of AQ, hence we are controlling yearly fixed effects in the time dimension and locational fixed effects of hierarchy above LSOA.

$$lnPrice_{ijt} = \beta_0 + \beta_1\, lnPollution_{ijt} + \beta_{HC}\, HC_{it} + \beta_{dist}\mathrm{Dist}_{it} + \beta_{NC}NC_{jt} + W_j + t_j + \varepsilon$$

Given that the neighbourhood characteristics are more likely to be homogenous within the same district, FE can control for those unobserved factors that would otherwise cause biased estimates. After controlling for neighbourhood FE at LAD level, table show that the estimates of WTP for *lnnox* are now negative, indicating the strong presence of positive OVB in OLS estimation due to the initial positive coefficient in the OLS model. After controlling for LAD level fixed effects, the model estimated that a 1% increase in *nox* is associated with 0.0089% decrease in house price, while a 1% increase in *PM10* is associated with 0.0360% decrease in house price. Lastly, we decided to take a more detailed approach where we controlled for neighbourhood effects at a Middle Super Output Area (MSOA) level, to explore additional neighbourhood effects on a more granular scale. After controlling for MSOA level fixed effects, the model estimated that a 1% increase in *nox* is associated with 0.0103% decrease in house price, while a 1% increase in *PM10* is associated with 0.0429% decrease in house price. Using both types of pollutants to run our regression model ultimately showed intuitive results as now they both reflected negative effects of pollution on house prices. Additionally, we noted that all added neighbourhood variables at the MSOA level indicated negative coefficients as well, as predicted.

However, which should be the more accurate measure of WTP for AQ? The dataset currently consists of 816 different MSOAs and 36 LADs. Consequently, we can expect roughly 27 times more observations for each LAD level as compared to MSOA level. Acknowledging that urban development within each LAD is usually managed by one local planning authority (Statistics, 2019), we can assume that the unobserved neighbourhood characteristics are likely to homogenous as well, even at a larger district area. Hence, we can satisfy the implicit assumption of homogeneity of unobserved neighbourhood characteristics at the LAD level and yet, preserving a decent number of observations within the location, by controlling FE at the LAD level. Hence our final coefficients for *lnNOx* and *lnPM10* are -0.00887 and -0.0360 respectively.

3.3 Instrumental Variable (IV)

Luechinger (2009) constructed instrumental variable by exploiting the mandated scrubber installation at powerplants and included wind effects in his model. Similarly, we suggest collecting data on the number of scrubbers installed and the level of pollution in the respective districts. The districts can then be classified as located "upwind" and "downwind" of central London. Our model specification for the 2SLS model:

$$lnPrice_{ijt} = \beta_0 + \beta_1 \, ln\widehat{Pollution}_{ijt} + \beta_{HC} \, HC_{it} + \beta_{dist}\mathrm{Dist}_{it} + \beta_{NC}NC_{jt} + W_j + t_j + \varepsilon_{ijt}$$

First-stage-least-square regression:

$$lnPollution_{ijt} = \alpha_0 + \alpha_1 \, (Scrubber_{jt} * Down\,wind_{jt}) + \alpha_{HC} \, HC_{it} + \alpha_{dist}\mathrm{Dist}_{it} + \alpha_{NC}NC_{jt} + W_j + t_j + u_{ijt}$$

where $Down\,wind_{jt}$ is a dummy variable with value 1 if district j lies downwind of central London.

Our assumptions to the model are (1) $Scrubber_{jt} * Down\,wind_{jt}$ is not correlated to ε and (2) $Scrubber_{jt} * Down\,wind_{jt}$ only affects $lnPrice_{ijt}$ through $lnPollution_{ijt}$. This is a potential IV that can be further studied if the data required is present.

4. Further Studies

In this paper, we have attempted to estimate the WTP for air quality as accurately as we could through a hedonic price model. However, there are still gaps to be filled and with further research and data, these estimates could be improved.

To start, studies done pointed out that outliers in data could result in heavily skewed estimations, causing inaccuracy of estimates (Smith & Huang, 1995). To identify the extent to which outliers will affect our estimates, we explored the possibility of removing the top and bottom 1 percentile of the data for variables *price*, *NOx*, and *PM10*. This is to preserve the integrity of the dataset by capturing at least 95% of the observations in the original dataset. After truncating these data, the signs of the variables remained the same and the magnitude did not differ much. Thus, we assumed minimal effects of outliers in our analysis which does not call for mitigation in this aspect.

lnprice	*Observations*	Coefficient of l*nNOx*	Adjusted R^2	RMSE
Before removing	434,205	-0.0088717	0.6298	0.3442
After removing	412,877	-0.0100593	0.6202	0.3155
lnprice	*Observations*	Coefficient of *lnPM10*	Adjusted R^2	RMSE
Before removing	434,205	-0.03599	0.6298	0.3442
After removing	412,877	-0.0281738	0.6202	0.3155

Figure 1: Comparison of coefficients of *lnNOx* and *lnPM10* before and after removing top and bottom 1 percentile for *price*, *NOx*, and *PM10*

Next, we could employ a similar approach as Chay & Greenstone which used changes in regulation i.e. Clean Air Act to simulate an experimental set up with control and treatment groups to estimate the WTP for air quality (Chay & Greenstone, 2005). Taking reference to this, we could use other policies as well that might affect air quality such as Congestion Charging Scheme (Santos, G., Button, K. & Noll, R. G., 2008) or Taxi Emission Strategy (Edie Newsroom, 2005). Lastly, we could consider the addition of interaction variables with pollution as it would allow for comparison of WTP across different subgroups of the population. This could be done by including a dummy variable that segments the population and multiplying it with pollution itself. The model would look as such:

$$lnPrice_{ijt} = \beta_0 + \beta_1\, lnPollution_{ijt} + \beta_{HC}\, HC_{it} + \beta_{dist}\mathrm{Dist}_{it} + \beta_{NC}NC_{jt} + \gamma\, Pollution * Dummy + \varepsilon i$$

Some interaction variables to consider could be education and income. Taking education for example, the dummy would equal one if households are beyond a specified level of education threshold and thus categorized as "educated", and zero if they are not, being "uneducated". The coefficient, γ, measures the additional WTP for air quality for the educated subpopulation. Likewise, for income, we can divide the population into "high income" and "low income" by setting a threshold and thus measure the difference in WTP between the subgroups by looking at γ.

This information will be particularly useful for the next section, application of WTP as it can aid policy makers by using the difference in WTP to target certain subsegments. For instance, if "uneducated" households are found to have lower value for air quality it implies they are uninformed of the severity of pollution and thus result in a lowered sensitivity to living in pollution ridden areas. This could be detrimental to their health and thus policies could target such households, by raising awareness or providing necessary information about pollution.

Similarly, if "low income" households are found to have lower WTP because of their financial situation, it implies their inability to pay to avoid pollution. As such, policy makers could target these households and provide them with subsidies or vouchers to relocate to areas with less pollution.

So far, we have been suggesting improvements for the model and measuring the accuracy of our estimates. However, apart from improving the methods of regression we could look into complementing our estimates obtained through the HPM alongside other valuation methods. This is as HPM alone is insufficient in providing an accurate estimate of WTP for air quality, given it does not capture non-market effects of poor air quality such as health effects. (Currie, 2015)

As such, we can attempt to include the effect of negative externalities arising from pollution by implementing the life satisfaction approach as proposed by Simon Luechinger. The life satisfaction approach is carried out through the regression of self-reported life satisfaction on the public good of interest, income, and various covariates such as age and marital status. The estimated WTP from the new approach will then be summed up with the estimated WTP from the hedonic model to obtain the refined WTP.

Similarly, we could employ contingent valuation method, which is a survey-based method which directly asks respondents to report their WTP for a non-market good, in this case, air quality. This is particularly useful as it indicates how households truly value air quality instead of observing market factors. There are several other methods such as dose response method and replacement cost method. However, these methods also have their complexities and more through research has to be done to use them as an effective complement to HPM.

5. Application of WTP

With our estimates obtained in the hedonic price model, we can approximate the cost of pollution in London. We use an arbitrary year of 2017 to base our estimations upon.

First, we obtained the average house price in London which was about £538,000 in 2017. (Statista, 2020). Then, we use our coefficients from the model to obtain the average WTP for a 1% increase in NOx and PM10, which amounted to approximately (0.00887%*£538,000) = £47.721 and (0.0360%*£538,000) = £193.68 respectively. These results were in line with that of similar studies conducted in other cities, after accounting for inflation (Smith & Huang, 1995). An interesting finding is the large difference in WTP for NOx and PM10. We hypothesised that this could be attributed to differences in toxicity and salience of each pollutant, as studies show that PM10 brings about more biological harm compared to NOx (Muller & Mendelsohn, 2009). Moreover, PM10 is significantly more visible than NOx (Sullivan, 2017) and as such, individuals may perceive the harm of PM10 to be more severe. Coupled with the characteristics of each pollutant, this could explain the differences in WTP to avoid PM10 and NOx.

Then, to obtain the total WTP of London households, we found the total number of households in London to be about 3.29 million and multiplied this with the estimated value of WTP to avoid pollutants. Due to data limitations, we focused on NOx emissions only. The estimated total WTP to avoid 1% of NOx of London households totalled to £47.721*3.29 million = £157 million, which is the cost of a 1% increase in NOx pollution in London.

This estimate can first be applied to assess if the London government has been spending enough to combat pollution. We attempted to do so by finding the percentage decrease in NOx emissions from 2016 to 2017, and multiplying it with the WTP to find the cost of NOx pollution in 2017. Then, we compared it to how much the London government has spent to combat NOx pollution for analysis.

We found that NOx emissions had decreased by 7.56% (Department for Environment Food & Rural Affairs, 2021) from 2016 to 2017, and using our total estimated WTP, the cost to avoid a 7.56% increase in NOx emissions is (7.56*£157 million) = £1.187 billion.

Next, we found the amount spent by the government on combatting NOx pollution. While we were unable to find London specific data, we found that the UK government has spent a total of £3.5 billion to combat NOx pollution in 2017 (Department for Environment Food & Rural Affairs, 2019). We attempted to estimate the amount spent in London alone by applying a weight to the amount spent, by using the proportion of NOx emissions in London as compared to UK. NOx emissions in 2017 was found to be 20860 tonnes in London (Statista, 2020), and 891000 tonnes in UK (Department for Environment Food & Rural Affairs, 2021). As such, London contributes to about $\frac{20860}{891000}$ = 2.34% of NOx emissions. Thus, we estimate the amount spent in London to be about (2.34%*£3.5 billion) = £81.9 million.

As such, we note that the government is not spending enough to fight pollution and should significantly increase its spending on the issue, in order to increase the welfare of its citizens. More generally, the estimated WTP can also be used by policy makers to perform cost benefit analysis for projects. These projects may be projected to bring about economic benefits but also increase pollution levels, thus policy makers can estimate the costs of pollution using the figures we obtained, and make an objective decision whether to approve certain projects.

7. Hedonic Pricing Index

The Hedonic Price Index (HPI) measures price changes in housing over time relative to a reference period, allowing us to observe trends in the housing market due to economic conditions. This can be constructed by adding a vector of time dummies in a specified hedonic regression model, then plotting the coefficients of the time dummies into a graph. Using our final specification of our FE model at the LAD level and adding quarterly time dummies, our estimation equation for the HPI is as follows:

$$lnPrice_{ijt} = \beta_0 + \beta_1\ lnPollution_{ijt} + \beta_{HC}\ HC_{it} + \beta_{dist}\mathrm{Dist}_{it} + \beta_{NC}NC_{jt} + + \delta_{2005Q1}\ Dum_{2005Q1}$$

$$+\ldots\ \delta_{2014Q1}\ Dum_{2014Q1} + \varepsilon i$$

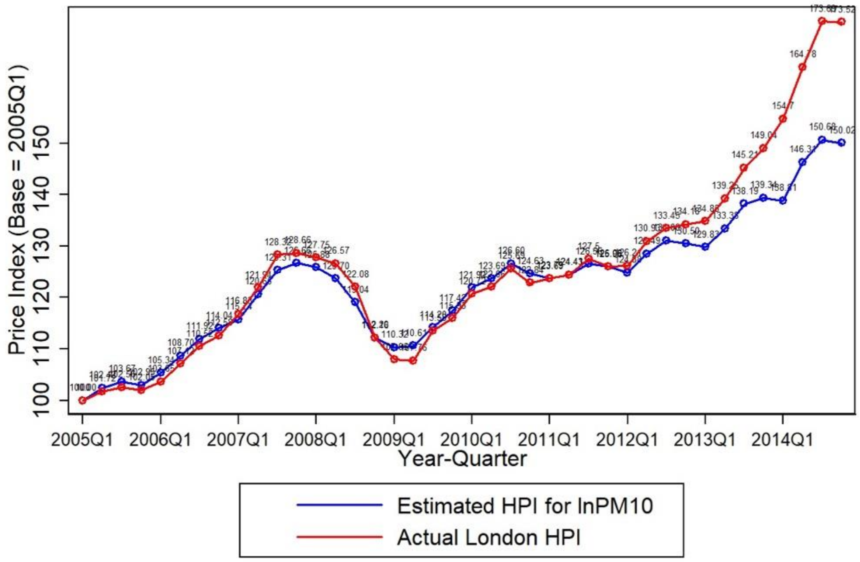

Figure 2: Comparison between the estimated HPI for lnPM10 and actual London HPI (Registry, 2021)

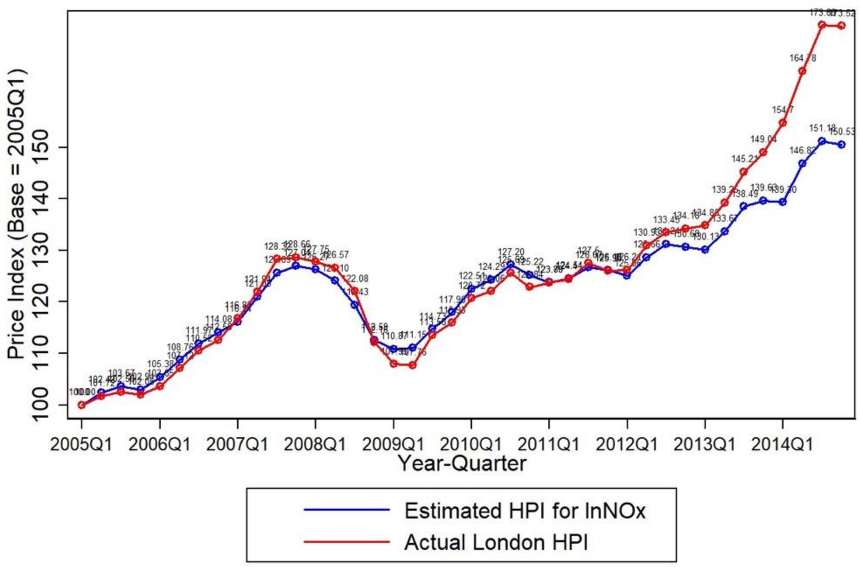

Figure 3: Comparison between the estimated HPI for lnNOx and actual London HPI (Registry, 2021)

As seen from above, there is a general upward trend in housing price which could be attributed to general inflationary pressures. We observe a steep increase in prices from 2005 to 2008, before prices start to fall, in line with the 2008 Global Financial Crisis (GFC). In 2009, prices rise rapidly reflecting the recovery from the GFC, before slowing to moderate growth in 2010. Another significant turning point lies in 2013 where we see large increases in prices again. This could be due to “Help to Buy” policy introduced in the same year, which aimed to enhance housing affordability in the UK. Comparing to the published London House Price Index (Registry, 2021), our estimated HPI for both are in tandem from 2005 to 2012. Deviations from the actual HPI starts to occur after 2012 where our HPI is underestimated compared to the actual HPI. The deviations could be due to the limitations of our dataset used in estimating the HPI. Additionally, the hedonic model that is used in our estimation is close to the limit of its internal validity.

8. Conclusion

Our paper employed pooled OLS, FE, and IV strategies in estimating the WTP for air pollutants. By controlling the FE model at a LAD level, we estimated that a 1% increase in NOx and PM10 causes house prices to decrease by approximately ₤47.721 (0.00887% of house price) and ₤193.68 (0.0360% of house price) respectively. To achieve more accurate WTP values, we can explore the possibility of including controls for other factors such as policy funding data that is being spent to combat air pollutions. Difference-in-difference strategies can also be adopted if data pertaining to changes in air pollution regulations is available. To conclude, estimating WTP for non-market goods such as air pollutants is not easy. From our paper, we found that using hedonic price model provides a relatively good estimate of the WTP. However, we should simultaneously employ other strategies such as contingent valuation method or even non-demand approaches to fully capture the actual WTP for air pollutants.

Appendix:

	lnnox	lnpm10	lnprice	detach~m	semi d~m	terrac~m	freehold	newbuild	p~tnoedu	percen..	percen~e	grossa~y	jobden~y	popdens	hoursw~d	unempl~t	thames~t	tube d~r	bus di~r	Bus~200m	Neares~t	her~200m	dist t~d
lnnox	1																						
lnpm10	0.7019*	1																					
lnprice	0.1166*	0.0161*	1																				
detached c	-0.0289*	-0.0500*	0.2468*	1																			
semi d dur	-0.0746*	-0.0760*	0.1182*	-0.0939*	1																		
terrace dur	-0.0533*	-0.0041*	0.0409*	-0.1350*	-0.2626*	1																	
freehold	-0.1137*	-0.0800*	0.2279*	0.2305*	0.4397*	0.6112*	1																
newbuild	-0.0182*	-0.0148*	0.0075*	-0.0456*	-0.1203*	-0.1639*	-0.2502*	1															
percentnoe	-0.1574*	-0.0788*	-0.4042*	-0.0140*	0.0942*	0.1694*	0.2158*	-0.0805*	1														
percentmir	-0.1019*	0.0171*	-0.2646*	-0.1033*	-0.0654*	0.0495*	-0.0432*	0.1045*	0.1975*	1													
percentlon	-0.0774*	-0.0186*	-0.3801*	-0.1205*	-0.0740*	0.1130*	-0.0040*	0.0146*	0.4823*	0.3943*	1												
grossannu	0.0427*	-0.0439*	0.2607*	-0.0815*	-0.1628*	-0.1441*	-0.2796*	0.1590*	-0.2419*	0.0267*	-0.1223*	1											
jobdensity	0.1000*	0.0386*	0.1076*	-0.0178*	-0.0370*	-0.0564*	-0.0833*	0.0371*	-0.0958*	-0.0525*	-0.0780*	0.3172*	1										
popdens	0.1604*	0.1244*	0.2882*	-0.1674*	-0.2747*	-0.1161*	-0.3719*	0.1045*	-0.2846*	0.0712*	-0.0682*	0.5935*	0.0119*	1									
hourswork	0.1080*	0.0739*	0.0122*	0.0497*	0.0849*	0.0642*	0.1374*	-0.1405*	0.0210*	-0.0728*	0.0458*	-0.5753*	-0.1843*	-0.2799*	1								
unemployr	-0.0822*	0.0993*	-0.1111*	-0.0936*	-0.1390*	0.0418*	-0.1046*	0.1086*	0.0040*	0.2784*	0.0688*	0.4243*	-0.0095*	0.3717*	-0.3337*	1							
thamesriv	-0.1603*	-0.1889*	-0.1354*	0.2218*	0.2582*	-0.0176*	0.2659*	-0.1116*	0.2168*	-0.0624*	0.0497*	-0.4473*	-0.0831*	-0.5712*	0.2027*	-0.3705*	1						
tube distn	-0.0492*	-0.0932*	-0.1825*	0.1226*	0.1390*	0.0541*	0.2034*	-0.0901*	0.1888*	-0.2217*	0.0927*	-0.3011*	-0.0687*	-0.3956*	0.2031*	-0.2871*	0.3357*	1					
bus distne	-0.0637*	-0.0481*	0.0742*	0.1258*	0.1162*	0.0717*	0.1999*	-0.0509*	0.0133*	-0.0567*	-0.0453*	-0.1076*	-0.0324*	-0.1484*	0.0561*	-0.0495*	0.1484*	0.0363*	1				
Bus sto~2	0.1009*	0.0809*	-0.0312*	-0.0966*	-0.1307*	-0.0938*	-0.2167*	0.0771*	-0.0487*	0.0527*	0.0169*	0.1241*	0.0479*	0.1810*	-0.0469*	0.0465*	-0.1611*	-0.0519*	-0.6755*	1			
Nearest Pa	-0.1539*	-0.2090*	-0.2524*	0.2210*	0.2200*	0.0288*	0.2807*	-0.0751*	0.2608*	-0.0759*	0.1084*	-0.3531*	-0.0995*	-0.5759*	0.1865*	-0.2497*	0.7694*	0.5283*	0.1496*	-0.1622*	1		
heritag~20	0.1175*	0.0844*	0.1320*	-0.0249*	-0.0485*	-0.0556*	-0.0993*	0.0314*	-0.0939*	-0.0188*	-0.0741*	0.1715*	0.1877*	0.1164*	-0.0577*	0.0209*	-0.1080*	-0.0741*	-0.0631*	0.0956*	-0.1387*	1	
dist to cbd	-0.2234*	-0.2227*	-0.3172*	0.2005*	0.2901*	0.1142*	0.3971*	-0.0801*	0.3582*	-0.0258*	0.1076*	-0.4633*	-0.1424*	-0.8250*	0.2100*	-0.2596*	0.6266*	0.3882*	0.1773*	-0.2112*	0.7604*	-0.1738*	1

* Significant at 5% level

Table 1 Pairwise correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	HCNOX	HCLCNOX	HCLCNCNOX	HCPM10	HCLCPM10	HCLCNCPM10	LADNOX	MSOANOX	LADPM10	MSOAPM10
VARIABLES	lnprice	lnprice	lnprice	lnprice	lnprice	lnprice	lnprice	lnprice	lnprice	lnprice
lnnox	0.16268***	0.06037***	-0.01432***				-0.00887***	-0.01029***		
	(0.00175)	(0.00149)	(0.00123)				(0.00128)	(0.00246)		
lnpm10				0.12614***	-0.15570***	-0.06077***			-0.03599***	-0.04293***
				(0.00400)	(0.00331)	(0.00291)			(0.00301)	(0.00398)
detached_dum	0.80429***	1.19172***	1.07278***	0.79038***	1.19273***	1.07304***	1.05673***	0.97644***	1.05695***	0.97642***
	(0.00401)	(0.00389)	(0.00364)	(0.00399)	(0.00388)	(0.00363)	(0.00350)	(0.00319)	(0.00350)	(0.00318)
semi_d_dum	0.32706***	0.63603***	0.61755***	0.30786***	0.63301***	0.61755***	0.60936***	0.61728***	0.60932***	0.61733***
	(0.00216)	(0.00207)	(0.00177)	(0.00217)	(0.00207)	(0.00177)	(0.00165)	(0.00160)	(0.00165)	(0.00160)
terrace_dum	0.19884***	0.36295***	0.44974***	0.18264***	0.36128***	0.44979***	0.45443***	0.48259***	0.45440***	0.48257***
	(0.00197)	(0.00178)	(0.00145)	(0.00199)	(0.00177)	(0.00145)	(0.00132)	(0.00132)	(0.00132)	(0.00132)
newbuild	0.14530***	0.14748***	0.17150***	0.13389***	0.13842***	0.17133***	0.20096***	0.18360***	0.20089***	0.18346***
	(0.00267)	(0.00234)	(0.00228)	(0.00268)	(0.00232)	(0.00228)	(0.00213)	(0.00208)	(0.00213)	(0.00208)
tube_distnear		-0.00003***	-0.00003***		-0.00002***	-0.00003***	-0.00001***	-0.00001***	-0.00001***	-0.00001***
		(0.00000)	(0.00000)		(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Nearest_Park_dist		0.00000	-0.00001***		-0.00000**	-0.00001***	-0.00000***	0.00001***	-0.00000***	0.00001***
		(0.00000)	(0.00000)		(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
heritage_count_200r		0.12412***	0.05761***		0.13493***	0.05769***	0.05775***	0.04382***	0.05783***	0.04387***
		(0.00338)	(0.00318)		(0.00339)	(0.00318)	(0.00317)	(0.00333)	(0.00317)	(0.00333)
dist_to_cbd		-0.04786***	-0.02093***		-0.04982***	-0.02113***	-0.03380***	-0.03136***	-0.03399***	-0.03149***
		(0.00018)	(0.00025)		(0.00018)	(0.00025)	(0.00045)	(0.00217)	(0.00046)	(0.00217)
Bus_stopscount_20C		-0.01073***	-0.00709***		-0.00938***	-0.00705***	-0.00528***	-0.00673***	-0.00523***	-0.00671***
		(0.00037)	(0.00031)		(0.00037)	(0.00031)	(0.00029)	(0.00029)	(0.00029)	(0.00029)
percentnoedu			-0.01233***			-0.01234***	-0.00970***	-0.00620***	-0.00971***	-0.00621***
			(0.00008)			(0.00008)	(0.00008)	(0.00008)	(0.00008)	(0.00008)
percentminorities			-0.00376***			-0.00375***	-0.00568***	-0.00361***	-0.00567***	-0.00361***
			(0.00003)			(0.00003)	(0.00004)	(0.00006)	(0.00004)	(0.00006)
percentlone			-0.01069***			-0.01069***	-0.00814***	-0.00505***	-0.00811***	-0.00505***
			(0.00012)			(0.00012)	(0.00012)	(0.00012)	(0.00012)	(0.00012)
grossannualpay			0.00002***			0.00002***	-0.00000***	0.00000	-0.00000***	0.00000
			(0.00000)			(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
hoursworked			0.11294***			0.11208***	0.00665***	0.01158***	0.00690***	0.01186***
			(0.00124)			(0.00124)	(0.00186)	(0.00180)	(0.00186)	(0.00179)
jobdensity			0.08183***			0.08408***	0.51250***	0.57293***	0.50599***	0.56838***
			(0.00174)			(0.00175)	(0.01510)	(0.01507)	(0.01510)	(0.01506)
unemployment			-0.04703***			-0.04573***	-0.01699***	-0.01581***	-0.01672***	-0.01529***
			(0.00034)			(0.00034)	(0.00059)	(0.00058)	(0.00059)	(0.00059)
popdens			0.00246***			0.00243***	0.00079***	0.00184***	0.00069***	0.00173***
			(0.00004)			(0.00004)	(0.00026)	(0.00027)	(0.00026)	(0.00027)
Constant	11.69107***	12.70357***	8.34805***	12.04835***	13.50916***	8.52639***	12.77079***	12.07059***	12.84648***	12.15394***
	(0.00814)	(0.00730)	(0.04944)	(0.01287)	(0.01104)	(0.05069)	(0.07534)	(0.07507)	(0.07572)	(0.07542)
Observations	450,885	450,885	434,205	450,885	450,885	434,205	434,205	434,203	434,205	434,203
R-squared	0.12159	0.35974	0.55968	0.10158	0.36000	0.55999	0.62981	0.68249	0.62990	0.68256

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

Table 2 OLS and FE estimation result

References:

Brasington, D. M., Hite, D., & Jauregui, A. (2015) House price impacts of racial, income, education, and age neighbourhood segregation. *Journal of Regional Science*, *55*(3), 442-467. https://ideas.repec.org/a/bla/jregsc/v55y2015i3p442-467.html

Campbell, J. Y., Giglio, S. & Pathak P. (2011) Forced sales and house prices. *American Economic Review* 101(5), 2108-2131. https://doi.org/10.1257/aer.101.5.2108

Chay, K. Y., & Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, 113(2), 376–424. https://doi.org/10.1086/427462

Currie, J., Davis, L., Greenstone, M., & Walker, R. (2015). Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings. *American Economic Review*, 105(2), 678–709. https://doi.org/10.1257/aer.20121656

Department for Environment Food & Rural Affairs. (2019). Clean Air Strategy 2019. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/770715/clean-air-strategy-2019.pdf

Department for Environment Food & Rural Affairs. (2021). National Statistics: Emissions of Air Pollutants in the UK – Nitrogen Oxides (NOx). https://www.gov.uk/government/statistics/emissions-of-air-pollutants/emissions-of-air-pollutants-in-the-uk-nitrogen-oxides-nox

Edie Newsroom. (2005). Taxi Emissions Strategy. https://www.edie.net/library/Taxi-Emissions-Strategy/2667

Land Registry. (2021). UK House Price Index. https://landregistry.data.gov.uk/app/ukhpi/browse?from=2000-01-01&location=http%3A%2F%2Flandregistry.data.gov.uk%2Fid%2Fregion%2Flondon&to=2018-04-01&lang=en

Le Boennec, R., & Salladarré, F. (2017). The impact of air pollution and noise on the real estate market. The case of the 2013 European Green Capital: Nantes, France. *Ecological Economics*, *138*, 82-89. https://doi.org/10.1016/j.ecolecon.2017.03.030

Luechinger, S. (2009). Valuing air quality using the life satisfaction approach. *The Economic Journal*, 119(536), 482–515. https://doi.org/10.1111/j.1468-0297.2008.02241.x

Muller, N., Mendelsohn, R. (2009). Efficient Pollution Regulation: Getting the Prices Right. *American Economic Review* 99 (5), 1714–1739. https://www.jstor.org/stable/25592534

Office for National Statistics. (2017). Total Number of Households by Region and Country of the UK, 1996 to 2017. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/adhocs/005374totalnumberofhouseholdsbyregionandcountryoftheuk1996to2015

Santos, G., Button, K., and Noll, R. G. (2008). London Congestion Charging. *Brookings-Wharton Papers on Urban Affairs*, 177–234. http://www.jstor.org/stable/25609551

Smith, V. K., & Huang, J. (1995). Can markets value air quality? A meta-analysis of hedonic property value models. *Journal of Political Economy*, 103(1), 209–227. https://www.jstor.org/stable/2138724

Statista Research Department. (2020). Total dwellings: Simple average house price in London 2000-2019. https://www.statista.com/statistics/286048/total-dwellings- simple-average-house-price-in-london/

Sullivan, D. M. (2017). The True Cost of Air Pollution: Evidence from the Housing Market http://www.danielmsullivan.com/pdf/Sullivan_Cost_of_Pollution_housing.pdf

Tiseo, I. (2020). Annual Road Transport NOx Emisions in London 2016-2019, by Location. *Statista*. https://www.statista.com/statistics/1176800/average-annual-nox-road-transport-emissions-london-uk/

Wen, H., Yue Xiao, Hui, E., & Ling, Z. (2018) Education quality, accessibility, and housing price: Does spatial heterogeneity exist in education capitalization? *Habitat International, 78*, 68-82. https://doi.org/10.1016/j.habitatint.2018.05.012