The Causal Effects of Income Support and Housing Benefits on Mental Well-Being: An Application of a Bayesian Network

Oznur Ozdamar

1. Adnan Menderes University, Aydın Faculty of Economics, Department of Econometrics, Aydın, Turkey
2. Bologna University, Department of Economics, Bologna, Italy

oznur.ozdamar@adu.edu.tr, oznur.ozdamar@unibo.it

Eleftherios Giovanis
Verona University, Department of Economics, Italy
giovanis95@gmail.com, eleftherios.giovanis@univr.it

The Causal Effects of Income Support and Housing Benefits on Mental Well-Being: An Application of a Bayesian Network

Abstract

This study explores the causal effects of air pollution, income support, housing benefits and household income on the subjective mental well-being in United Kingdom (UK). Additionally, the analysis considers the effects of air pollution and weather conditions. The estimates are based on data from the British Household Panel Survey (BHPS). The results show that those who are unemployed or who have a low income and who claim the benefits report higher levels of mental well-being than those who do not claim them. Moreover, the marginal willingness to pay (MWTP) for an improvement on air quality are lower in the case of the Bayesian Network.

Keywords: Air Pollution, Bayesian Networks, Housing Benefits, Income Support, Subjective Mental Well-Being

JEL Codes: I31, H41, Q53

1. Introduction

The benefits due to improving health in the last years are clear and significant. The motivation of this study is to examine the causal effects of housing and income support on mental well-being using the Life Satisfaction Evaluation (LSE) approach and Bayesian Networks (BNs). More specifically, among those who are qualified for these benefits, the mental well-being between those who receive and those who do not receive the benefits is compared. In the case where a positive effect of these benefits on mental well-being of the claimants is found relatively to those who are qualified but do not claim them for, then these benefits might actually have positive implications in various fields of economy, including improvement on well-being, labour market participation recovery and increase in productivity among others.

The analysis in this study considers air pollution and weather conditions as additional factors that can have significant impact on individuals' well-being. Regarding the air pollution and weather conditions, one of the main advantage of the LSE approach is that it does not require assumptions of causal relationships. It just assumes that pollution and weather leads to change in life satisfaction. However, one of the main drawbacks of the LSE approach is the reverse causality between income and well-being. For instance, , Pischke (2011) shows that the effect of the income-life satisfaction is mostly causal; however people who are happier might earn more due to the existence of the reverse causality. Similarly, there might be a possible degree between mental well-being and the benefits (housing and income support) examined in this study. A solution for this issue is to use instrumental variables approach (Luechinger, 2009; Ferreira and Moro, 2010). However, Stutzer and Frey (2012) suggest that instrumental variable approaches are difficult to convince especially in the case of happiness and life satisfaction, because it is almost impossible to find an appropriate instrumental variable, since any factor can determine and affect an individual's overall well-being. In addition, the main issue in all the methods and approaches, including IV, natural and randomized experiments, as well as, the Bayesian Networks, which is proposed in this case, is the unobserved confounders which may affect the treatment.

The analysis in this study relies on detailed micro-level data, based on local authority districts, which provides more precise air and weather mapping on individual's residence instead of using cities or counties like other studies did before (Ferreira and Moro, 2010; Luechinger, 2009; Ferreira et al., 2013). Thus, the advantage of using a more detailed geographical reference in order to map air pollution and weather conditions, implies more precise and more robust estimates. Secondly, the analysis relies on individual level panel data, so that unobserved individual level and geographical characteristics can be accounted for. Well-being may be correlated with some unobserved amenities that also affect pollution levels and benefits, thus in the case of cross sectional data, the LSE may be biased. Thirdly, this study uses the non-movers sample, who are those that have not moved to another location or residence. The reason of considering this sample is an effort to limit the endogeneity which comes from the "sorting" problem that it can be plausible when people choose where to reside.

Next, the marginal willingness-to-pay (MTWP) for an improvement on air quality is calculated. This is used as an example in order to show that the income and benefits effects on well-being are significantly stronger, hence the MWTP is lower, considering possible endogenous and selection biases, as well as, over-control bias with the BN framework. This has important implications, especially in the case of the valuation of public goods, as is the air pollution in this paper. Additionally, three major air pollutants are explored; ozone (O_3), nitrogen oxides (NO_X) and carbon monoxide (CO).

Studies as by Luechinger (2009), Levinson (2012), Ferreira and Moro (2010) among others found a systematic negative effect of air pollution on life satisfaction and happiness, while other studies show evidence of the adverse health effects of air pollution (Chung et al., 2011; Patankar et al., 2011; Gonzalez-Barcala et al., 2013). However, studies as by Ferreira and Moro (2010), Ferreira et al. (2013) and MacKerron and Mourato (2009) rely on cross sectional data and do not account for the endogeneity of pollution; i.e. areas with high pollution levels are likely to also have some other amenities that negatively affect well-being. For this reason, this study employs an analysis using panel data.

Regarding the impact of housing benefits on mental well-being, theoretically can be positive. Housing benefits, which imply support with housing can improve the health and well-being of individuals and lead to demand reduction for health and social care services (Johnson, et al., 2006; Bolton, 2009). On the other hand, the literature shows a strong evidence of reverse causal effect that mental health can lead also to homelessness (Johnson, et al., 2006; Bolton, 2009). More specifically, people with mental health problems are less likely to own a house and less likely to live in a stable environment. To summarise, this study tries to fill a gap in the previous literature by examining additional factors on mental well-being, such as air pollution and weather conditions and employing Bayesian Networks for causal inference.

The paper is organized as follows. Section 2 presents a short description of the income support and housing benefits in UK. Section 3 describes the methodology, while in section 4 the data sample is provided. In section 5 the empirical results are reported and in section 6 the concluding remarks are presented.

2. Income Support and Housing Benefits

This section describes the requirements for the individuals and households who are qualified for income support and housing benefits. These benefits refer to a certain group of people who do not have enough money to live on and to pay their rent. These benefits are for adult people who are more than 18 years old. Those who are 16-17 years old can claim the income support benefit in the case they have a child or are pregnant. Individuals without partners must either not be working at all or they should work less than 16 hours a week. In the case they have partner, he /she must work under 24 hours a week. Finally, a claimant should not have a capital income more than £16,000, where capital income is defined as the sum of asset income and private retirement pensions (Bardasi et al., 1999; Fräßdorf, 2011). Regarding the housing benefits those who have capital income more than £16,000 and are full time students are not eligible to claim the benefit.

3. Methodology

3.1 Fixed Effects

Self-assessed well-being measures can serve as empirically valid and adequate approximations of individual welfare. The following model of subjective well-being for individual i, in area j at time t is estimated using the life satisfaction approach (LSA):

$$GHQ_{i,j,t} = \beta_0 + \beta_1 \log(y_{i,t}) + \lambda' e_{j,t} + \delta' ben_{i,t} + \phi' z_{i,j,t} + \gamma W_{j,t} + \mu_i + l_j + \theta_t + l_j T + \varepsilon_{i,j,t} \quad (1)$$

$GHQ_{i,j,t}$ is the subjective well-being *caseness measure*. The vector $e_{j,t}$ is a vector of the measured air pollutants in location j and in time t, $log(y_{i,t})$ denotes the logarithm of the household income, $ben_{i,t}$ denotes the examined social benefits such as income support and housing benefits. Vector z includes household and demographic factors, discussed in the next section. W is a vector of meteorological variables, in location j and in time t. Set μ_i denotes the individual-fixed effects, l_j represent the location (local authority) fixed effects, θ_t is a time-specific vector of indicators for the day and month the interview took place and the survey wave, while l_jT is a set of area-specific time trends. Finally, $\varepsilon_{i,j,t}$ expresses the error term which we assume to be *iid*. Standard errors are clustered at the area specific local authority district level.

In the case of cross sectional data, ordered Probit or Logit are the most appropriate approaches in order to capture the non-linearities of the subjective well-being. However, in the case of panel data, as in this study, it is not possible to apply ordered Probit or Logit with fixed effects, but only with random effects. In this case, the adapted Probit fixed effects (FE) approach proposed by van Praag and Ferrer-i-Carbonell (2004) is applied. Van Praag and Ferrer-i-Carbonell (2004; 2006) show both heuristically and in several applications that the adapted Probit is virtually identical to the traditional ordered Probit analysis. The second approach is the “Blow-Up and Cluster” (BUC) estimator (Baetschmann et al., 2015). An alternative estimator is the Ferrer-i-Carbonell and Frijters (FCF) estimator developed by Ferrer-i-Carbonell and Frijters (2004), but is not employed in this study as it is inconsistent due to its way of choosing the cutoff point based on the outcome that produces a form of endogeneity and leads to large loss of data (Baetschmann et al., 2015).

The marginal willingness-to-pay (MWTP) for an improvement in air pollution can be derived from differentiating (1) and setting dGHQ=0. This is the income drop that would lead to the same reduction in life satisfaction than an increase in pollution. Thus, the MWTP can be computed as:

$$MWTP = -\frac{dy}{de} = -\frac{\partial f}{\partial e} / \frac{\partial f}{\partial y} \tag{2}$$

3.2 Dynamic panel regressions

The second model which can be considered is the Generalized Methods of Moments (GMM) system and it can be defined as:

$$\begin{aligned} GHQ_{i,j,t} &= \beta_0 + \beta_1 \log(y_{i,t}) + \beta_2 GHQ_{i,j,t-1} + \lambda' e_{j,t} + \delta' ben_{i,t} + \phi' z_{i,j,t} + \gamma W_{j,t} + \\ &\mu_i + l_j + \theta_t + l_j T + \varepsilon_{i,j,t} \end{aligned} \tag{3}$$

The dynamic models are useful because the lagged dependent variable controls for a dependent variable that follows an autoregressive-AR(1) process and it shows how an individual changes his or her adaptation level to living conditions represented by the stimulus level in the preceding period. The most important issue of (3) is the reverse causality between income and well-being, as well as between benefits and well-being, thus these regressors may be correlated with the error term. Furthermore, time-invariant fixed effects personal, demographic and geographical characteristics may be correlated with the explanatory variables. Function (3) presents the mentioned problems when *T*, denoting time, is short. More specifically, the Blundell – Bond estimator was designed for small-*T* and large-*N* panels, where *N* denotes the region or individual effects. Therefore this study examines the Blundell-Bond (1998) system GMM.

3.3 Bayesian Networks

This section discusses the directed acyclic graphs (DAGs) and describes the Bayesian Network used in this study for causal inference. DAGs consist of three elements: variables (nodes, vertices), arrows (edges), and missing arrows. *Arrows* represent the possible *direct causal effects* between pairs of variables and order the variables in time. The arrow between *T* and *F* in figure 1 means that *T* may have a direct causal effect on *F*. The variables that are directly caused by a given variable are called its *children*. Considering figure 1, *B* has three children that are *C*, *D*, and *T*. All variables directly or indirectly caused by a given variable are called its *descendants*. For example, the *descendants* of *B* are *C, D, T* (B's children), *E* (D's and T's child), *F* (T's child) and *Y* (child of *A, C, D, E, F*). On the other hand, *parents* are the variables that directly cause other variable(s). Coming back to figure 1 the only parent of *F* is *T*. The opposite of *descendants* are called *ancestors* which are the variables that directly and indirectly cause of other variable(s).

(Insert figure 1)

Definition 1. (Markovian parents) (Pearl, 2000): *Let $V = \{X_1, X_2, . . . X_v\}$ be set of variables, and let $P(v)$ be the joint probability distribution on these variables. A set of variables PA_j is said to be Markovian parents of X_j if PA_j is a minimal set of predecessors of X_j that renders X_j independent of all its other predecessors.*

$$p(x) = \prod_{i=1}^{m} p(x_i \mid par_i) \tag{4}$$

Applying the chain rule of probability, we have:

$$p(x) = \prod_{i=1}^{m} p(x_i \mid x_1, ..., x_{i-1}) \tag{5}$$

Relation (5) uses the back-door criterion. More specifically, estimating the effect of a factor of interest *X* on the outcome of interest *Y*, a back-door path is an undirected path between *X* and *Y* with an arrow into *X*. These paths create confounding, by providing an indirect non causal channel along which information can flow. Thus, a set of conditioning variables or controls *Z* satisfies the backdoor-criterion when *Z* blocks every back-door between *X* and *Y*. Moreover, no node in *Z* is a descendant of *X* or both descendent of *X* and ancestor of *Y* because it will block the causal path between *X* and *Y*. Based on (5) and the back-door criterion, the causal effect of *C* to *Y* in figure 1 can be simply estimated by running a regression of *A*, *B* and *C* on *Y*, since *C* has a direct effect and variables *A* and *B* are its parents and confounders. There is also no other confounder exists in this case. As it has been mentioned, a set of variables *Z* blocks every back-door path between *X* and *Y*. Similarly for the causal effects of *E* on *Y*, will be a regression from *E* and its parents *D* and *T*. For example *T* blocks every back-door path between *F* and *Y* and there is no node-variable in set *Z* which is a descendant of *F*. This is true; since no node in *T* (*B*, *A*, *C*) is descendant or child of *F* and they do not block the causal path.

On the other hand, figure 1 shows that *F* blocks the causal effect from *T* to *Y*. This implies that there is not any indirect effect from *T* to *Y* and *Y* is independent from *T* give *F* $Y \perp T | F$. In this case, using the back-door criterion, a partial regression of *T* conditioning on its parent *B* on *Y* takes place but excluding *F* since it creates over-control bias. In that case set *Z*=*B* meets the back-door criterion, since it is parent and not descendant of *T* and it does not block the path. A test for conditional independence is therefore a test for partial correlation between the variables and the partial correlations can be estimated, via regression analysis. The DAG is estimated with PC algorithm and a pseudo-code is reported in figure 2 (Spirtes et al., 2000). In other words Spirtes et al., (2000) suggest to use the Fisher's Z to test the independence between X and Y testing the following hypothesis that *X* and *Y* are independent given on a set of variables *C*:

$$\rho_{XY|C} = 0 \tag{7}$$

$$z(\rho_{XY|C}, n) = \frac{1}{2}\sqrt{n-|C|-3}\,\log\frac{(|1+\rho_{XY|C}|)}{(|1-\rho_{XY|C}|)}$$

(8)

|C| is the number of variables in C and *n* is the length of the sample. If X,Y,C~N under the null hypothesis of zero partial correlation, it is:

$$z(\hat{\rho}_{XY|C}, n) \sim N(0,1)$$

(9)

(Insert figure 2)

Based on this we discuss next the *d-separation* condition (Pearl, 1988; Spirtes et al., 2000; Neapolitan, 2003) which is especially important and useful in constructing a BN because it controls possible confounders and tests if the effect of one variable to another is identifiable. Graphically, *d*-separation exhibits two main cases: firstly $X \rightarrow S \rightarrow Y$ and secondly $X \leftarrow S \rightarrow Y$. Thus, in the first as we have shown before, it is implied that $Y \perp X | S$ and the causal effects of *X* on *Y* can be found by using front-door criterion, while the effect of *S* on *Y* is direct. The second relation is very important and it tells us $X \leftarrow S \rightarrow Y$ that the factor of interest *X* and the outcome of interest *Y* have a common cause which is the confounder *S*. This is what is desired in regression analysis, since we want to control for variables that cause both *X* and the outcome *Y*. Thus, so far the merits of Bayesian networks are mainly four: Firstly it is possible to see the direction of the effect among the variables. Secondly, it is possible to find the causal effect even if this is blocked by another variable. Thus, BN can be used as robustness checks or priory to explore the causal relations graphically. Thirdly, the relation $X \leftarrow S \rightarrow Y$

guarantees that the appropriate control variables and confounders are considered. This is another very useful information which can be further used in the regression model. Fourthly, the *d-separation* and DAG do not include the relation $X \rightarrow S \leftarrow Y$, which is selection bias. More specifically, it is selection since *X*, the factor or treatment of interest, and the outcome of interest *Y* are conditioned or they cause variable *S*. This is very important, since priori the information on selection bias can be unknown, thus, conditioning on a variable *S* which is caused both by *X* and the dependent variable *Y* will lead to selection bias. Concluding, BN can provide information for the quality of variables and which ones should be included into the regression, depending on the factor of interest or the treatment –intervention- variable, accounting for confounding and endogenous bias. Overall, BN put discussions about causality on a solid mathematical basis and the logic is that the relationship can be measured at least between three variables where one of them can act as a "virtual control" for the relationship of the other two so to no be always necessary to conduct experiments. For instance knowing the marital status it is possible to examine what will be the effect of job status (e.g. increasing employment) intervention on income support.

4. Data

The British Household Panel Survey (BHPS) is used for the entire analyses which started in 1991 and it is an annual survey of individuals of a nationally representative sample of more than 5,000 households in United Kingdom. Individuals moving out or into the original household are also followed (Taylor et al. 2010). The data period used in the current study covers the waves 1-18, for the years 1991-2009. The BHPS has been extensively used for the empirical work on life satisfaction / happiness (Clark and Oswald 1994). Based on the literature, the demographic and household variables of interest in this paper are household income[1], gender, age, age squared, household size, labour force status, house tenure, marital status, education level and local authority districts. The income is measured in thousands of pounds and the year basis is 2010. The regressions control for the day of the week, month of the year and the wave of the survey. The area-specific trends are included as additional controls, since these variables are likely to be correlated both with health status and air pollution level.

Furthermore, the weekly average of the air pollutants preceding the interview is computed, in order to reduce the variation, to increase the robustness of the estimations and in an effort to capture the missing values. This is considered for the reason that the specific time of the interview is unknown and thus the air pollution on the same day may have little or even insignificant effect on well-being. This is especially related when the interview is attending during the early morning hours. In addition, the household income of the last month is considered. In order to limit endogeneity, the non-movers sample is selected. This sample includes the individuals up to their first move to another location and the fraction over the total sample is around 70 per cent. The mental well-being measure examined in this study is the General Health Questionnaire (GHQ) "Caseness Scores" used by Clark and Oswald (1994). More specifically, the GHQ score combines the answer of twelve questions, each on a four-point scale. The GHQ level of mental distress score ranges from 0 to 12, where 12 is the lowest feeling of well-being, and 0 indicates the lowest mental distress. Thus, a negative sign of the coefficient will imply that the specific factor has a positive effect on well-being.

Three major air pollutants are examined: O_3, NO_X and CO and are measured in $\mu g/m^3$. In order to match the air pollution emissions with the individuals the following steps are followed. Firstly, the exact location of the air monitoring stations is known and it is expressed on grid points –eastings and northings- which can be found on DEFRA (http://uk-

[1] The analysis was also conducted using individual level income; however this is affected by labour force participation which we do not explicitly model here.

air.defra.gov.uk) and the London Air Quality Network (http://www.londonair.org.uk) websites. The weather data have been derived by the MIDAS database of the British Atmospheric Data Centre (BADC) and the National Climatic Data Center (NCDC) which has available data for many countries around the world. Secondly, there is special access to the individuals' local authority district (LAD) level, which is also expressed on grid references provided by the Institute for Social and Economic Research (ISER) at the University of Essex.

In order to convert the point data from the monitoring stations into the data up to LAD Level the inverse distance weighting (IDW), which is a GIS-based interpolation method, is employed. In IDW, the weight of a sampled data point is inversely proportional to its distance from the estimated value. Firstly the centroid of each LAD is calculated and then the distance between the air pollution monitor and the centre of the LAD is measured using the Euclidean distance and a radius of 10 km. The unique feature of these data is the information that they provide about the location of individual's residence down to a disaggregated level which allows us to identify more precisely than other geographical references, including cities or counties.

In table 1 and panel A, the summary statistics for the GHQ are reported. A lower mean value implies better levels of health status. In the case of the total sample, the average value is 1.9 implying a better mental well-being status than those who are eligible but either receive or not receive the benefits examined. For instance, the average GHQ for those who are eligible and receive the housing benefits and the income support are 2.55 and 2.76 respectively, indicating that this group presents lower levels of mental well-being than the overall BHPS sample. However, those who are eligible and do not receive the benefits, are more likely to report higher levels of GHQ, 2.95 and 3.01 for housing benefits and income support respectively, implying lower levels of mental well-being. In panel B the summary statistics of air pollutants, income and weather conditions are presented. As it is observed, the standard deviations among the air pollutants are significantly varied; for this reason the standardized air pollutants are considered in the regression analyses.

In table 2, the correlation matrix between the various pollutants, GHQ measure and benefits is reported. The correlation between nitrogen oxides and carbon monoxide is positive, while ground-level ozone is negative correlated with the other air pollutants examined. The negative correlation between O_3 and the other pollutants induced by seasonal variations in the occurrence of these pollutants, as O_3 is well known as the summer smog and its formation depends on solar radiation and temperature. As it was expected the correlation between air pollutants and the GHQ mental health measure is positive indicating that air pollutants might have a negative effect on psychological health. This will be examined in more details in section 5. Similarly, the correlation between income support and housing benefits and GHQ is negative, indicating that individuals who receive these benefits are more likely to report higher levels of mental health. The correlation between air pollutants and the benefits is positive, showing that individuals who reside in high polluted areas are more likely to receive these benefits. One explanation for this association could be that poorer households are located in more deprived –based on air quality- areas and thus might need additional income support and housing benefits. This can be also seen by the negative relationship between household income and air pollutants, as well as, the negative association between income and benefits, since households belonging in higher income classes either are not eligible or are not in need. As it was expected the association between GHQ and household income is negative indicating higher levels of mental well-being. Lastly, the individuals who receive housing benefit is more likely to receive income support too.

(Insert tables 1-2)

5. Empirical results and discussions

It should be noticed that the main population of interest in this study is the non-movers since it is plausible that the decision to move is correlated with the factors of interest- air pollution, housing benefits and income support. More specifically, income and job status do not remain stable across areas and thus housing benefits and income support will change as well. One possible way is to employ panel data analysis which will eliminate the area fixed effects for the non-movers leading to more robust MWTP. On the other hand, the error term for the movers sample will contain the difference of area fixed effects moving across different areas and locations. Since, this difference may be well correlated with the difference in air pollution levels, as well as, in income and thus in benefits and income support across the two locations, may lead to biased estimates.

In table 3, the adapted Probit FE results are reported. In columns 1 and 2, the estimates for those are eligible for income support and housing benefits are respectively presented. The results confirm the positive effects of household income on mental well-being, indicated by the negative sign. In addition, those who claim the housing benefits and income support are more likely to report better mental well-being levels.

The rest of the estimated coefficients are consistent with previous studies (Clark and Oswald, 1994; Benzeval, 2000; Contoyannis et al., 2004; Levinson, 2012; Giovanis, 2014). More specifically, a quadratic relationship between mental health and age is presented. This indicates that mental health is improved and there is a peak at a certain point on life cycle, where after this point of age is more likely to be associated with the probability occurrence of mental health presence. Air pollution has significant and negative impact on mental well-being with the exception of CO regarding the sample eligible for income support in column (1). In addition, O_3 has the strongest negative impact. Regarding the air pollutants we interpret the coefficients by saying that an increase of a standard deviation in air pollutants, results on average, in an increase of $\lambda'^{*}s_y$ in the dependent variable. The parameter λ' denotes the standardised coefficient of the air pollutant, while s_y denotes the standard deviation of the dependent variable, which is the GHQ. Consistent with the previous literature (Levinson, 2012; Giovanis, 2014) average temperature and the difference between maximum and minimum temperature improve mental health, while wind speed presents negative effects, as it is usually associated with low temperature and cold days. The results show that precipitation is insignificant. Household size is insignificant, while the non-smokers are more likely to report better levels of mental well-being. However, this does not imply any causality, as for people who suffer from mental health problems, it may be more likely to smoke. Regarding the job status, there is no difference between employed and self-employed people, while those who are unemployed report lower levels of mental well-being, as it was expected. Regarding the retired individuals it is more likely to report significant lower levels of well-being when the sample of people who eligible for housing benefits is used for the analysis. This can be explained due the old age of these individual and who are more likely to face problems with good housing quality and housing payments. Concerning the marital status and education level, divorced and widowed individuals report significantly lower levels of mental well-being, as these individuals are less able to afford living costs and those with lower education level might earn less income. Finally, there is no difference between the mental well-being of out-rightly house owners and the mental well-being of house owners with mortgage, but the individuals who reside in rented house report lower levels of well-being.

Furthermore, the results of the Bayesian Network are presented in figure 3 and in columns (3)-(4) of table 3 the BN and DAG estimates are reported. In figure 3, it is observed that there is a direct causal effect from income support and household income on GHQ. Also figure 3 shows the *parents*, *children*, *ancestors* and *descendants*. Taking for example income support, there are three *parents*, the job status, education, household income and marital status (*jbstat*, *educ, house_income* and *mastat* respectively in figure 3). This can be explained by various factors, which are not explicitly exploited here. For example, job status can cause income support depending on whether the individual is employed part time or full time of whether is unemployed, determining this way the eligibility on income support. Causal paths between weather conditions and air pollutants are observed confirming the natural properties of air pollutants (Harrison, 2001), as air pollutants are correlated and are dependent on weather conditions. Moreover, ozone is dependent on both NO_X (nitro in the graph) and CO, as well as, on temperature (Harrison, 2001).

Finally, a causal one-direction path from household income to well-being GHQ is observed, which allows us to calculate robust MWTP values. Thus, in order to estimate the causal effect of household income on GHQ based on relation (5) and the back-door criterion, it should be the regression of household income and its parents- household size, weather conditions, CO, job status and education level. Therefore, BN can be a useful graphical tool which allows us which control variables should be included in the regression analysis, avoiding selection and over-control bias and considering confounding. For instance conditioning also to house tenure or income support which is a descendant of household income will distort the income effect, since it creates a selection bias. Regarding O_3 and NO_X it should be regressed considering also the other air pollutants and weather conditions, while concerning CO only the weather factors are important. In addition, a causal path from NO_X and weather variables to household income is observed, which can be explained by the productivity and educational outcomes. There has been a long literature on exploring the effects of air pollutants on cognitive performance, educational outcomes, productivity and income. Similarly the effects of weather or air pollution on education and job status, can be explained through the health status channel, since job status and education may be dependent on the health status of people and whether are able to work full time, be productive or whether are capable to participate in the labour market due to health problems (Gilliland et al., 2001; Mohai et al., 2011; Ponce, 2012). However, these effects are out of the current study's scope, as well as, weather factors are additionally considered. The causal effect of the social benefits examined and the income are even stronger, indicating that their causal effects are underestimated based on the previous estimates. In addition, the estimated coefficients of the air pollutants remain almost the same indicating that the short time frame and their assignment and mapping on highly disaggregated spatial level is proper and exogenous. Furthermore, the effect of CO on GHQ is significant, while it was found insignificant in the case of the fixed effects regressions in column (1). This again is explained by the fact that the regression is conditioning on education level and other variables which are descendants of CO and block the causal path, as there is no also indirect effect on GHQ. Moreover, the effect of income support on GHQ should include itself and its parents, such as the household income, but not tenure, since both household income and income support cause tenure and this will lead to selection bias. Thus, if we are interest on the effects of tenure on GHQ then the income can be included, but in this case the effect of tenure on GHQ will be explored. Therefore, BN provides us with a graphical representation of the associations among variables, where in some cases are dependent and in other cases become independent. In table 4 the p-values of the causal independence tests for the air pollutants, household income and social benefits explored in this study are presented. According to these values the null hypothesis of independence is rejected and thus it is concluded that GHQ is dependent on the air pollutants, income and social benefits examined in this study. In figure 4 the DAG for housing benefits is presented. In this case the relationships remain the same with the exception that the parents of housing benefits differ, as temperature and precipitation are also parents. This can be explained by the fact that housing benefits are given also in cases housing conditions, floods and disasters which are captured by weather conditions.

(Insert tables 3-4)

(Insert figure 3)

Next the MWTP values for a unit reduction in air pollutants are calculated and are reported in table 5. Regarding the adapted fixed effects model and those who are eligible for income support (column 1) are willing to pay more for air pollutants, than those who are eligible for housing benefits (column 1); £1,550 versus £1,100 for O3, £1070 versus £800 for NO_X and £830 versus £430 for CO. However the MWTP for the income support regarding CO is insignificant, as the estimated air pollutants coefficient is insignificant. This is due to that relation (2) becomes smaller because of the higher income effects on well-being for those who are eligible for housing benefits. The MWTP derived by BN (columns 3-4) are lower than the respective ones calculated with the adapted fixed effects model. This is due the fact that the household income effect for this social class of individuals on well-being is significantly more important than the previous estimates shown. This results to lower MWTP values by almost 20-30 per cent. The MWTP values refer to changes in standard deviation. For example, based on the summary statistics in table 1, one standard deviation in O_3 is equal at 17 and its average value is 35 amounting to a change of slightly over 48 per cent. Thus, the MWTP found in table 5 correspond to this percentage change. The percentage change for the NO_X and CO are respectively 53 and 90.

(Insert table 5)

The comparison of MWTP values with previous studies is mixed. For comparison reasons, the MWTP of these studies has been converted into British pounds based on 2010 as reference year. For instance, using a cross-sectional dataset of 54 countries in 1990 and 1995, Welsch (2002) found that the MWTP is equal at; £145 for a one $\mu g/m^3$ increase in Nitrogen Dioxide (NO_2). In another study, Welsch (2006) used the Eurobarometer survey during the period 1990-1997 for 10 European countries and he found that the MWTP is equal at £175 and £460 for a one μg/m^3 increase in Lead (Pb) and NO_2 respectively. MacKerron and Mourato examined NO_2 in London and they found the MWTP equal at £1,550, while Ferreira et al. (2006) explored the effects of Particulate matter (PM_{10}) on life satisfaction in Ireland. The MWTP was found equal at £950.

The results' consistency with previous studies are mixed for the following reasons. Firstly, the sample examined in this study covers only the specific households that are eligible for income support and housing benefits, while the other studies consider the total samples. Secondly, the interest of population as it has been discussed is the non-movers sample, which has been considered only in the study by Luechinger (2009). Thirdly, some of these studies assign the air pollution based on large geographical areas (Welsh, 2002, 2006) or they employ cross-sectional data (Ferreira et al., 2006; 2013; MacKerron and Mourato). In addition, MacKerron and Mourato (2009) explored only London, which can be highly polluted due the high traffic volume. Fourthly, the frequency of the air pollutants in the study by Luechinger (2009) is annual and the air pollution mapping is based on county level. Finally, the majority of the studies explores the MWTP for one unit increase and not a unit change in standard deviation, which the latter is more appropriate since the standard deviation of the air pollutants is significantly different. The results of this study are closer to those found in the studies by Levivson (2012), who used panel data and the studies by MacKerron and Mourato (2009) and Ferreira et al. (2006).

In table 6, the adapted fixed effects for the total sample and movers sample are reported for additional robustness checks. It is clear that income support and the air pollutants, except O_3, are insignificant which estimates can be biased for the reasons discussed before. In addition, it should be noticed that mover and non-movers sample do not sum up to the total sample as there are also other movers, including those who moved into UK from abroad, as well as, other categories, such as missing and dead. These samples are not considered because the estimates can be even more biased, as well as, the individuals' history in those samples are not always observed during the period examined.

(Insert table 6)

In table 7, additional models for the non-movers sample as robustness checks are considered. It should be noticed that the coefficients of BUC and ordered Logit models are not the same and this is due the fact that these models are different. Thus, in order to be comparable with the rest of the models, the MWTP values are compared. It is observed that the MWTP values are lower in all cases than those derived by the adapted Probit FE. More specifically, concerning income support sample, the MWTP calculated based on the BUC and GMM estimates are £1,320-£1,350 for $O_{3,}$ £960 -£990 for NO_X and £770-£790 for CO. Similarly, for housing benefits sample, the MWTP values are £960-£980 for $O_{3,}$ £680 -£700 for NO_X and £420 -£430 for CO. Regarding GMM, the results are robust based on the Sargan statistic, where the null hypothesis of no endogeneity is not rejected, while the null hypothesis of AR(2) is accepted as well.

If the income support and housing benefits amounts are considered they can be compared with the MWTP values. For example a person older than 25 years old is eligible for the weekly amount of income support equal at £3,800, which is significantly higher than the MWTP for improvement on air quality. Thus, policies that reduce the air pollution can create plausible savings on the public finance system as well.

(Insert table 7)

Bayesian networks can have important policy implication, as causal inference has a central role in well-being, including life satisfaction and other measures of well-being, such as leisure and health with various implication to public health, such as the examination of public goods, which is the air quality explored in this study and the effects of income and benefits on well-being. Therefore, the determination that an association is causal indicates the possibility for intervention and thus for policy making and causation can have profound consequences on well-being and public health among other sectors.

However, Bayesian Networks share the same drawbacks with other causal inference approaches, including natural and randomized experiments and instrumental variables. More specifically, natural experiments are not under the control of the investigator and the variation in the level of outcome can vary also in many other ways and some of them can also affect the treatment, even if the *parallel trend* assumption test used in the differences-in-differences models shows the opposite. In addition there might be still problems about unobserved confounders in all these approaches, and the design of the randomized experiments. However, natural experiments is very difficult to be found and meet the above conditions, such as the *parallel trend* assumption, or the suitability of the instrumental variables, where as it has been discussed previously, it is almost impossible to find an instrument which does not determine or is not related to life satisfaction. Thus, BN provide an alternative approach for using observational data, when natural experiments or instrumental variables are difficult to be implemented. In addition, BN can have applications in randomized experiments (Pearl, 2000; Spirtes et al., 2000).

6. Conclusions

This study has used a set of panel micro-data on self-reported mental well-being from the British Household Survey and it examined the causal effects of housing and income support benefits, as well as, income on well-being. Various econometric approaches have been applied for robustness checks.

The importance of this study comes from the fact that the analysis relies on detailed micro-level data and controls additionally for air quality and weather conditions, using highly spatially disaggregated data based on local authority districts, capturing more precise the air pollution effects which are not captured in previous studies. Furthermore, future applications and alternative approaches are suggested, such as the Random Effect Generalized Ordered Probit and Logit models, which account for slope heterogeneity. Furthermore, personality traits and social norms can be considered for future research, as for instance unemployment can be even more hurtful when regional unemployment is considered, and consumer behaviour and preferences can be dependent on social norms (Winkelmann, 2009; Woersdorfer, 2010; Binder and Ward, 2013). The same can also hold for public goods as the air quality.

Finally, BN framework has been proposed, which accounts for confounding and endogenous bias. Therefore, BN is suggested for future research and for applications on causal effects and policies, especially in the cases where natural experiments are very difficult to be applied and instrumental variables are not available, not convincible and which may lead to selection bias. This will help the quest for causality, which is very important for policy design and implications.

Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable comments, suggestions and constructive comments that greatly contributed to the improvement of the quality of this paper. Any remaining errors or omissions remain the responsibility of the authors.

This research has been funded under the Scientific and Technological Research Council of Turkey (TUBITAK) Scientist Support Directorate (BİDEB) - Postdoctoral Research Grants Funding Scheme 2219.

This work was based on data from the British Household Panel Survey, Waves 1-18, 1991-2009: Special Data Service Access, produced by the Institute for Social and Economic Research (ISER) at the University of Essex, sponsored by the Economic and Social Research

Council (ESRC). The data are the copyright of ISER. The use of the data in this work does not imply the endorsement of ISER, ESRC or the Secure Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. University of Essex. Institute for Social and Economic Research, British Household Panel Survey, Waves 1-18, 1991-2009: Special Data Service Access, Colchester, Essex: UK Data Archive [distributor], August 2010. SN: 6340.

References

Baetschmann, G., Staub, K.E., Winkelmann, R. (2015): 'Consistent Estimation of the Fixed Effects Ordered Logit Model', *Journal of the Royal Statistical Society: Series A*, 178(3), pp. 685-703

Bardasi, E., Jenkins, S.P., and Rigg, J.A. (1999): 'Documentation for derived current and annual net household income variables, BHPS Waves 1–7', Working Paper 99–25, Institute for Social and Economic Research, University of Essex.

Benzeval, M., Taylor, J., Judge, K. (2000): 'Evidence on the relationship between low income and poor health: Is the Government doing enough?' *Fiscal Studies*, 21(3), pp. 375-99.

Binder, M. and Ward, F. (2013): 'The Structure of Subjective Well-being: A Vector Autoregressive Approach', *Metroeconomica*, 64(2), pp. 361-400.

Blau, F., Kahn, L. (1997): 'Swimming Upstream: Trends in the Gender Wage Differential in the 1980s', *Journal of Labor Economics*, 15(1), pp. 1-42.

Blundell, R., Bond S. (1998): Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), pp. 115-43.

Bolton, J. (2009): *The use of resources in adult social care: a guide for local authorities*. London, Department of Health

Clark, E.A., Oswald J.A. (1994): 'Unhappiness and Unemployment', *The Economic Journal*, 104(424), pp. 648-59.

Chung, K.F., Zhang, J., Zhong, N. (2011): 'Outdoor air pollution and respiratory health in Asia', *Respirology*, 16(7), 1023-26.

Contoyannis, P., Jones, M.A., Rice, N. (2004): 'The dynamics of health in the British Household Panel Survey', *Journal of Applied Econometrics*, 19(4), pp. 473-503.

Contoyannis, P., Jones, M.A. (2004): 'Socioeconomic status, health and lifestyle', *Journal of Health Economics*. 23(5), pp. 965-95.

Ferreira, S., Moro, M., Clinch, J.P. (2006): ‘Valuing the environment using the life-satisfaction approach’, Planning and Environmental Policy Research Series Working Paper No. 06/03, School of Geography, University College Dublin.
Ferreira, S., Moro, M. (2010): ‘On the use of subjective well-being data for environmental valuation’, *Environmental Resource Economics*, 46(3), pp. 249-73.
Ferreira, S., Akay, A., Brereton, F., Cuñado, J., Martinsson, P., Moro, M., Ningal, T.F. (2013): ‘Life Satisfaction and Air Quality in Europe’, *Ecological Economics*, 88, pp. 1-10
Ferrer-i-Carbonell, A., Frijters, P. (2004): ‘How Important is Methodology for the estimates of the determinants of Happiness?’ *Economic Journal*, 114, pp. 641-59.
Fräßdorf, A., Grabka M.M., Schwarze, J. (2011): ‘The impact of household capital income on income inequality—a factor decomposition analysis for the UK, Germany and the USA’, *The Journal of Income Inequality*, 9(1), pp. 35-56.
Giovanis, E., (2014): ‘Relationship between Well-Being and Recycling Rates: Evidence from Life Satisfaction Approach in Britain’, *Journal of Environmental Economics and Policy*, 3(2), pp. 201-14.
Gonzalez-Barcala, F.J., Pertega, S. Garnelo, L. Castro, T.P. Sampedro, M. Lastres, J.S., san Jose Gonzalez, M.A., Bamonde, L. Valdes, L., Carreira, J-M., Silvarrey, A.L. (2013): ‘Truck traffic related air pollution associated with asthma symptoms in young boys: a cross-sectional study’, *Public Health*, 127(3), 275-81.
Harrison, R.M. (2001): *Pollution: Causes, Effects and Control*. 4th Edition, The Royal Society of Chemistry, UK.
Johnson, R., Griffiths, C., Nottingham, T. (2006): *At home? Mental health issues arising in social housing*. National Institute for Mental Health in England.
Kalisch, M., Buhlmann, B. (2007): ‘Estimating high-dimensional directed acyclic graphs with the PC-Algorithm’, *Journal of Machine Learning Research*, 8, 613-36
Kapteyn, A. Smith, P.J., van Soest, A. (2004): ‘Self-reported work disability in the US and The Netherlands’, RAND Labor and Population Working Paper
Katz, L., Autor, D. (1999): ‘Changes in the Wage Structure and Earnings Inequality’, In Ashenfelter O., Card D. (Eds), *Handbook of Labor Economics*, 3A, pp. 1463-555. Elsevier
Kenkel, S.D. (1995): ‘Should you eat breakfast? Estimates form health production functions’, *Health Economics*. 4(1), pp. 15-29.
Kerkhofs, M., Lindeboom, M. (1995): ‘Subjective health measures and state dependent reporting errors’, *Health Economics*. 4(3), pp. 221-35.
King, G., Murray, C.J.L., Salomon, J.A., Tandon, A. (2004): ‘Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research’, *American Politcal Science Review*, 98(1), pp. 191–207.
Levinson, A. (2012): ‘Valuing Public Goods Using Happiness Data: The Case of Air Quality’, *Journal of Public Economics*, 96(9-10), pp. 869-80.
Lindeboom, M., van Doorslaer, E. (2004): ‘Cut-point shift and index shift in self-reported health’, *Journal of Health Economics*, 23(6), pp. 1083-099.
Luechinger, S. (2009): ‘Valuing Air Quality Using the Life Satisfaction Approach’, *The Economic Journal*. 119(536), pp. 482-515.
Mackerron, G., Mourato, S. (2009): ‘Life satisfaction and air quality in London’, *Ecological Economics*, 68(5), pp. 1441-453
Mulligan, C., Rubinstein, Y. (2008): ‘Selection, Investment, and Women’s Relative Wages Over Time’, *Quarterly Journal of Economics*, 123(3), pp. 1061–110.
Neapolitan, R.E. (2003): *Learning Bayesian Networks*. Prentice Hall Series in Artificial Intelligence: Pearson

Patankar, A.M., Trivedi, P.L. (2011): ‘Monetary burden of health impacts of air pollution in Mumbai, India: Implications for public health policy’, *Public Health*, 125(3), pp. 157-64

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems*. San Mateo, California, Morgan Kaufmann.

Pearl, J. (2000) *Causality. Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.

Pearl, J. (2009): ‘Causal inference in statistics: An overview’, *Statistical Surveys*, 3, pp. 96–146.

Pischke, J.S. (2011): ‘Money and Happiness: Evidence from the Industry Wage Structure’, Discussion Paper No. 5705, IZA

Spirtes, P., Glymour, C., Scheines, R. (2000): *Causation, Prediction, and Search*. 2nd ed, MIT Press, Cambridge, MA.

Stutzer, A., Frey, B. (2012): Recent developments in the economics of happiness: A selective overview. Discussion Paper No 7978. The Institute for the Study of Labour (IZA), Bonn.

Taylor, M.C., Brice, J., Buck, N., Lane, E.L. (2010): ‘British Household Panel Survey User Manual Volume A: Introduction, technical report and appendices’, Colchester: University of Essex.

van Doorslaer, E.J., Jones, A.M. (2003): ‘Inequalities in self-reported health: validation of a new approach to measurement’, *Journal of Health Economics*, 22(1), pp. 61-87.

van Praag, B., Ferrer-i-Carbonell, A. (2004): *Happiness quantified: A satisfaction calculus approach*. Oxford: Oxford University Press.

van Praag, B., Ferrer-i-Carbonell, A. (2006): ‘An almost integration-free approach to ordered response models’, Tinbergen Institute Discussion Paper TI 2006-047/3.

Welsch, H. (2002): ‘Preferences over prosperity and pollution: Environmental valuation based on happiness surveys’, *Kyklos*, 55(4), pp. 473–494.

Welsch, H. (2006): ‘Environment and happiness: Valuation of air pollution using life satisfaction data’, *Ecological Economics*, 58(4), pp. 801–813.

Winkelmann, R. (2009): Unemployment, Social Capital, and Subjective Well-Being’, *Journal of Happiness Studies*, 10(4), pp. 421-30

Woersdorfer, J.S. (2010): ‘When Do Social Norms Replace Status-Seeking Consumption? An Application To The Consumption Of Cleanliness’, *Metroeconomica*, 61(1), pp. 35-67.

Figure 1. An example of a Directed Acyclic Graph (DAG)

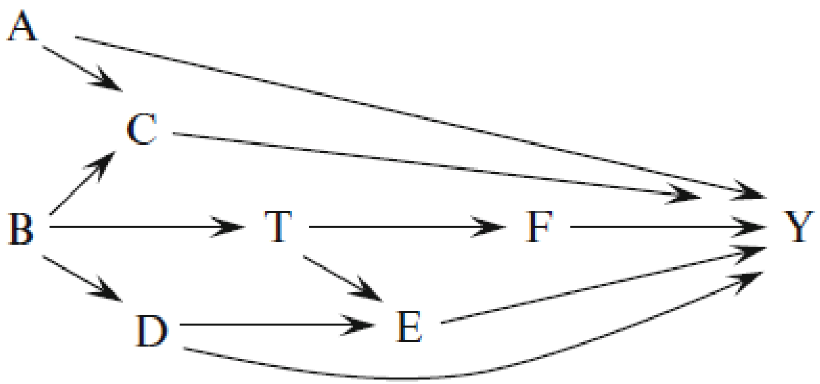

Figure 2. PC algorithm for the estimated DAG

Step 1:
Start with the complete undirected graph, $C^{\sim}$ with vertices $V = X_1, \ldots, X_p$. Then:
Step 2:
Set $l = -1$ and $C = C^{\sim}$
Step 3:
Increase l by one. For all pairs of adjacent nodes:

- Check for conditional independence
- Remove edge (X_i, X_j) if $X_i \perp\perp X_j|rest$

Step 4:
Repeat step 2 until $l = m$ or until each node has fewer than $l - 1$ neighbours
And let mr each $\in$ max l, m denote the stopping level of the algorithm and q be the maximum number of neighbours

In plain words the above pseuso-code of the PC algorithm works on the following simple steps.

- For each X and Y, see if X $\perp$Y; if so, remove their edge.
- For each X and Y which are still connected, and add third variable Z1, see if X $\perp$ Y|Z1; if so, remove the edge between X and Y.
- For each X and Y which are still connected, and add third and fourth variables
- Z1 and Z2, see if X $\perp$ Y|Z1,Z2; if so, remove their edge.

For each X and Y which are still connected, see if X $\perp$ Y| all the p − 2 other variables; if so, remove, their edge
In more details it will be:
Step 1. Form the complete undirected graph G on the set of variables V;
Step 2. For each pair of variables X and Y that are adjacent in the current G such that adj(G, X)\{Y} or adj(G, Y)\{X} has at least n elements, check through the subsets of adj(G, X)\{Y} and the subsets of adj(G, Y)\{X}that have exactly n variables. If a subset S is found conditional on which X and Y are independent, remove the edge between X and Y in U, and record S as separation set- Sepset(X, Y) and repeat until for each ordered pair of adjacent variables X and Y , adj(G, X)\{Y} has less than n elements.
Step 3. Let P be the graph resulting from step 2. For each unshielded triple {A, B, C} in P, orient it as $A \rightarrow B \leftarrow C$ iff. B is not in Sepset(A, C).
Step 4. Execute the following orientation rules until none of them applies:
a If $A \rightarrow B - C$, A and C are not adjacent, orient as $B \rightarrow C$.
b If $A \rightarrow B \rightarrow C$ and $A - C$, orient as $A \rightarrow C$.
c If $A \rightarrow B \leftarrow C$, $A–D - C$, $D - B$, and A and C are not adjacent, orient $D - B$ as $D \rightarrow B$.

Table 1. Summary statistics of income and air pollutants

Variables	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*
Panel A:GHQ				
GHQ	1.9044	2.963	0	12
GHQ (receive the housing benefits)	2.5542	3.122	0	12
GHQ (no-receive the housing benefits)	2.9525	3.604	0	12
GHQ (receive the income support)	2.7650	3.022	0	12
GHQ (no-receive the income support)	3.0165	3.778	0	12
Panel B: Continuous variable				
Household income	1,159.55	980.563	0.0	31,635.07
Ozone (O_3)	35.314	17.357	0.5	124
Nitrogen Oxides (NO_X)	68.747	36.366	8.031	1,780
Carbon Monoxide (CO)	0.418	0.375	0.0	6.7
Average temperature	50.368	7.342	13	81.4
Wind speed	8.374	4.037	0.0	35.2
Precipitation	3.531	1.587	0.69	6.800
Minimum Temperature	44.593	4.022	31.385	53.206
Maximum Temperature	55.725	3.947	41.806	63.667

* The air pollutants are measured in micrograms per cubic meter ($\mu g/m^3$)

Table 2. Correlation between Air Pollutants, Social Benefits and GHQ Well-Being Measure

	Ground-Level Ozone	Nitrogen Oxides	Carbon Monoxide	Income Support Benefit	Housing Benefit	GHQ Caseness Scores
Nitrogen Oxides	-0.5204*** (0.000)					
Carbon Monoxide	-0.0042*** (0.000)	0.2676*** (0.000)				
Income Support Benefit	0.0055** (0.0142)	0.0066*** (0.0022)	0.0107*** (0.000)			
Housing Benefit	0.0071 *** (0.0018)	0.0008 (0.7238)	0.0070*** (0.0010)	0.4161*** (0.000)		
GHQ Caseness Scores	0.0266*** (0.000)	0.0097*** (0.0003)	0.0088** (0.0146)	-0.1203*** (0.000)	-0.1043*** (0.000)	
Household Income	-0.0126*** (0.000)	-0.0451*** (0.000)	-0.0144*** (0.000)	-0.2385*** (0.000)	-0.2272*** (0.0139)	-0.0774*** (0.000)

p-values are reported between brackets, *** and ** indicate significance at 1% and 5% level.

Table 3. Adapted Probit Fixed Effects and BN Estimates

Model	**Adapted Probit FE (1)**	**Adapted Probit FE (2)**	**BN (1)**	**BN (2)**
Household Income	-0.0158**	-0.0384***	-0.0317**	0.0535***
	(0.0061)	(0.0108)	(0.0147)	(0.0157)
Income Support Benefit	-0.0269*		-0.0501**	
	(0.0142)		(0.0237)	
Housing Benefit		-0.0473***		-0.0725***
		(0.0191)		(0.0224)
O_3	0.0052**	0.0049***	0.0055***	0.0051**
	(0.0022)	(0.0010)	(0.0019)	(0.0019)
NO_X	0.0036**	0.0035**	0.0039***	0.0035***
	(0.0015)	(0.0017)	(0.0011)	(0.0011)
CO	0.0028	0.0019**	0.0025*	0.0020*
	(0.0018)	(0.0008)	(0.0013)	(0.0011)
Age	-0.0202**	-0.0452**	-0.0273***	-0.0345***
	(0.0085)	(0.0211)	(0.0074)	(0.0084)
Age Square	0.0002**	0.0005***	0.0003***	0.0004***
	(2.2e-0.4)	(4.4e-0.5)	(2.7e-0.5)	(2.5e-0.5)
Average Temperature	-0.0025*	-0.0025**	-0.0028***	-0.0023**
	(0.0013)	(0.0012)	(0.0008)	(0.0010)
Maximum-Minimum Temperature	-0.0018***	-0.0013*	-0.0031**	-0.0031**
	(0.0006)	(0.0006)	(0.0014)	(0.0014)
Wind Speed	0.0016***	0.0012*	0.0028**	0.0011
	(0.0007)	(0.0007)	(0.0012)	(0.0007)
Precipitation	0.0160	0.0033	0.0142***	0.0115**
	(0.0183)	(0.0021)	(0.0018)	(0.0048)
Household size	-0.0177	-0.0028	-0.0018	-0.0066
	(0.0270)	(0.0144)	(0.0016)	(0.0121)
Job Status (ref=self-employed)				
Job Status (Unemployed)	0.2667**	0.4105***	0.2273***	0.4358***
	(0.1111)	(0.0438)	(0.0874)	(0.0165)
Job Status (Employed)	0.0097	0.0287	0.0104	0.0287
	(0.0072)	(0.0398)	(0.0085)	(0.0398)
Job Status (Retired)	0.1201	0.1394***	0.1275***	0.0705***
	(0.1148)	(0.0414)	(0.0351)	(0.0117)
Marital Status (ref=married)				
Marital Status (Living as couple)	0.1471	0.0155	0.1812	0.0110
	(0.1159)	(0.0522)	(0.1231)	(0.0265)
Marital Status (Widowed)	0.6959***	0.2231***	0.6271***	0.1829***
	(0.1389)	(0.0328)	(0.1149)	(0.0203)
Marital Status (Divorced)	0.4646***	0.0964***	0.4693***	0.0524***
	(0.1759)	(0.0443)	(0.1809)	(0.0211)
Tenure (ref=owned outright)				
Tenure house (Owned with mortgage)	0.0276	0.0166	0.0512**	0.0351***
	(0.0493)	(0.0428)	(0.0224)	(0.0074)
Tenure house (Rented)	0.3077*	0.0822*	0.2158***	0.1115***
	(0.1691)	(0.0477)	(0.0089)	(0.0057)
Education (ref=Higher degree)				
Education Level (First Degree)	-0.0613	0.6174	-0.0587	-0.0887
	(0.0370)	(0.5392)	(0.0375)	(0.4295)
Education Level (Teaching, HNC)	0.0658	0.7197	0.0738	0.0639
	(0.0733)	(0.5818)	(0.0584)	(0.0511)
Education Level (A Level)	-0.0159	0.0774**	0.0400**	0.0852**
	(0.0193)	(0.0353)	(0.0184)	(0.0414)
No obs.	7,848	26,539	7,848	26,539
R square	0.3564	0.3912	0.3446	0.3888

Standard errors between brackets, clustered standard errors on wave area specific trends
***, ** and * indicate significance at 1%, 5% and 10% level

Table 4. P-values for Causal Effects Tests

Associations	P-values
O_3 causes GHQ given temperature, difference in temperature, CO, and NO_X	0.0059
NO_X causes GHQ given CO and precipitation	0.0308
CO causes GHQ given precipitation and temperature	0.0070
Household Income causes GHQ given job status, education level, household size, CO and weather factors.	0.0000
Income support causes GHQ given household income job status, education, marital status	0.0000
Housing benefit causes GHQ given household income, job status, education, temperature and precipitation	0.0000

Table 5. MWTP estimates

	Income support FE	Housing FE	Income support BN	Housing BN
MWTP for a standard deviation reduction in O_3 per year	£1,550	£1,100	£1,100	£730
MWTP for a standard deviation reduction in NO_X per year	£1,070	£800	£710	£500
MWTP for a standard deviation reduction in CO per year	£830	£430	£460	£300

Table 6. Adapted Probit Fixed Effects Estimates for the total sample and movers

Model	Total Sample		Movers	
Household Income	-0.0203***	-0.0299***	-0.0191*	-0.0233**
	(0.0063)	(0.0084)	(0.0104)	(0.0112)
Income Support Benefit	-0.0158**		-0.0082	
	(0.0064)		(0.0102)	
Housing Benefit		-0.0416**		-0.0267*
		(0.0189)		(0.0143)
O_3	0.0086**	0.0089**	0.0044*	0.0041*
	(0.0042)	(0.0043)	(0.0024)	(0.0022)
NO_X	0.0068**	0.0067**	0.0062	0.0058
	(0.0032)	(0.0032)	(0.0056)	(0.0045)
CO	0.0044	0.0041	0.0052	0.0054
	(0.0026)	(0.0025)	(0.0074)	(0.0079)
No obs.	13,313	37,714	4,464	8,473
MWTP for a standard deviation reduction in O_3 per year	£1,404	£1,215	£862	£709
MWTP for a standard deviation reduction in NO_X per year	£1,110	£984	£1,215	£1,003
MWTP for a standard deviation reduction in CO per year	£718	£574	£1,019	£934
R square	0.4566	0.4890	0.3007	0.3214

Standard errors between brackets, clustered standard errors on wave area specific trends
***, ** and * indicate significance at 1%, 5% and 10% level

Table 7. Robustness checks GHQ Regressions

Model	**BUC (1)**	**BUC (2)**	**Ordered Logit RE (3)**	**Ordered Logit RE (4)**	**GMM System (5)**	**GMM System (6)**
GHQ one lag					0.0905***	0.1585***
					(0.0172)	(0.0087)
Household Income	-0.0580***	-0.0715**	-0.0609***	- 0.0758***	-0.0227**	-0.0405***
	(0.0249)	(0.0291)	(0.0249)	(0.0254)	(0.0110)	(0.0082)
Income Support Benefit	-0.0516**		-0.0539**		-0.0337**	
	(0.0226)		(0.0259)		(0.0158)	
Housing Benefit		-0.0720**		-0.0605***		-0.0521**
		(0.0296)		(0.0132)		(0.0259)
O_3	0.0077**	0.0073***	0.0074**	0.0065**	0.0054**	0.0050**
	(0.0036)	(0.0010)	(0.0034)	(0.0028)	(0.0029)	(0.0023)
NO_X	0.0058**	0.0055**	0.0053**	0.0049**	0.0037***	0.0034**
	(0.0017)	(0.0017)	(0.0022)	(0.0016)	(0.0009)	(0.0016)
CO	0.0045	0.0033*	0.0032	0.0028**	0.0031	0.0021*
	(0.0033)	(0.0017)	(0.0044)	(0.0013)	(0.0028)	(0.0011)
No obs.	7,762	26,482	7,837	26,531	7,372	21,962
MWTP for a standard deviation reduction in O_3 per year	£1,350	£980	£1,280	£930	£1,320	£960
MWTP for a standard deviation reduction in NO_X per year	£990	£700	£870	£620	£960	£680
MWTP for a standard deviation reduction in CO per year	£770	£420	£710	£360	£790	£430
Wald Chi-square	559.85	1,458.24	730.15	1,639.45		
	[0.000]	[0.000]	[0.000]	[0.000]		
Wald Statistic					699.34	3,836.41
					[0.000]	[0.000]
P-value for Sargan Statistic endogeneity					32.29	27.51
					[0.472]	[0.542]
P-value for Arellano-Bond test for AR(2)					0.59	1.25
					[0.556]	[0.316]

Standard errors between brackets, p-values between square brackets, ***, ** and * indicate significance at 1%, 5% and 10% level

Figure 3. Estimated DAG for income support

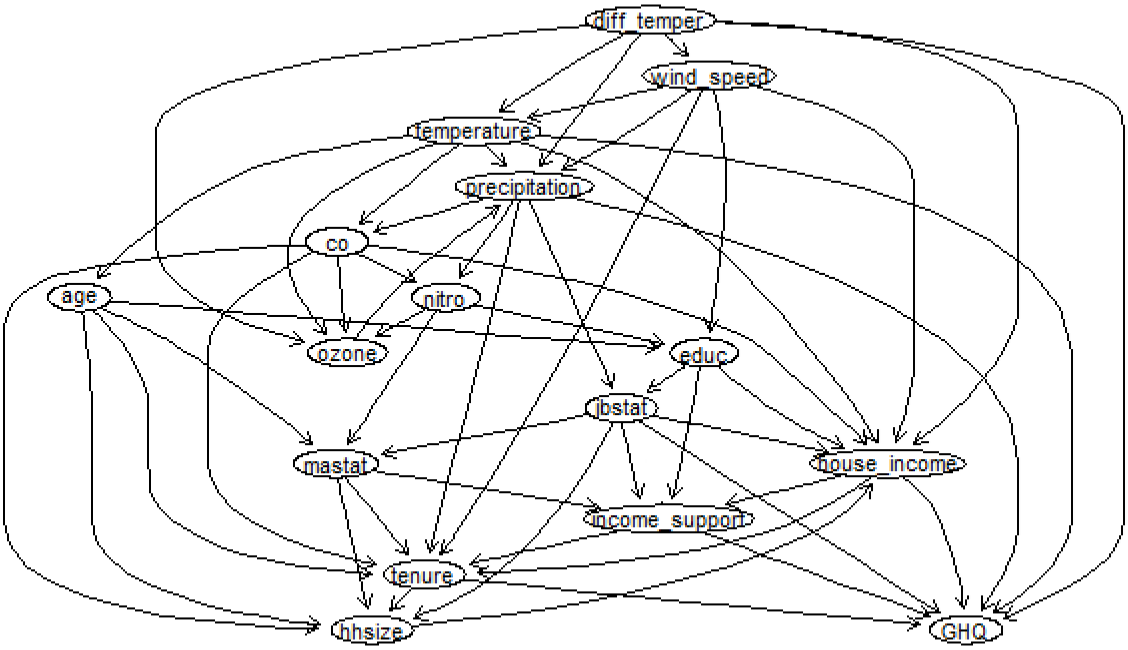

Figure 4. Estimated DAG for housing benefits