Science of the Total Environment 630 (2018) 1269–1282

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Contingent valuation of health and mood impacts of $PM_{2.5}$ in Beijing, China

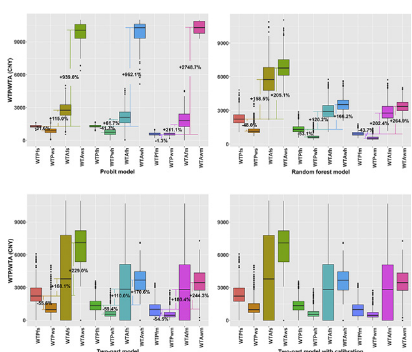
Hao Yin [a,b], Massimo Pizzol [b], Jette Bredahl Jacobsen [c], Linyu Xu [a,*]

[a] State Key Laboratory of Environmental Simulation and Pollution Control, School of Environment, Beijing Normal University, No. 19, Xinjiekouwai Street, Haidian District, Beijing 100875, China
[b] Department of Planning, Danish Centre for Environmental Assessment, Aalborg University, Rendsburggade 14, 9000 Aalborg, Denmark
[c] University of Copenhagen, Department of Food and Resource Economics and Centre for Macroecology, Evolution and Climate, Rolighedsvej 23, 1959 Frederiksberg C, Denmark

HIGHLIGHTS

- The perceived welfare loss caused by $PM_{2.5}$ in Beijing are CNY 2286.1/person/year.
- Both health and mood impacts are investigated via contingent valuation.
- "Willingness to pay" (WTP) and "willingness to accept" (WTA) formats are applied.
- Both face-to-face and web-based survey modes are used.
- WTP/WTA estimates from different models, including Random Forest, are compared.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 16 October 2017
Received in revised form 31 January 2018
Accepted 23 February 2018
Available online xxxx

Editor: SCOTT SHERIDAN

Keywords:
$PM_{2.5}$, welfare loss
WTP/WTA
Health impacts
Mood impacts
Random forest

ABSTRACT

Air pollution from $PM_{2.5}$ affects many cities worldwide, causing both health impacts and mood depression. One of the obstacles to implementing environmental regulations for $PM_{2.5}$ reduction is that there are limited studies of $PM_{2.5}$ welfare loss and few investigations of mood depression caused by $PM_{2.5}$. This article describes a survey study conducted in Beijing, China to estimate the welfare loss due to $PM_{2.5}$. In total, 1709 participants completed either a face-to-face or online survey. A contingent valuation method was applied to elicit people's willingness to pay to avoid $PM_{2.5}$ pollution and willingness to accept a compensation for such pollution. The payment/compensation was evaluated for two outcome variables: perceived health impacts and mood depression caused by $PM_{2.5}$ pollution. This is one of few papers that explicitly studies the effects of $PM_{2.5}$ on subjective well-being, and to the authors' knowledge, the first to estimate welfare loss from $PM_{2.5}$ using a random forest model. Compared to the standard Turnbull, probit, and two-part models, the random forest model gave the best fit to the data, suggesting that this may be a useful tool for future studies too. The welfare loss due to health impacts and mood depression is CNY 1388.4/person/year and CNY 897.7/person/year respectively, indicating that the public attaches great

Abbreviations: WTP_{fh}, willingness to pay to avoid health impacts caused by $PM_{2.5}$ in face-to-face survey; WTP_{fm}, willingness to pay to avoid mood depression caused by $PM_{2.5}$ in face-to-face survey; WTP_{fs}, the sum of willingness to pay to avoid health impacts and mood depression caused by $PM_{2.5}$ in face-to-face survey; WTA_{fh}, willingness to accept to avoid health impacts caused by $PM_{2.5}$ in face-to-face survey; WTA_{fm}, willingness to accept to avoid mood depression caused by $PM_{2.5}$ in face-to-face survey; WTA_{fs}, the sum of willingness to accept to avoid health impacts and mood depression caused by $PM_{2.5}$ in face-to-face survey; WTP_{wh}, willingness to pay to avoid health impacts caused by $PM_{2.5}$ in web-based survey; WTP_{wm}, willingness to pay to avoid mood depression caused by $PM_{2.5}$ in web-based survey; WTP_{ws}, the sum of willingness to pay to avoid health impacts and mood depression caused by $PM_{2.5}$ in web-based survey; WTA_{wh}, willingness to accept to avoid health impacts caused by $PM_{2.5}$ in web-based survey; WTA_{wm}, willingness to accept to avoid mood depression caused by $PM_{2.5}$ in web-based survey; WTA_{ws}, the sum of willingness to accept to avoid health impacts and mood depression caused by $PM_{2.5}$ in web-based survey; WTP_f, willingness to pay survey in face-to-face format; WTP_w, willingness to pay survey in web-based format; WTA_f, willingness to accept survey in face-to-face format; WTA_w, willingness to accept survey in web-based format; FTF, face-to-face survey; WB, web-based survey.

* Corresponding author.
E-mail address: xly@bnu.edu.cn (L. Xu).

https://doi.org/10.1016/j.scitotenv.2018.02.275

importance to mood, feelings and happiness. The study provides scientific support to the development of economic policy instruments for $PM_{2.5}$ control in China.

1. Introduction

Pollution from particulate matter with an aerodynamic diameter equal to or smaller than 2.5 μm ($PM_{2.5}$) generates substantial negative health impacts and social concern in China. $PM_{2.5}$ pollution affects many major Chinese cities in recent years (Guan et al., 2014; Shen et al., 2014). In Beijing, the average annual $PM_{2.5}$ concentration level in 2015 was 80.6 μg/m^3 (BMEPB, 2016), which is more than two times higher than the China National Ambient Air Quality Standard for annual $PM_{2.5}$ concentration (35 μg/m^3) (PRCMEP, 2012).

Several studies highlight how the health impacts due to $PM_{2.5}$ pollution ultimately lead to economic loss and represent a cost for society (Wu et al., 2017; Xie et al., 2016b; Yin et al., 2017). Estimates of these externalities or so-called "external costs" of $PM_{2.5}$ are therefore a valuable tool in the design of economic policies to combat air pollution, such as green taxes or "cap and trade" systems (Hoveidi et al., 2013). The key element of studies that estimate the external costs of air pollution is the valuation of people's willingness to pay to avoid the air pollution damages (Daly, 1992). As the most well-known stated preference technique, contingent valuation is widely used in valuing environmental amenities and damages (Freeman III et al., 2014). Contingent valuation studies can be used to estimate the willingness to pay (WTP) for reducing pollution impacts (Filippini and Martínez-Cruz, 2016; Istamto et al., 2014b; Wang et al., 2015) or the willingness to accept (WTA) them (Breffle et al., 2015).

Contingent valuation studies on air quality improvement and health damage assessment have been conducted in many areas worldwide (Chattopadhyay, 1999; Chestnut et al., 1997; Desaigues et al., 2011; Lee et al., 2011; Navrud, 2001). In China, empirical contingent valuation studies on air pollution have been implemented in several cities, including Beijing (Wang et al., 2006), Ji'nan (Wang and Zhang, 2009), Chongqing (Wang and Mullahy, 2006) and Anqing (Hammitt and Zhou, 2006). One study in Shanghai showed that parents' WTP for air quality improvement to reduce their children's respiratory diseases was between USD $68 to USD $80 (Wang et al., 2015). According to another recent contingent valuation study, the WTP for smog reduction is CNY 380/year (around USD $56) (Sun et al., 2016a). A value of a statistical life of $34458 is estimated from the WTP for reducing health risks caused by air pollution in Chongqing (Wang and Mullahy, 2006). Other studies elicit the WTP for different health impacts caused by air pollution, such as a cold, lower respiratory tract infection, and chronic bronchitis (Alberini et al., 1997; Hammitt and Zhou, 2006). Despite the many contingent valuation studies that estimate the WTP to reduce overall air pollution or smog (Desaigues et al., 2011; Sun et al., 2016a; Sun et al., 2016b; Wang et al., 2015), only a few studies estimate specifically the WTP for reducing $PM_{2.5}$ pollution (Lee et al., 2011; Wei and Wu, 2016). Studying the welfare consequence of a single type of pollution control like $PM_{2.5}$ is important for three reasons: 1) $PM_{2.5}$ is one of the primary air pollutants in many cities in China (PRCMEP, 2017); 2) $PM_{2.5}$ has more serious health and mood influence than coarse particles ($PM_{2.5-10}$); 3) the policy and technical measures to avoid pollution of small particles are different from those to avoid pollution from large particles.

High $PM_{2.5}$ pollution levels do not only impair people's physical health status but also their psychological and mental health (Evans et al., 1988; Zijlema et al., 2016). First of all, $PM_{2.5}$ reduces visibility as the wavelength of visible light falls in the similar range of the $PM_{2.5}$ size (Liu et al., 2014; Sisler and Malm, 1994). The pollution results in the decrease of sunny days. Previous studies show that sunshine hours positively increase mood scores - a criterion for the ranking the mood conditions (the higher a person's mood score is, the happier is this person) (Sanders and Brizzolara, 1982). Clear skies are also a highly prized human amenity (Watson and Chow, 1994). Thus, $PM_{2.5}$ decreases individual's aesthetic enjoyment. It is then reasonable to think that the reduction in atmospheric visibility due to $PM_{2.5}$ pollution may cause a decrease of mood conditions. Secondly, epidemiological studies report that $PM_{2.5}$ pollution is related to anxiety symptoms (Power et al., 2015; Zijlema et al., 2016), which are often comorbid with depression (Lamers et al., 2011). Lastly, toxicological studies indicate that the oxidative stress and inflammation induced by $PM_{2.5}$ pollution impairs brain, cognitive and neurological functioning psychological condition (Calderón-Garcidueñas et al., 2015; Peters et al., 2006). Though there are limited studies reporting on the quantitative correlation between $PM_{2.5}$ exposure and mood impacts, many studies provide evidence of its influence on individual's mood and neuropsychological conditions (Guxens and Sunyer Deu, 2012; Suades-González et al., 2015). It is therefore imperative to quantify the perceived mood impacts caused by $PM_{2.5}$. However, previous contingent valuation studies of $PM_{2.5}$ have focused primarily on its impact on physical health without addressing its effect on subjective well-being.

The choice of using WTP or WTA for monetary valuation via contingent valuation studies depends on the implicit assumptions of the 'property right' ascribed to the status quo or to the post-policy situation (Pearce, 2002). As $PM_{2.5}$ pollution is a public good (bad), these property rights are unclear, and people may have heterogeneous perceptions of it (Hanley et al., 2009). Previous studies have addressed the problem of unclear property rights by estimating both a measure of WTP for e.g. reducing pollutions levels and a WTA for bearing the pollution. This allows researchers to test the discrepancy between the two measurements (Sayman and Öncüler, 2005).

In contingent valuation studies, face-to-face (FTF) surveys have for a long time been considered the best data collection method (Arrow et al., 1993), though opinions are changing (Johnston et al., 2017). The use of web-based surveys (WB) via the internet is increasing due to the low cost and quick responses that this format allows (Nielsen, 2011). WB surveys can also ensure respondents privacy (Nielsen, 2011) and, given that it is often difficult to ask busy interviewees to participate in surveys, can also be a flexible and convenient way of data collection (Szolnoki and Hoffmann, 2013). The empirical setting in China may make these issues more or less important than seen in the literature from elsewhere in the world, and consequently, we test both surveys modes in this context.

In contingent valuation studies, many classical parametric regression models, such as two-part model, probit model, and logit model, are traditionally applied to quantify the relationship between the explanatory and dependent variables (WTP/WTA), and to calculate the estimates of welfare loss. However, the limitation of parametric regressions is their distributional assumptions. The random forest model introduced by Breiman (2001) is a machine learning algorithm with high prediction accuracy that can also estimate the relative importance of each explanatory variable (Archer and Kimes, 2008). Random forest performs well with large datasets and thousands of input variables (Archer and Kimes, 2008) as well as in cases of a small number of observations (Grömping, 2009; Strobl et al., 2007), highly imbalanced data (Khalilia et al., 2011), or high dimensional datasets with non-linear and complex interactions among variables (Cutler et al., 2007). However, according to the authors' knowledge, there is no contingent valuation study that uses random forest to estimate welfare loss.

In this context, the present study aims at providing monetary estimates of health and mood impacts of $PM_{2.5}$ pollution in Beijing, China based on the contingent valuation method. To address the problems of unclear property rights and sampling issues, the study estimates two different measures of individuals' preferences and elicits these preferences using two different survey methods. Both WTP and WTA are investigated and both FTF and WB data collection is performed. WTP/WTA estimates are calculated with different models to identify how the choice of model affects the results.

This paper is organized as follows: Section 2 describes the contingent valuation study design and the methodology used for data analysis including the econometric models used. Section 3 presents the study results in terms of WTP/WTA measures obtained with FTF and WB surveys respectively. In Section 4, the study results, implications of the results and limitations of the study are critically discussed. Section 5 concludes with an outlook on the applications of the study's findings and on directions for future research.

2. Materials and methods

2.1. Study area and sampling

Beijing, the capital city of China with a population of more than 20 million, is located in northern China and is severely affected by $PM_{2.5}$ pollution. Fig. 1 shows the $PM_{2.5}$ pollution distribution in 16 districts in Beijing. The central and southern part of Beijing suffers the most severe levels of $PM_{2.5}$ pollution. We visited 13 districts for the questionnaire survey to make sure the interviewees are evenly distributed according to the population density distribution in Beijing and to test the potential heterogeneity of WTP/WTA in different districts. The investigated districts include Dongcheng, Xicheng, Haidian, Chaoyang, Fengtai, Shijingshan, Tongzhou, Shunyi, Fangshan, Daxing, Changping, Huairou, and Pinggu districts.

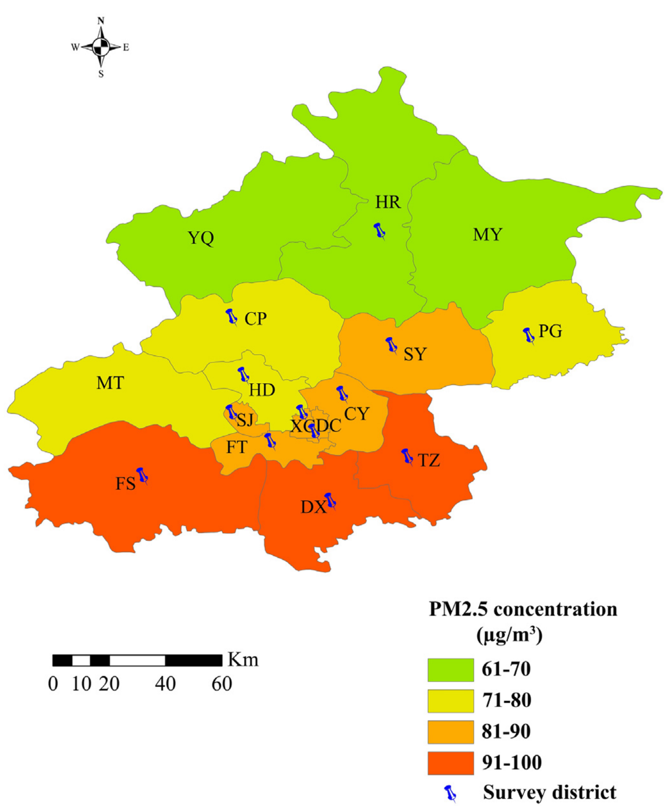


Fig. 1. $PM_{2.5}$ questionnaire sampling sites and concentration in 16 districts of Beijing.

(DC, XC, HD, CY, FT, SJ, MT, TZ, SY, FS, DX, CP, HR, PG, MY and YQ refer to the districts of Dongcheng, Xicheng, Haidian, Chaoyang, Fengtai, Shijingshan, Mentougou, Tongzhou, Shunyi, Fangshan, Daxing, Changping, Huairou, Pinggu, Miyun and Yanqing. The $PM_{2.5}$ data were retrieved from the Beijing Environmental Statement (BMEPB, 2016).

2.2. Questionnaire design

In this study, a payment card was used for WTP/WTA elicitation as proposed by Mitchell and Carson (1981). Payment card can be used when it is difficult for respondents to state a value for certain goods, thereby increasing the response rate. A payment card provides a range of bids and asks respondents to select the highest WTP (smallest WTA) (Cameron and Quiggin, 1994). With the payment card technique, respondents' WTP is higher than the value chosen and lower than the next, and vice versa for WTA. The use of a payment card format assumes that respondents can distinguish between values. Therefore, the payment card should provide noticeable differences between two values. Weber's law describes the difference between physical magnitude and perceived intensity, which allows respondents to discriminate different stimuli (Leshowitz et al., 1968; Panek and Stevens, 1966). With Weber's law, we set the payment card levels which show noticeable differences among those bids (Rowe et al., 1996). As Weber's law is broadly applied for perception discrimination (Deco et al., 2007), we obtained the payment card bid sequence based on the Eq. (1) (Rowe et al., 1996):

$$P_n = P_1 \times (1+k)^{n-1} \tag{1}$$

where P_n refers to the nth payment value, and k is the positive constant. The value of P_1 and k are selected so that P_1 and $P_1 \times (1+k)^{n-1}$ equal the smallest and largest value posted on the payment card respectively (Rowe et al., 1996). The payment card uses a P_1 of "5" and a k value of 0.5. Here, we exclude the "0" bid, and choose "5" as the smallest bid for the rest of P_n calculation.

The questionnaire is designed with the help of a pretest survey in advance of the formal survey. We conducted pretest interviews ($n = 76$) in the Xicheng and Haidian Districts to get a basic knowledge of respondents' reaction and understanding of the questionnaire and their WTP range. The initial bid and scale of the payment card can influence respondents' WTP/WTA (Rowe et al., 1996; Van Exel et al., 2006), thus the pretest questionnaire used an open-ended vehicle to estimate people's willingness to pay/accept, which provided and fixed the payment card setting levels in the latter formal interviews.[1] According to respondents' answers and reactions, we set the final payment card levels, revised and improved the expressions of the questions that were difficult for respondents to answer. For example, some respondents did not fully understand what $PM_{2.5}$ exactly is, thus we added a figure before the questions describing what and how $PM_{2.5}$ damages our health and mood. Further explanation was also provided whenever respondents face difficulties in answering the questions. In the final questionnaire, there are $n = 16$ bid levels in the payment card and an additional open-ended choice at the end of payment card to allow for higher bids. This can avoid the anchoring effect and potential truncated issue to a large extent.

Respondents are split into two groups, either receiving a WTP or a WTA questionnaire. The questionnaire contents are the same for the FTF and the WB surveys. The WTP survey elicits respondents' willingness to pay for $PM_{2.5}$ pollution reduction to the national $PM_{2.5}$ limit (35 μg/m^3) and the WTA survey elicits respondents' willingness to

[1] An open-ended format was not chosen for the final survey because it tends to provide unrealistically high payments (Johnston et al., 2017) Johnston, R.J., Boyle, K., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R., 2017. Contemporary guidance for stated preference studies. J Assoc Environ. Econ., 4, 319–405.

accept a compensation if the status quo $PM_{2.5}$ pollution level is maintained. There are three main reasons why we set the goal of annual $PM_{2.5}$ concentration at this level. First, Ministry of Environmental Protection of China set 35 μg/m^3 as interim $PM_{2.5}$ standard in urban areas (PRCMEP, 2012). Secondly, under the circumstances of current pollution-control techniques and socioeconomic level in China, it is unrealistic to cut down pollution to 10 μg/m^3 or even lower in a short time. Lastly, background $PM_{2.5}$ concentration (arising from natural or transported process) is around 35 μg/m^3 in Beijing (Sun, 2012; Zhao et al., 2013), which indicates it is difficult to reduce annual $PM_{2.5}$ concentration to a lower level. Thus, we select annual concentration of 35 μg/m^3 as $PM_{2.5}$ reduction goal in the questionnaire survey.

In the survey design, we followed the guidelines from the NOAA report (Arrow et al., 1993) and further looked at formulations of more recent work. Most of the cases respondents are provided a scenario description and asked how much they are willing to pay for a certain project for air pollution mitigation (Sun et al., 2016b; Wang et al., 2006) or how much they are willing to pay for extending life expectancy (Istamto et al., 2014a) or avoiding diseases (Wei and Wu, 2017) due to pollution reduction under family or individual income budget (Wang et al., 2015).

Specifically, two hypothetical scenarios are presented to the respondents in this study: 1) "*If providing funds for the measures that you choose above could cut the annual* $PM_{2.5}$ *pollution from around* 80 μg/m^3 *to the China air quality standards (*$PM_{2.5} \leq 35$ μg/m^3*) effectively and avoid the impacts of* $PM_{2.5}$ *on your health and mood to a large extent, would you be willing to pay a fee for this every month? Please be aware that your payment should be within your income budget. All the money you have paid would be guaranteed to devote to the implementation of these measures.*"; 2) "*If the above measures could cut the annual* $PM_{2.5}$ *pollution from around 80 μg/m3 to the China air quality standards (*$PM_{2.5} \leq 35$ μg/m^3*) effectively and avoid the impacts of* $PM_{2.5}$ *on your health and mood to a large extent, but the measures could not be implemented at present, according to the health and mood impacts caused by current* $PM_{2.5}$ *pollution, would you be willing to accept certain amount of economic compensation monthly?*". Overall, the questionnaire is designed to retrieve information about 1) respondents' background in terms of age, education, and place of living; 2) self-reported socio-economic and health status, including individuals' health conditions, expenditures on medical costs, and exposure to $PM_{2.5}$ pollution; 3) self-reported knowledge of $PM_{2.5}$ adverse impacts and preferred $PM_{2.5}$ pollution control measures; 4) WTP or WTA, starting with a screening of whether respondents want to pay/accept at all, followed by a question of how much they are willing to pay, and follow-up questions about the reasons for their choices. Respondents state their subjective valuation on impacts of $PM_{2.5}$ pollution through the corresponding questions. The full original questionnaire in Chinese and its English translation are provided in Appendix.

Before answering the survey questions, all respondents were provided with basic information about $PM_{2.5}$ pollution including $PM_{2.5}$ definition, causes and sources of $PM_{2.5}$ pollution, $PM_{2.5}$ impacts on health and visibility, and $PM_{2.5}$ concentration on the survey date. To make it clear, we insert a figure to explain $PM_{2.5}$ pollution sources, pathways, impacts on health and atmospheric visibility. In particular, trained investigators also give explanation whenever respondents have difficulties in understanding the survey contents. The daily $PM_{2.5}$ concentration on the date of the interview was retrieved from the Municipal Environmental Protection Bureau (BMEPB, 2015) and was coded on each questionnaire. For the WB survey, we created our questionnaire on the professional survey website Music Survey (MusicSurvey, 2015). The service provider facilitates data collection and promotes participation by providing users a reward for completing the questionnaire and for sharing it. In the WB survey, the relevant background information and relevant explanations were provided in detail before participants started answering the questions.

2.3. Data collection

The target population of this survey is permanent residents[2] living in Beijing who are affected by $PM_{2.5}$ pollution. In the formal survey, respondents were selected by approaching individuals in selected parks, commercial districts and recreational places with a large flow of people. The survey sites are popular places distributed in different districts in Beijing, which allowed us to approach people from different districts. We recruited 11 undergraduate students assisting the questionnaire survey. To make sure objective knowledge can be delivered to the respondents, we trained the interviewers in detailed knowledge of PM2.5 pollution, how to approach interviewing, how to introduce our questionnaire survey, and how to explain frequent asked questions about PM2.5 pollution characteristics, sources, pathways and impacts. From April to July 2015, we mainly conducted the survey from 9 a.m. to 8 p.m. on Fridays, weekends and holidays when more people have leisure time in public areas. The questionnaire survey covers 13 out of the 16 districts of Beijing (Fig. 1). In the survey, we randomly select people who stay alone or in a small group taking a rest or having no important things to do at that moment. A stratified sampling approach was applied to choose the respondents in the FTF surveys according to the population density, gender and age structures in Beijing. All FTF questionnaires were administered by trained interviewers who were able to explain to participants the background information and the meaning of the questions and choice options.

For the WB survey, links to the questionnaires along with advertisement of the survey were posted on Weibo (Weibo, 2015) and Wechat (Wechat, 2015), the most widely used social media in China, to invite people to participate and repost the WB questionnaires. After 30 days, we logged in to the account of the survey website to export the survey data.

In total, we investigated 1751 participants among which there are 1709 valid questionnaires including 1189 WTP questionnaires (727 face-to-face and 462 WB questionnaires) and 520 WTA questionnaires (308 face-to-face and 212 WB questionnaires).

2.4. Data analysis

It is common to remove unrealistic and protest answers to the payment questions before estimation. We excluded unrealistic answers with WTP or WTA much higher than CNY 440,000 per year (i.e. the value of statistical life according to the study from (Hammitt and Zhou, 2006)) or WTP higher than the declared income. Furthermore, protest bidders were removed (Morrison et al., 2000) because their responses do not represent their true WTP for $PM_{2.5}$ reduction (Strazzera et al., 2003). Protest bidders were identified as respondents who expressed a zero WTP/WTA and chose one of the following three explanatory reasons among those provided in the questionnaire: '*the hypothetical goal of* $PM_{2.5}$ *reduction is unlikely to be achieved*', '*the costs have already been included in the taxes*', '*the government or the polluters should pay*'.

The investigation of WTP, WTA and all potential explanatory variables is under analysis of four statistical models including Turnbull model, two-part model, probit model and random forest model. These four models are chosen so ensure robustness in the result, and furthermore also because each is able to show different aspects of the WTP and WTA. The Turnbull model is a non-parametric model which does not depend on the assumption of data distribution. Thus, it is insensitive to assumptions of functional form or distributional effects (Haab and McConnell, 1997). The

[2] i.e. who have lived there for more than half a year. Respondents were asked this as a first screening question.

Turnbull estimator can provide a lower-bound WTP and WTA estimate (Ready et al., 2001). After this distribution-free estimation, we apply a two-part model and a probit model to analyze explanatory variables of WTP or WTA. The two-part model handles zero-bidders influence explicitly and combines it with a linear regression (Duan et al., 1983). Both the Turnbull model and the two-part model assume respondents' bids to be an exact value. The probit model instead assumes the answer to a given bid as just a threshold – lower or higher than this and thus more in accordance with the way the bid questions are formulated. Finally, the random forest is applied to test performance of WTP (or WTA) estimation without distribution assumption.

2.4.1. $PM_{2.5}$ pollution welfare measures

To interpret the estimates of WTP and WTA, consider a conceptual model of welfare measures where a change of environmental quality (Q) has impacts on individual utility but little influence on the price of market goods. Welfare changes are measured as the compensating variation (CV) and equivalent variation (EV) based on the following theoretical functions (Maler and Vincent, 2005). For further explanations and interpretations of the CV and EV, see Freeman III et al. (2014). According to the Hicksian welfare theory, WTP and WTA are equal to the compensating variation and equivalent variation, equivalently. Thus, the welfare changes due to the $PM_{2.5}$ pollution are estimated with WTP/WTA through contingent valuation method in this study.

The initial (U^0) and proposed (U^1) utility level are expressed in Eq. (2a) and Eq. (2b)

$$U_i^0 = V\left(P^0, Q^0,\ I_i^0\right) \tag{2a}$$

$$U_i^1 = V\left(P^0, Q^1, I_i^0\right) \tag{2b}$$

where U is the utility function; P, Q and I_i denote the prices of goods, qualities of goods and the individual's income. The superscripts 0 and 1 denote the initial and subsequent states of the relevant parameters. Here we assume the income is restrictive and fixed.

If the environmental quality changes from Q^0 to Q^1, then there should be certain amount of money taken from (given to) the individual i to make him/her as well off as (s)he was before the environmental quality changes happened. This is the compensating variation see Eq. (2c).

$$U_i^0 = V\left(P^0, Q^1,\ I_i^0 - CV\right) = V\left(P^0, Q^0,\ I_i^0\right) \tag{2c}$$

Another measure takes the subsequent situation as the utility baseline. Specifically, a certain amount of money would be given to (taken away from) the individual to make him/her as well off as s/he could have been with the environmental quality change. This is the equivalent variation as expressed in Eq. (2d):

$$U_i^1 = V\left(P^0, Q^1, I_i^0\right) = V\left(P^0, Q^0,\ I_i^0 + EV\right) \tag{2d}$$

According to Eqs. (2c)–(2d), it is straightforward to obtain the functions of CV and EV:

$$CV = e\left(P^0, Q^0, U_i^0\right) - e\left(P^0, Q^1, U_i^0\right) \tag{2e}$$

$$EV = e\left(P^0, Q^0, U_i^1\right) - e\left(P^0, Q^1, U_i^1\right) \tag{2f}$$

2.4.2. Turnbull model

For a non-parametric estimation of WTP/WTA, we applied the Turnbull estimator to analyze the payment card data (Haab and McConnell, 2002). The lower-bound estimate of WTP is obtained with Eq. (3):

$$E_{LB}(\text{WTP}) = \sum_{j=1}^{M} P_k * f_k \tag{3}$$

where E_{LB}(WTP) is the lower bound of expected willingness to pay, M is the number of bids, P_k refers to the kth payment value, and $f_k = T_k/T$ where T is the total number of respondents and T_k is the number of respondents who pick P_k.

According to Eq. (3), we estimate the Turnbull lower bound mean WTP values. If $T_k > T_{k+1}$ then pooling may be applied, but that was not needed in our case. The same approach was used for the WTA data. The 95% confidence intervals were calculated by the bootstrap method (Wang et al., 2006; Wu, 1986) with replacement for 1000 replications.

2.4.3. Two-part model

To take explicitly into account that there are relatively many valid zero-bids, and still rely on a linear model, we used a two-part model (Duan et al., 1983). This model deals with the zero bid responses by separating the decision behavior into two stages: 1) the respondents have positive WTP/WTA; 2) the level of WTP/WTA of the respondents. A probit model was used to determine the probability estimation of positive WTP/WTA, Eq. (4a); and then a linear regression model was applied for the estimation of positive WTP/WTA, Eqs. (4b), (4c).

$$\text{Part I}: Prob(\text{WTP}{>}0) = f(X) \tag{4a}$$

$$\text{Part II}: \text{WTP}(\text{WTP}{>}0) = \alpha + \beta * X + \varepsilon, \varepsilon \sim N\left(0, \sigma^2\right) \tag{4b}$$

$$\text{EWTP} = Prob(\text{WTP}{>}0) \times \text{WTP}(\text{WTP}{>}0) \tag{4c}$$

where f() is the standard normal distribution, β denotes a vector of coefficients and X is a vector of the associated sociodemographic variables, α is a constant term of the linear regression model, ε is an error term assumed to be normally distributed with mean value zero and variance σ^2, EWTP is the expected WTP estimated with the two-part model.

2.4.4. Probit model

A linear regression model does not take into account the intervals used in the payment cards, which assumes that the maximum bid is an exact value. To solve this problem, a binary probit model was also fitted to determine the respondents' WTP/WTA for $PM_{2.5}$ reduction. Each level on the payment card is understood as a bid, to which respondents answer yes (1) or no (0) (Haab and McConnell, 2002). Each questionnaire sample has 17 levels assigned with "1" or "0", which expands the observations in probit model. Consequently, the probability of accepting a given bid can be estimated as Eq. (5a).[3]

Relying on the random utility framework, an individual will answer yes to a given bid, given that the utility of that policy scenario is larger than without it, i.e. if

$$Prob\left(yes_j\right) = Prob\left(u_1\left(y_j - t_j, X_j, \varepsilon_{1j}\right) > u_0\left(y_j, X_j, \varepsilon_{0j}\right)\right) \tag{5a}$$

where u is the utility of the alternative "Yes" (1) or "No" (0), y_j is the income of an individual j, t_j is the payment for the policy or measures for $PM_{2.5}$ pollution control, X_j is a vector of sociodemographic variables, and ε_{0j} and ε_{1j} are error terms.

Assuming a linear utility function, and splitting the utility up into a deterministic utility (v), and a stochastic term (ε), the deterministic

[3] Notation in the following follows Haab and McConnell (2002)

part of the policy/measures may be written as Eq. (5b):

$$v_{1j}\left(y_j - t_j\right) = \alpha_1 X_j + \beta_1\left(y_j - t_j\right) \tag{5b}$$

where α is a vector of parameters to be estimated, and the utility of the status quo as Eq. (5c):

$$v_{0j}\left(y_j\right) = \alpha_0 X_j + \beta_0 y_j \tag{5c}$$

Assuming we are looking at a marginal change, so that the marginal utility of income is constant, the utility difference becomes (Eq. (5d)). Although marginal utility varies among individuals, the income effects are explicitly controlled in the regression model.

$$v_{1j} - v_{0j} = \alpha X_j + \beta t_j \tag{5d}$$

And consequently the probability of answering yes can be written as:

$$Prob\left(yes_j\right) = Prob\left(\alpha X_j + \beta t_j - \varepsilon_j\right) \tag{5e}$$

where $\varepsilon_j = \varepsilon_{1j} - \varepsilon_{0j}$. This may be estimated by a standard probit model.

Solving for t_j, WTP or WTA can be calculated as Eq. (5f):

$$\mathrm{WTP}_j = \frac{\alpha X_j}{\beta} + \frac{\varepsilon_j}{\beta} \tag{5f}$$

We used the Krinsky & Robb procedure (Haab and McConnell, 2002) to obtain standard errors. This procedure estimates WTP (or WTA) through randomly selecting of coefficients (β) from multiple mean coefficients and variance-covariance matrix obtained from a probit model (Cooper, 1994).

2.4.5. Random forest model

The so-called "random forest" modeling approach has been widely applied in ecology (Nam et al., 2015), geography (Mutanga et al., 2012), biology (Boulesteix et al., 2012) and generally in the social sciences (Bravo Sanzana et al., 2015). As a powerful machine-learning technique, random forest possesses two appealing merits: bootstrap sampling and random feature selection techniques (Jiang et al., 2007). A random forest model can be used for a wide range of regression and prediction problems, even though the sampling data are nonlinear and involve complex interaction effects (Strobl et al., 2007). Furthermore, random forest model can be used with relatively small number of observations (Grömping, 2009) and has high prediction accuracy based on the attempts of a large number of various trees classification (Strobl et al., 2007; Svetnik et al., 2003). Consequently, we think it could be useful to apply a random forest model here to test which variables are the best predictors of WTP/WTA.

Random forest model combines an ensemble of trees that grow under the classification and regression tree guidance (Breiman et al., 1984). The basic algorithm of random forest modeling is to perform binary splitting of the explanatory variables recursively in a large number of random trees (Ma, 2005) to minimize the impurity at each node (Breiman et al., 1984). Each explanatory variable is tested for the level of impurity reduction. The smaller the impurity is, the more homogeneous the data are at the specific node (Ma, 2005). It is a criterion to test how well each split classifies the data according to certain explanatory variables. All the above procedure is repeated until the nodes can't be split any longer. The impurity is measured with a Gini index calculated as the difference of impurity between before and after the split of a sampling tree based on a certain variable.

First, the initial Gini index before split is computed as following equations:

$$I(D) = 1 - P(D_+)^2 - P(D_-)^2 \tag{6a}$$

where I(D) refers to the impurity of data D; +/− denotes to the class of attribute; $P(D_+)$ is the proportion of data with "+" attribute.

After the split, the impurity for the left and right child nodes could be expressed as Eqs. (6b)–(6c):

$$I(D_l) = 1 - P(D_{l+})^2 - P(D_{l-})^2 \tag{6b}$$

$$I(D_r) = 1 - \mathrm{P}(D_{r+})^2 - P(D_{r-})^2 \tag{6c}$$

where $P(D_{l+})$ is the proportion of left subset data with attribute "+"; $P(D_{l-})$ is the proportion of left subset data with attribute "−"; $P(D_{r+})$ is the proportion of right subset data with attribute "+"; $P(D_{r-})$ is the proportion of right subset data with attribute "−".

Finally, the formula for Gini index is Eq. (6d):

$$Gini(D) = I(D) - p_l * I(D_l) - p_r * I(D_r) \tag{6d}$$

where p_l and p_r are the proportions of left and right subsets data.

The WTP/WTA is predicted based on the recursive partitioning of the explanatory variables in a large number of random tree-samples and each sampling tree generates a prediction value of the dependent variable (WTP/WTA).

3. Results

3.1. Descriptive statistics

In total, 1709 respondents participated in the WTP/WTA survey. About 0.7%, 1.5%, 2.0% and 6.2% of the total responses were removed as non-valid responses in $\mathrm{WTP_f}$, $\mathrm{WTP_w}$, $\mathrm{WTA_f}$ and $\mathrm{WTA_w}$ surveys respectively. Of the rest, the rates of valid zero-bidders were 10.8% for $\mathrm{WTP_{fh}}$, 24.9% for $\mathrm{WTP_{fm}}$, 3.2% for $\mathrm{WTP_{wh}}$, 10.2% for $\mathrm{WTP_{wm}}$, 37.9% for $\mathrm{WTA_{fh}}$, 41.0% for $\mathrm{WTA_{fm}}$, 8.0% for $\mathrm{WTA_{wh}}$, 9.4% for $\mathrm{WTA_{wm}}$ respectively. The descriptive statistics for the independent variables are reported in Table 1 for the different samples. Comparing the respondents of the $\mathrm{WTP_f}$ or $\mathrm{WTA_f}$ survey format samples, their background was similar in terms of gender, education, income etc. (Table 1). However, age and education differed between FTF and WB questionnaire collection modes: the WB questionnaire had more young-aged and high-educated respondents. In the survey, we classify respondents' knowledge of $PM_{2.5}$ pollution into four levels: *Very well*, *Largely*, *Partly*, *Nothing*. Among the questionnaire samples, 69%–87% of the respondents at least "*partly*" know that $PM_{2.5}$ is harmful to the health and mood, but only around 3%–8% of the respondents understand it "*very well*". The average outdoor time of interviewees is around 2.6 h/day in the face-to-face survey, which is slightly higher than the outdoor time of WB survey participants. In all survey modes, more than 90% of the respondents reported to have experienced at least one health impact and showed depressed mood symptom due to $PM_{2.5}$ pollution. Specifically, over 70% of the respondents stated that they have experienced symptoms of respiratory diseases including asthma attacks and chronic bronchitis.

3.2. The probability of having a positive WTP/WTA for $PM_{2.5}$ pollution

The probability of having a positive $\mathrm{WTP_f}$, $\mathrm{WTA_f}$, $\mathrm{WTP_w}$ and $\mathrm{WTA_w}$ were estimated. To keep the focus on the main findings, we will mainly present the results for $\mathrm{WTP_f}$/$\mathrm{WTA_f}$, and show the results from $\mathrm{WTP_w}$/$\mathrm{WTA_w}$ in the Appendix. However, whenever results differ substantially between the two, we will comment on it in the text. The probabilities of positive WTP/WTA estimated via probit

Table 1
Descriptive summary for primary independent variables.

Parameters	Description	Levels	Population statistics[a]	WTP_f	WTP_w	WTA_f	WTA_w
GE	1 if male, 0 if female	Male	51%	48%	43%	44%	51%
		Female	49%	52%	57%	56%	49%
AGE	1 under 18; 2 19–28; 3 29–60; 4 over 60	<18	13%	15%	0.5%	20%	1.5%
		19–28	21%	36%	79%	14%	80%
		29–60	50%	39%	19%	56%	17%
		>60	17%	10%	1.5%	10%	1.5%
EDU	1 College and below; 2 bachelor; 3 master; 4 PhD	PhD	3%	7%	13%	3%	18%
		Master	4%	15%	42%	13%	50%
		Bachelor	31%	36%	49%	37%	27%
		College and below	62%	42%	6%	47%	5%
In (CNY)	Individual net income/month	Average	4408	5630	4040	6320	4240
		sd	–	4710	4280	5320	4040
SE	1 No; 2 sometimes; 3 often; 4 frequently	Frequently	–	25%	24%	33%	18%
		Often		45%	56%	45%	60%
		Sometimes		21%	14%	13%	18%
		No		8%	6%	9%	4%
PK	1 Nothing; 2 partly; 3 largely; 4 very well	Very well	–	6%	3%	8%	7%
		Largely		37%	28%	33%	24%
		Partly		42%	38%	45%	48%
		Nothing		15%	31%	13%	21%
OT (hrs/day)	Individual outdoor exposure time	Average	–	2.6	1.9	2.7	1.8
		sd		1.9	1.4	2.2	1.1
DM	1 if yes	Yes	–	91%	95%	86%	95%
		No		9%	5%	14%	5%
HI	1 if with certain health impacts, 0 if without impacts	Respiratory disease	–	77%	89%	72%	85%
		Allergy		2%	0.7%	4%	0.5%
		Lung disease		20%	10%	23%	14%
		None		1%	0.3%	1%	0.5%
DT	1 if central districts, 0 if peripheral districts	Central	56%	76%	85%	71%	91%
		Peripheral	44%	24%	15%	29%	8%
Cost	Cost due to $PM_{2.5}$	Average	–	–	–	109	79
		sd		–	–	149	91

Note: "GE" = Gender, "AGE" = age of the respondents, "EDU" = education levels, "In" = Income, "SE" = Smoking exposure, "PK" = $PM_{2.5}$ knowledge, "OT" = Outdoor time, "HI" = Health impacts, "DM" = Depressed in mood, "DS" = Districts where the respondents live, "PM" = daily $PM_{2.5}$ concentration on the survey dates, "Cost" = Medical and work-loss cost caused by $PM_{2.5}$. Districts in central area refer to Dongcheng, Xicheng, Haidian, Chaoyang and Fengtai district, and other districts are grouped into the peripheral districts. The medical and work-loss costs were only investigated in the WTA survey, which was set to avoid the unrealistic answers in certain degree. "–" refers to no available data. The variable of $PM_{2.5}$ knowledge is determined by a self-reported question.

[a] The population statistics were from Beijing Statistics Yearbook 2016 (BMBS, 2016).

model for health and mood impacts respectively are reported in Table 2. The probability of positive WTP was higher than the positive WTA in the FTF survey (Table 2); whereas, the probability of positive WTP was similar to the positive WTA in the WB survey (see Table A.1 in the Appendix). Income showed a significant positive effect on the probability of positive WTP, whereas no significant effect of income

Table 2
Probit model for the probability of positive WTP_f/WTA_f.

Explanatory variables	Probability of WTP/WTA (Coef (Std. error))					
	WTP_{fh} (N = 610)	WTP_{fm} (N = 610)	WTP_{fs} (N = 610)	WTA_{fh} (N = 255)	WTA_{fm} (N = 255)	WTA_{fs} (N = 255)
GE	−0.1 (0.2)	0.2 (0.1)	−0.01 (0.2)	−0.8 (0.2)***	−0.7 (0.2)***	−0.8 (0.2)***
SE	0.1 (0.1)	0.1 (0.1)	0.1 (0.1)	−0.2 (0.1)*	−0.1 (0.1)	−0.2 (0.1)**
In	7.05E−05 (2.13E−05)***	1.87E−05 (1.26E−05)	8.32E−05 (2.32E−05)***	−2.96E−06 (1.63E−05)	−1.83E−05 (1.61E−05)	−5.15E−06 (1.63E−05)
PK	−0.02 (0.1)	0.2 (0.1)**	0.003 (0.1)	0.1 (0.1)	0.04 (0.1)	0.1 (0.1)
HI	0.3 (0.2)*	0.2 (0.1)	0.2 (0.2)	0.2 (0.2)	0.3 (0.2)	0.2 (0.2)
OT	0.01 (0.04)	−0.04 (0.03)	0.001 (0.04)	0.1 (0.04)	0.1 (0.04)**	0.1 (0.04)
DS	0.1 (0.2)	0.2 (0.1)	0.2 (0.2)	0.01 (0.2)	0.1 (0.2)	0.04 (0.2)
GT	4.7 (127.2)	0.3 (0.1)**	4.7 (125.9)	5.7 (180.6)	5.7 (182.9)	5.6 (180.7)
$PM_{2.5}$	0.001 (0.002)	−0.001 (0.001)	0.001 (0.002)	0.001 (0.002)	0.001 (0.001)	0.001 (0.002)
Cost	–	–	–	0.001 (0.001)**	0.001 (0.001)*	0.001 (0.001)*
Constant	0.2 (0.5)	−0.2 (0.4)	0.1 (0.5)	0.2 (0.5)	0.2 (0.5)	0.4 (0.5)
Observations	610	610	610	255	255	255
Log likelihood	−181.9	−333.1	−174.8	−139.7	−142.9	−139.3
Akaike inf. crit.	383.8	686.1	369.7	301.4	307.7	300.6
Probability	0.89 (0.08)	0.75(0.08)	0.90 (0.08)	0.62 (0.2)	0.59 (0.22)	0.62 (0.22)

Note: "GE" = Gender, "AGE" = age of the respondents, "EDU" = education levels, "In" = Income, "SE" = Smoking exposure, "PK" = $PM_{2.5}$ knowledge, "OT" = Outdoor time, "HI" = Health impacts, "DM" = Depressed in mood, "DS" = Districts where the respondents live, "PM" = daily $PM_{2.5}$ concentration on the survey dates, "Cost" = Medical and work-loss cost caused by $PM_{2.5}$, "GT" = Government trust.

* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Table 3
Turnbull estimates (and confidence intervals) for WTP and WTA (Unit: CNY/year, 1 CNY ≈ 0.163 US dollars in 2015).

EWTP/EWTA	WTP_{fh}	WTP_{fm}	WTP_{wh}	WTP_{wm}	WTA_{fh}	WTA_{fm}	WTA_{wh}	WTA_{wm}
Mean	1388.4 (437.9, 3234.2)	897.7 (222.0, 2328.3)	689.9 (228.0, 1938.5)	558.9 (192.0, 1632.2)	2916.9 (624.0 5664.9)	2791.2 (540.0, 6000.0)	3467.5 (1127.7, 6432.6)	3295.7 (1174.8, 6120.3)
Median	1140.0 (1089.0, 1191.0)	762.0 (726.0, 792.0)	564.0 (545.1, 576.1)	468.0 (449.6, 482.7)	2844.0 (2760.0, 2940.0)	2646.0 (2535.0, 2740.0)	3330.0 (3246.0, 3423.0)	3312.0 (3219.0, 3405.0)
1 μg/m^3 $PM_{2.5}$/year	30.8 (9.7, 71.9)	19.9 (4.9, 51.7)	15.3 (5.1, 43.1)	12.4 (4.3, 36.3)	64.8 (13.9, 125.9)	62.0 (12.0, 133.3)	77.1 (25.1, 142.9)	73.2 (26.1, 136.0)

on the probability of positive WTA was observed. The probability of positive WTP_{fm} tended to be high when the respondents had high-level of $PM_{2.5}$ knowledge. The probability of positive WTP_{fm} was also higher when the respondents trust the government more. Men tended to have a lower probability of positive WTA for the $PM_{2.5}$ compensation compared with women. Smoking exposure was positively related to the probability of positive WTP, and negatively related to the probability of positive WTA. The respondents from the central area were willing to pay or accept more than the respondents from the peripheral districts. In the WTA regression models, the variable "cost" including medical cost and work loss due to $PM_{2.5}$ was highly correlated with the probability of positive WTA.

3.3. Estimation of WTP/WTA by parametric and non-parametric models

The Turnbull model, probit model, two-part model, and random forest model were applied for the WTP/WTA estimations and comparisons. The estimated results with two-part model and random forest model were in close agreement with the Turnbull model results.

3.3.1. Turnbull model for WTP/WTA

Table 3 shows the Turnbull estimates of lower-bound WTP and WTA and corresponding 95% confidence intervals calculated with the bootstrap method. In general, WTP obtained from the WB surveys was lower than WTP of FTF surveys and WB WTA was higher than that from the FTF surveys. The discrepancies of WTP and WTA were larger in the mood samples than the health samples.

The average WTP_{fh} and WTP_{fm} were CNY 1388.4 and CNY 897.7 per person per year respectively for the $PM_{2.5}$ reduction (Table 3). Respondents valued their mood depression around 65%, 81%, 96% and 95% of health impacts caused by $PM_{2.5}$ in WTP_f, WTP_w, WTA_f, WTA_w surveys.

As expected, WTA_f and WTA_w were much larger than the corresponding WTP estimates. Specifically, WTA_{fh} doubled WTP_{fh} and WTA_{fm} tripled WTP_{fm}; in terms of WB survey, WTA_{wh} was around five times the WTP_{wh} and WTA_{wm} was six times the WTP_{wm}. As linear exposure-response model is widely applied in epidemiological studies (Madaniyazi et al., 2015; Sun et al., 2013), thus we assume that annual $PM_{2.5}$ concentration (80 μg/m^3 in 2015) is linear to the WTP/WTA, the marginal WTP_{fh}, WTP_{fm}, WTP_{wh}, WTA_{wm} of 1 μg/m^3 $PM_{2.5}$ were around CNY 30 and CNY 20 per person per year respectively (Table 3). The estimated median values of WTP/WTA were slightly lower than the corresponding mean values in all the samples.

3.3.2. Probit model for WTP/WTA

The probit model showed that the WTP is positively correlated with the individual's gender, income, health impacts, outdoor time, districts and government trust (Table 4). Specifically, people who had experienced certain health impacts due to $PM_{2.5}$ were more willing to pay for $PM_{2.5}$ reduction and more willing to accept a compensation due to the pollution. Central area residents tended to pay more for $PM_{2.5}$ reduction and also accept more for $PM_{2.5}$ compensation than the peripheral area residents. The income elasticity of WTP[4] for $PM_{2.5}$ reduction was estimated around 0.31, 0.26, 0.33, 0.07, 0.04, 0.10 in samples of WTP_{fh}, WTP_{fm}, WTP_{fs}, WTA_{fh}, WTA_{fm}, WTA_{fs} respectively.

3.3.3. Two-part model for WTP/WTA

The explanatory variables coefficients of positive WTP_f/WTA_f estimated with the linear regression in the two-part model are reported in Table 5. WTP_f tended to be higher for men, people with higher income and people who trust the government. In terms of the WTA_f, people accepted lower compensation when they were exposed more frequently to smoke. Additionally, the individual income and cost had a positive effect on the WTA.

3.3.4. Random forest for WTP/WTA

The random forest algorithm generates the rankings of the variable importance according to the reduction level of the impurity after the split of explanatory variables of the sampling trees. For different survey modes, the variable-importance rankings were different from each other. Variables of "Income", "$PM_{2.5}$ concentration", "smoking exposure" were identified as the top three important variables influencing the impurity of the tree samples in the WTP_{fh} and WTP_{fm} (Fig. 2). The variable of "health impacts" reduced more of the impurity in WTP_f than in the WTA_f survey, which indicated that "health impacts" did not influence much on the WTA. In WTA surveys, the "cost" is the most important variable that reduces the impurity after the splits.

With thousands of trees built in a random forest and their recursive split nodes, it is plausible to show the density distribution of WTP/WTA estimated with random forest (Fig. 3). The distributions tell where a certain value of WTP/WTA lies on the distribution and what the probability of a certain value of WTP/WTA is. Results show that the distributions of WTP/WTA were a bit right-skewed, thus the mean WTP/WTA values were slightly larger than the median ones. Taking WTP_{fh} as an example, the mean was CNY 1369.7 and median was CNY 1311.5. Additionally, in the WTP probability density distributions, the probability of WTP/WTA fluctuated slightly when they were close to the right tail. As the respondents prefer to the round numbers of bids such as 100, 200 instead of numbers such as 70, 120. Thus, more people choose the numbers of multiple hundreds in payment card, which results in bimodal peaks in the probability density plot (Fig.3).

(Unit: 100 CNY/year).

3.3.5. Comparisons between different model estimations

WTP and WTA estimates with probit model, two-part model and random forest model are compared in Table 6 by estimating the aggregated WTP/WTA for the population. Notice that the estimated sum of health and mood is not exactly identical to the sum of the results from the two separate models (see previous sections). All the confidence intervals were estimated with the bootstrap percentile method. Table 6 shows that mean WTP and WTA estimated with the two-part and random forest models were quite similar to the results estimated via the Turnbull model. Most models show similar results for the different samples, the exception being the probit model for WTA in the WB survey. Here we find a large WTA, one possible reason being the distributional assumptions underlying the model, causing it to be quite sensitive to outliers. Thereby, the WTP/WTA was estimated by probit model without

[4] i.e. the sensitivity of WTP to changes in income.

Table 4
Parameters for a probit model for the WTP/WTA (cf. Eq. (5d)).

Explanatory variables	WTP/WTA samples (coef. (std. error))					
	WTP_{fh} (N = 610)	WTP_{fm} (N = 610)	WTP_{fs} (N = 610)	WTA_{fh} (N = 255)	WTA_{fm} (N = 255)	WTA_{fs} (N = 255)
GE	0.01 (0.03)	0.12 (0.03)***	0.10 (0.02)***	−0.50 (0.05)***	−0.48 (0.05)***	−0.50 (0.02)***
SE	−0.04 (0.02)**	0.02 (0.02)	−0.03 (0.01)***	−0.18 (0.01)***	−0.13 (0.03)***	−0.19 (0.01)***
In	5.59E−05 (3.35E−06)***	4.57E−05 (3.26E−06)***	5.84E−05 (1.77E−06)***	1.18E−05 (4.14E−06)***	6.25E−06 (4.10E−06)	1.60E−05 (2.15E−06)***
PK	−0.10 (0.02)***	0.06 (0.02)***	−0.05 (0.01)***	0.15 (0.03)***	0.10 (0.03)***	0.14 (0.01)***
HI	0.13 (0.04)***	0.08 (0.04)**	0.08 (0.02)***	0.23 (0.10)***	0.26 (0.05)***	0.27 (0.03)***
OT	0.03 (0.01)***	−0.01 (0.01)*	0.02 (4.25E−03)***	0.07 (0.01)***	0.09 (0.01)***	0.08 (0.01)***
DS	0.07 (0.03)**	0.10 (0.03)***	0.09 (0.02)***	−0.01 (0.05)	−0.03 (0.05)	−0.03 (0.03)
GT	0.24 (0.04)***	0.34 (0.04)***	0.27 (0.02)***	1.13 (0.09)***	1.01 (0.08)***	1.08 (0.04)***
$PM_{2.5}$	−5.62E−04 (3.54E−04)	−1.42E−03 (3.58E−04)***	−6.71E−04 (1.95E−04)***	−4.25E−04 (3.62E−04)	1.35E−04 (3.58E−04)	−2.69E−04 (1.88E−04)
Bid	−4.81E−03 (1.40E−04)***	−0.01 (1.76E−04)***	−2.86E−03 (3.99E−05)***	−2.10E−03 (8.95E−05)***	−1.99E−03 (8.78E−05)***	−1.18E−03 (2.80E−05)***
Cost	–	–	–	1.55E−03 (1.57E−04)***	1.27E−03 (1.51E−04)***	1.69E−03 (1.88E−04)***
Constant	0.34 (0.10)***	−0.29 (0.10)***	−0.02 (0.05)	0.07 (0.13)	−0.06 (0.13)	−0.07 (0.07)
Observations	10,370	10,370	37,210	4335	4335	15,555
Log likelihood	−181.9	−333.1	−174.8	−139.7	−142.9	−139.3
Akaike inf. crit.	383.8	686.1	369.7	301.4	307.7	300.6

Note: "GE" = Gender, "AGE" = age of the respondents, "EDU" = education levels, "In" = Income, "SE" = Smoking exposure, "PK" = $PM_{2.5}$ knowledge, "OT" = Outdoor time, "HI" = Health impacts, "DM" = Depressed in mood, "DS" = Districts where the respondents live, "PM" = daily $PM_{2.5}$ concentration on the survey dates, "Cost" = Medical and work-loss cost caused by $PM_{2.5}$, "GT" = Government trust.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

sociodemographic variables as comparisons. This analysis indicated that, without sociodemographic information, the WTA_w estimates were relatively consistent with the two-part model and random forest model results (Table 6), which demonstrated that the unrepresentativeness of the sampling respondents had a substantial effect on the probit model results. The two-part model showed a relatively larger confidence interval compared with the random forest estimates. The estimated WTP_s and WTA_s were consistent with WTP_h + WTP_m and WTA_h + WTA_m in two-part model and random forest model.

To test if the difference between FTF and WB surveys were caused by the socio-demographic structure of the population in Beijing, according to the demographic statistics in Table 1, we calibrated the independent variables of gender, income, districts distribution and $PM_{2.5}$ concentration in the two-part model. The predicted results indicated that the WTP and WTA estimates with calibration were quite close to the results without calibration and the detailed results could be accessed in Appendix. The obvious WTP/WTA gaps could still be seen between FTF- and WB-survey modes with the population statistics calibration.

4. Discussion

It is well-established that the health impacts caused by $PM_{2.5}$ pollution result in substantial welfare losses (Nagashima et al., 2017; Rabl and Spadaro, 2000; Weidema, 2009). Despite the attention put on health impacts, however, the welfare loss due to depression and anxiety induced by $PM_{2.5}$ has been barely investigated. Consequently, we

Table 5
Two-part model. Parameters (cf. Eq. (4b)) for a linear regression for a positive WTP_f/WTA_f.

Explanatory variables	Positive WTP_f/WTA_f (coef (std. error))					
	WTP_{fh} (N = 544)	WTP_{fm} (N = 458)	WTP_{fs} (N = 547)	WTA_{fh} (N = 158)	WTA_{fm} (N = 151)	WTA_{fs} (N = 159)
GE	33.6 (17.3)*	26.5 (15.6)*	59.5 (27.5)**	−25.1 (62.2)	−35.8 (65.1)	−64.2 (117.2)
SE	−9.2 (9.88)	3.0 (9.0)	−5.3 (15.7)	−62.5 (32.6)*	−49.1 (34.3)	−86.6 (60.7)
In	0.008 (0.002)***	0.07 (0.002)***	0.014 (0.003)***	0.013 (0.005)**	0.017 (0.006)***	0.026 (0.010)**
PK	−20.7 (11.0)	−2.6 (10.1)	−16.2 (17.6)	44.8 (33.0)	68.8 (34.2)**	102.1 (61.3)*
HI	5.0 (21.3)	−13.3 (18.9)	1.5 (33.9)	10.5 (68.3)	34.4 (71.0)	39.9 (128.9)
OT	2.7 (4.6)	−1.6 (4.3)	−0.4 (7.3)	20.8 (13.7)	19.0 (14.2)	50.5 (25.8)*
DS	3.1 (19.8)	6.5 (18.1)	10.5 (31.6)	12.5 (61.0)	−65.3 (64.0)	−46.9 (115.1)
GT	9.7 (19.8)	51.6 (18.3)**	54.5 (31.5)*	202.3 (78.4)**	161.1 (79.7)**	394.4 (147.5)***
$PM_{2.5}$	−0.1 (0.2)	−0.1 (0.2)	−0.2 (0.3)	−0.4 (0.4)	−0.2 (0.4)	−0.6 (0.8)
Cost	–	–	–	0.6 (0.2)***	0.6 (0.2)***	1.1 (0.4)***
Constant	130.3 (0.5)**	47.1 (51.9)	150 (91.2)	243.3 (175.8)	162.9 (179.1)	0.4 (0.5)
Observations	544	458	547	158	151	159
Residual std. error	194.9	161.6	310.6	343.1	348.6	648.2
F statistic	3.8***	4.2***	4.2***	3.4***	3.4***	3.6***

Note: "GE" = Gender, "AGE" = age of the respondents, "EDU" = education levels, "In" = Income, "SE" = Smoking exposure, "PK" = $PM_{2.5}$ knowledge, "OT" = Outdoor time, "HI" = Health impacts, "DM" = Depressed in mood, "DS" = Districts where the respondents live, "PM" = daily $PM_{2.5}$ concentration on the survey dates, "Cost" = Medical and work-loss cost caused by $PM_{2.5}$, "GT" = Government trust.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

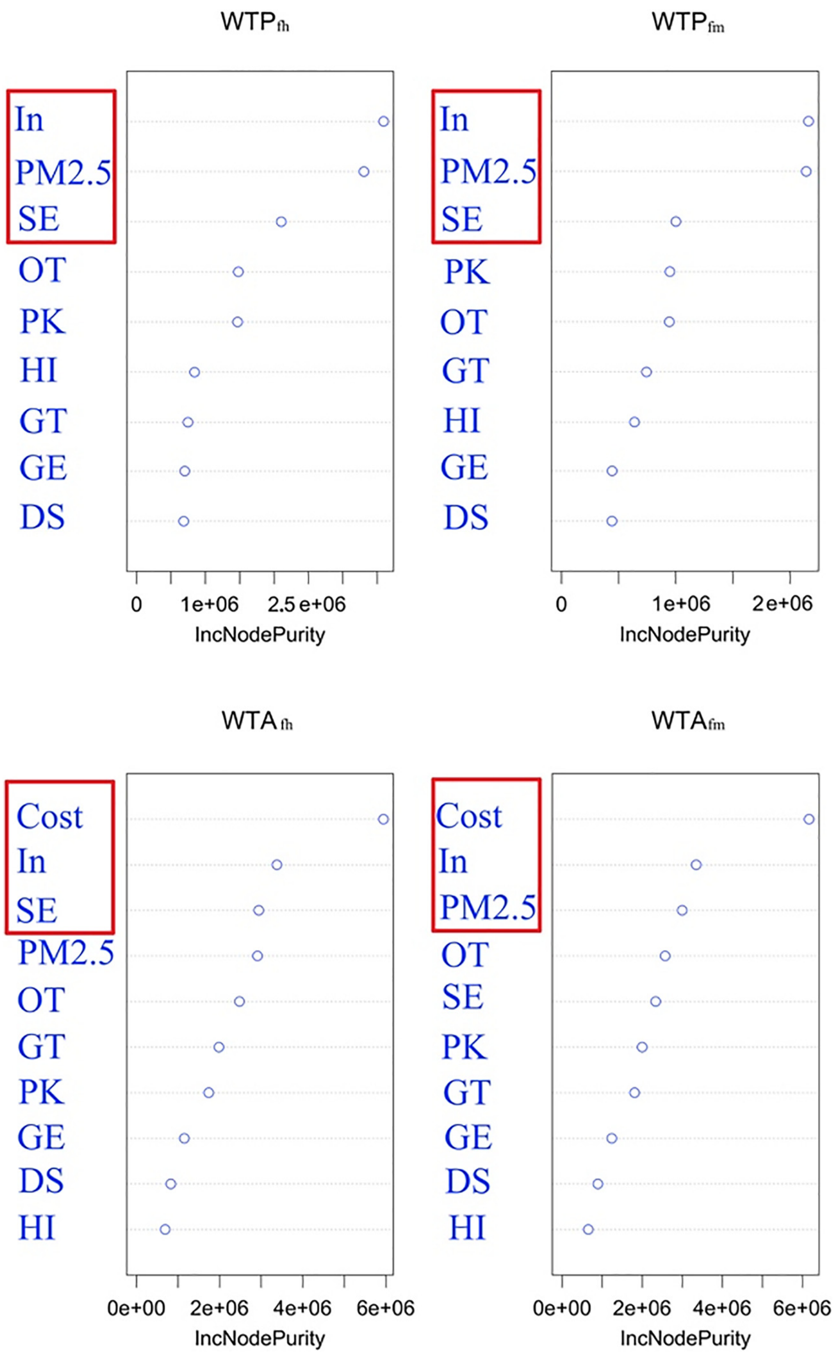


Fig. 2. Variables importance in random forest. Note: 'IncNodePurity' denotes the increase of node purity after each binary split in the random forest algorithm, which is the same value with the reduction of the node impurity. "GE" = Gender, "AGE" = age of the respondents, "EDU" = education levels, "In" = Income, "SE" = Smoking exposure, "PK" = $PM_{2.5}$ knowledge, "OT" = Outdoor time, "HI" = Health impacts, "DM" = Depressed in mood, "DS" = Districts where the respondents live, "PM" = daily $PM_{2.5}$ concentration on the survey dates, "Cost" = Medical or work-loss cost caused by $PM_{2.5}$, "GT" = Government trust.

estimated the perceived impacts on both health and mood due to $PM_{2.5}$ pollution in Beijing. Further, we test various estimation and survey data collection methods. Results show that the perceived welfare loss of the $PM_{2.5}$ related mood depression constitutes a large portion of the total welfare loss. Therefore, the estimated perceived welfare loss accounting for health impacts and mood depression caused by $PM_{2.5}$ was higher than previous results focusing on health impacts alone (Ito and Zhang, 2016; Sun et al., 2016a). In other words, the reduction of $PM_{2.5}$ pollution has the benefits of not only public health improvement but also improvements in the mental health of individuals.

4.1. Main findings

The perceived welfare loss due to health impacts caused by $PM_{2.5}$ was around CNY 29.3 (9.2, 66.9) billion/year, and in terms of mood depression, the perceived welfare loss was around CNY 19.7 (5.2, 50.8) billion/year. For the whole society, the perceived welfare loss including health and mood impacts was around CNY 49.3 (16.1, 113.8) billion, which equates to 2.2% (0.7%, 4.95%) of regional GDP in Beijing in 2015. In terms of WTA survey, the results showed that the social welfare loss was around CNY 63.3 (13.5, 127.6) billion and CNY 59.9 (10.9,

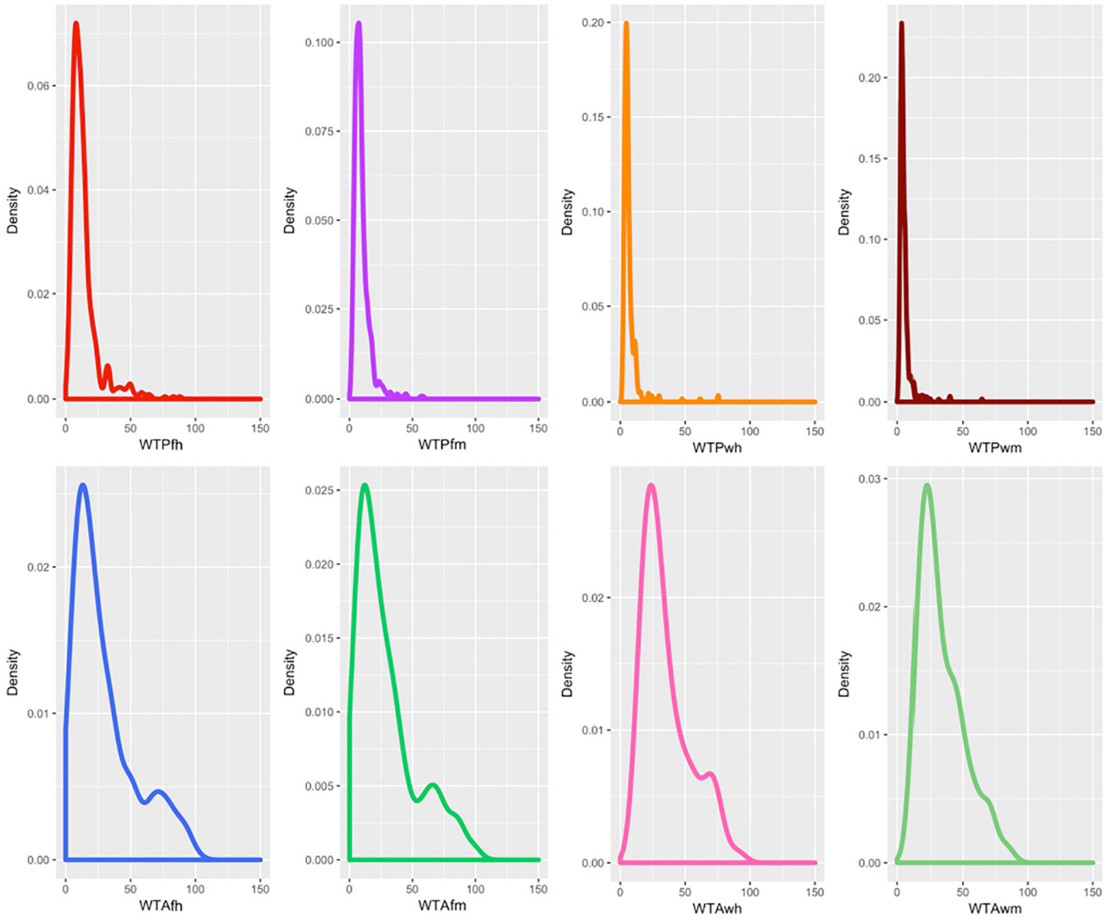


Fig. 3. The WTP/WTA probability density distribution with random forest.

119.8) billion for health impact and mood depression respectively. The perceived welfare loss was around 125.5 (30.7, 290.7) billion in 2015 accounting 5.5% (1.3%, 12.7%).

The lower-bound estimation with the Turnbull model showed that the perceived welfare loss/person/year including health impacts and mood depression was around CNY 2286.1 (720.0, 4903.2). For health

Table 6
WTP/WTA estimation with different models.

		WTP_f			WTP_w			WTA_f			WTA_w		
		WTP_{fh}	WTP_{fm}	WTP_{fs}	WTP_{wh}	WTP_{wm}	WTP_{ws}	WTA_{fh}	WTA_{fm}	WTA_{fs}	WTA_{wh}	WTA_{wm}	WTA_{ws}
Probit model	Mean (95% CI)	1297.3 (1051.4, 1533.4)	589.7 (336.6, 851.0)	1289.2 (1056.5, 1527.7)	756.2 (127.7, 1376.6)	581.8 (338.0, 820.2)	881.7 (362.9, 1362.1)	2097.9 (568.4, 3528.9)	1834.2 (301.9, 3427.8)	2772.1 (1366.5, 4190.6)	13,778.0 (9012.5, 18,328.5)	16,798.0 (11,845.3, 21,688.9)	13,394.0 (9000.0, 18,025.6)
	Total (billion)	28.2 (22.8, 33.3)	12.8 (7.3, 18.5)	28.0 (22.9, 33.2)	16.4 (2.8, 29.9)	12.6 (7.3, 17.8)	19.2 (7.9, 29.6)	45.5 (12.3, 76.6)	39.8 (6.6, 74.4)	60.2 (29.7, 91.0)	299.1 (195.6, 397.8)	364.6 (257.1, 470.8)	290.7 (195.3, 391.2)
Probit model excluding sociodemographics	Mean (95% CI)	1335.05 (1331.2, 1338.9)	607.3 (604.6, 610.1)	1364.2 (1361.1, 1367.3)	647.6 (645.7, 649.5)	519.4 (517.9, 521.0)	635.9 (633.0, 638.9)	1886.1 (1875.33, 1896.94)	1471.0 (1459.6, 1482.3)	2224.9 (2213.6, 2236.2)	3937.7 (3919.4, 3956.0)	3546.1 (3535.7, 3556.6)	5831.6 (5823.6, 5839.5)
	Total (billion)	29.0 (28.9, 29.1)	13.2 (13.1, 13.3)	29.6 (29.5, 29.7)	14.1 (14.0, 14.1)	11.3 (11.2, 11.3)	13.8 (13.7, 13.9)	40.9 (40.7, 41.2)	31.9 (31.7, 32.2)	48.3 (48.0, 48.5)	85.5 (85.1, 85.9)	77.0 (76.7, 77.2)	126.6 (126.4, 126.8)
Two-part model	Mean (95% CI)	1357.4 (920.7, 1835.1)	914.0 (511.6, 1360.8)	2287.5 (1546.9, 3077.1)	659.8 (386.2, 986.1)	541.9 (292.5, 874.7)	1260.8 (799.0, 1871.1)	2815.7 (1189.9, 4573.7)	2720.1 (1043.1, 4368.8)	5639.7 (2583.9, 8942.6)	3443.6 (2543.6, 4300.8)	3326.8 (2395.0, 4159.3)	6779.9 (5014.8, 8289.5)
	Total (billion)	29.4 (19.6, 39.2)	20.0 (11.0, 29.1)	49.7 (33.6, 65.5)	14.5 (8.6, 22.1)	11.7 (6.4, 18.9)	27.3 (16.5, 41.4)	60.9 (27.3, 98.4)	60.1 (25.8, 98.6)	123.4 (52.3, 200.6)	75.1 (52.4, 93.7)	71.8 (51.2, 89.4)	147.3 (108.0, 180.4)
Random forest	Mean (95% CI)	1369.7 (882.6, 2137.6)	946.2 (663.4, 1415.4)	2294.4 (1503.0 3394.1)	696.4 (473.5, 1289.8)	547.3 (387.3, 962.3)	1216.0 (881.6, 2139.0)	2944.4 (1716.2, 4331.9)	2848.7 (1740.7, 4176.7)	5838.7 (3073.0, 8965.3)	3524.4 (2540.1, 4606.5)	3376.2 (2526.3, 4355.4)	6871.3 (5207.7, 8876.8)
	Total (billion)	29.4 (19.2, 45.3)	20.4 (12.9, 33.5)	50.0 (33.5, 75.5)	15.0 (10.4, 27.8)	12.0 (8.3, 21.3)	26.8 (19.1, 46.5)	64.4 (38.5, 96.9)	62.2 (36.2, 91.5)	126.1 (71.7, 184.5)	76.2 (55.5, 100.9)	73.2 (54.2, 94.9)	149.0 (114.0, 188.3)

and mood comparison, the WTP_m/WTA_m were around 60%–90% of the WTP_h/WTA_h, which suggested that the respondents attached great importance on the mood and psychological feelings influenced by the $PM_{2.5}$ pollution. For different survey formats, WTA was about 50%–300% higher than WTP (Table 6). In a comparison of different survey modes, WTP_w was 50% lower than the WTP_f and WTA_w was >20% higher than the WTA_f. WTA from the WB survey was the highest estimates compared with all the other samples (Table 6).

In the study, young and highly educated respondents participated the WB survey, a similar effect as observed in the previous questionnaire-based studies (Canavari et al., 2005; Szolnoki and Hoffmann, 2013). This could be influenced by the fact that young and high-educated people use the internet more frequently than the aged and low-educated people (Bakker and De Vreese, 2011; Meerkerk et al., 2009; Szolnoki and Hoffmann, 2013). As there could be unrepresentative of people who participated FTF- and WB-survey modes, we calibrated variables according to the socio-demographic statistics in Beijing to verify the difference of WTP/WTA between FTF and WB surveys. The predicted results showed that the demographic statistics were not the crucial reasons for the variance between FTF and WB data collection modes. Therefore, the significant difference between FTF and WB surveys could be mainly due to the social desirability effect (Nielsen, 2011). This could be explained by the fact that the respondents wanted to show more socially desirable or responsible characteristics (Krosnick, 1999; Leggett et al., 2003). So the presence of the interviewers could lead the respondents to provide a higher WTP or a lower WTA to please the interviewers or behave in line with the societal norms or expectations (Fisher, 1993).

The Turnbull model, probit model, two-part model and random forest were applied for the WTP/WTA estimation. In comparisons with the other models, the probit model produced much higher estimates for the WTA_w but not for the other sample. A possible explanation is that this sample is the one that differs the most in terms of representativeness, which may cause that findings extrapolated from the sample to the population to exaggerate the effects. Furthermore, the probit model relies on a specific distribution and if there are outliers the effect may be more accentuated in this model. Both the two-part model and the random forest model generate similar results to the non-parametric estimation of the Turnbull model However, the two-part model has relatively larger confidence intervals compared with the random forest model. Except for the estimation of WTA_w, all the models showed similar estimates of WTP and WTA in both the health and mood samples. Although random forest cannot test the correlation between WTP/WTA and the explanatory variables, the advantage of this method is that it can rank the relative importance of variables and does not need any distribution assumptions. This indicates that random forest is a precise and effective estimation tool for future contingent valuation studies.

4.2. Previous studies comparison

The welfare loss from the health and mood impacts of $PM_{2.5}$ pollution is approximately CNY 50.0 (33.5, 75.5) billion per year, which is equivalent to 2.2% (1.5%, 3.3%) of the regional GDP of Beijing in 2015. In comparison, a previous study estimated that the health loss from particulate pollution was 1.03% of the regional GDP in Shanghai (Kan and Chen, 2004). Our study results showed that the WTP_{fh} was CNY 1388/person/year; whereas previous CVM surveys of WTP for smog reduction was only CNY 428/person/year (Sun et al., 2016a) and for 80% $PM_{2.5}$ reduction in Beijing-Tianjin-Hebei region was CNY 602/person/year (Wei and Wu, 2017). The large difference between our results and those from prior studies could be due to differences in the valuation of goods, pollution awareness and demographic characteristics of respondents between Beijing and Tianjin-Hebei regions. On the other hand, another WTP study for smog mitigation was estimated about CNY 1590.36/person/year in China (Sun et al., 2016b), which was larger than the WTP_{fh}, but lower than the total WTP_{fs} estimates in this study. Another study reported that the WTP for 1 $\mu g/m^3$ of $PM_{2.5}$ reduction was CNY 539/person/year in 2014 (Zhang et al., 2017), which was around 17 times the estimates with Turnbull model in our study. In terms of welfare loss, some studies investigated the welfare loss due to air pollution with a Computable General Equilibrium (CGE) model and the results showed that the welfare loss was around CNY 227.6 billion, or 6.5% of China's GDP in 2007 (Chen and He, 2014). Another $PM_{2.5}$ pollution economic estimation, also applying a CGE model, found that the welfare loss including morbidity and mortality cost due to $PM_{2.5}$ pollution will be around 2.26%–3.14% of regional GDP in Shanghai in 2030 (Wu et al., 2017) and for around 2.75%–3.92% of regional GDP in Beijing in 2030 (Xie et al., 2016a). The above estimates are lower than the welfare loss estimated in our case which was around 2.2%–5.5% of regional GDP in Beijing in 2015.

Overall, the $PM_{2.5}$ WTP for health estimated in this study is relatively consistent with other studies. However, most previous studies do not consider the mood impacts of $PM_{2.5}$. As a non-market environmental good, clean air has multidimensional values/attributes including health services (Howarth and Farber, 2002) visibility (Fischhoff and Furby, 1988), scenic services (Howarth and Farber, 2002), and other non-use values. Therefore, both the $PM_{2.5}$ related deterioration of health conditions and of visibility of scenes should be considered in the contingent valuation. With two separate elicitations of health and mood, we could evaluate respondents' value of clean air and figure out what they lose due to the $PM_{2.5}$ pollution in multiple dimensions. Thus, our monetary estimates are higher than those of previous contingent valuation studies.

4.3. Study limitations

Our study faces some limitations. First of all, the sampling size is too small to create a representative sample of the population of Beijing. Chinese people have very busy working hours, which makes it difficult to access the residents in Beijing. It is also difficult to do households surveys because most of the residential communities are gated in Beijing, thus the available survey sites are limited to public parks, shopping malls and open communities. One way that we have dealt with this issue is by using two different data collection methods – face-to-face, and online survey. The questionnaire respondents in the survey could be unrepresentative and not randomly selected. As a result, we calibrated the regression model for the verification of the socio-demographic influence on the WTP/WTA.

$PM_{2.5}$ is a public "bad", and people could be both polluters and victims at the same time. Consequently, it is not obvious whether WTP or WTA would provide the best estimates in terms of property rights. Therefore, we conducted both WTP and WTA surveys in FTF and WB modes and applied four models for the WTP/WTA estimates, which could provide a "bigger picture" of welfare loss. The WTA/WTP ratios were 1.5–4.0, and possible reasons for the difference are the substitution effect, income effect (Hanemann, 1991) and endowment effect (Morrison, 1998). Thus, WTA could be overrated in a certain degree, and we propose the WTP as the conservative estimate for the $PM_{2.5}$ related health and mood loss.

We assumed the respondents could fully understand the questions asked. In the FTF surveys, the investigators could give respondents explanations as clear as possible, which could avoid misunderstanding of relevant expressions or questions, to a certain degree. We expected this problem to be more serious in the WB surveys and to overcome this issue we added some further explanations in the WB questionnaire. Still, the respondents may have had difficulties making decisions on the hypothetical assumptions of non-market goods transactions.

The payment card elicitation method was applied in this study, as it could provide information for the respondents and save survey time. As the bid number, values and intervals would influence the respondents' choices, the format could cause an anchoring effect on the proposed bids (Chanel et al., 2016). In payment card vehicle, the true WTP/WTA is assumed to be higher than the bid chosen and lower than the next higher one. Thus, the WTP and WTA elicited through a lower bound payment card technique are conservative estimates.

In the study, we did not point out if the WTP/WTA is a mandatory or voluntary payment or a donation to a governmental department. A donation can induce free-riding effects. This could influence the response rate and payment levels (Shah et al. 2017). Further, if a donation was assumed, the lack of trust for the government or the concerns of corruption would cause the respondents to be unwilling to express their true WTP/WTA. And finally the lack of specified payment method results in lack of consequentiality. Looking at the good itself, clean air is a public good, and paying for it may therefore include free-riding possibilities (Samuelson, 1954), leading to an under-estimation of the WTP (or WTA). On the other hand, the hypothetical setting, where people do not actually have to pay, likely lead to an overestimation. In conclusion, these aspects may lead to under or overestimation.

5. Conclusions

The study proposed a WTP/WTA estimation, including both perceived health impacts and mood depression, through FTF and WB surveys. The results showed that in addition to the welfare loss from negative health effects, respondents attached great importance to their psychological feelings, suggesting that the mental health should be also quantitatively estimated in future studies. A random forest model was successfully applied in the contingent valuation for the first time to our knowledge, with good performance, and is recommended for further testing in future contingent valuation studies.

Acknowledgements

This study was supported by the National Key Research and Development Program of China (No. 2016YFC0502802), the Fund for Innovative Research Group of the National Natural Science Foundation of China (No. 51721093), the National Natural Science Foundation of China (No. 41271105) and China Scholarship Council (No. [2015]3022), the project “Developing an indicator system for economic benefit and loss assessment of environmental impacts” of EIA Management in 2017 administrated by the Ministry of Environmental Protection, China. Jette Bredahl Jacobsen would also like to thank the Danish National Research Foundation (No. DNRF96) for supporting the research at the Centre for Macroecology, Evolution and Climate. We would like to thank Jilin Hu for introducing random forest method to the study and also thank for all the participants' endeavor in the questionnaire surveys.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2018.02.275.

References

Alberini, A., Cropper, M., Fu, T.T., Krupnick, A., Liu, J.T., Shaw, D., Harrington, W., 1997. Valuing health effects of air pollution in developing countries: the case of Taiwan. J. Environ. Econ. Manag. 34, 107–126.
Archer, K.J., Kimes, R.V., 2008. Empirical characterization of random forest variable importance measures. Comput. Stat. Data Anal. 52, 2249–2260.
Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOAA panel on contingent valuation. Fed. Regist. 58, 4601–4614.
Bakker, T.P., De Vreese, C.H., 2011. Good news for the future? Young people, Internet use, and political participation. Commn. Res. 38, 451–470.
BMBS, 2016. Beijing statistical yearbook 2016, Beijing. http://www.bjstats.gov.cn/nj/main/2016-tjnj/zk/indexch.htm.
BMEPB, 2015, March 14. Beijing municipal environmental protection bureau website. http://www.bjepb.gov.cn/.
BMEPB, 2016. Beijing environmental statement 2015, Beijing. http://www.bjepb.gov.cn/bjepb/323474/324034/324735/index.html.
Boulesteix, A.L., Janitza, S., Kruppa, J., König, I.R., 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Data Min. Knowl. Disc. 2, 493–507.
Bravo Sanzana, M., Salvo Garrido, S., Muñoz Poblete, C., 2015. Profiles of Chilean students according to academic performance in mathematics: an exploratory study using classification trees and random forests. Stud. Educ, Eval. 44, 50–59.
Breffle, W.S., Eiswerth, M.E., Muralidharan, D., Thornton, J., 2015. Understanding how income influences willingness to pay for joint programs: a more equitable value measure for the less wealthy. Ecol. Econ. 109, 17–25.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. Classification and Regression Trees. CRC press.
Calderón-Garcidueñas, L., Vojdani, A., Blaurock-Busch, E., Busch, Y., Friedle, A., Franco-Lira, M., Sarathi-Mukherjee, P., Martínez-Aguirre, X., Park, S.-B., Torres-Jardón, R., 2015. Air pollution and children: neural and tight junction antibodies and combustion metals, the role of barrier breakdown and brain immunity in neurodegeneration. J. Alzheimers Dis. 43, 1039–1058.
Cameron, T.A., Quiggin, J., 1994. Estimation using contingent valuation data from a. J. Environ. Econ. Manag. 27, 218–234.
Canavari, M., Nocella, G., Scarpa, R., 2005. Stated willingness-to-pay for organic fruit and pesticide ban: an evaluation using both web-based and face-to-face interviewing. J. Food Prod. Mark. 11, 107–134.
Chanel, O., Makhloufi, K., Abu-Zaineh, M., 2016. Can a circular payment card format effectively elicit preferences? Evidence from a survey on a mandatory health insurance scheme in Tunisia. Appl. Health Econ. Health Policy 1–14.
Chattopadhyay, S., 1999. Estimating the demand for air quality: new evidence based on the Chicago housing market. Land Econ. 22–38.
Chen, S., He, L., 2014. Welfare loss of China's air pollution: how to make personal vehicle transportation policy. China Econ. Rev. 31, 106–118.
Chestnut, L.G., Ostro, B.D., Vichit Vadakan, N., 1997. Transferability of air pollution control health benefits estimates from the United States to developing countries: evidence from the Bangkok study. Am. J. Agric. Econ. 79, 1630–1635.
Cooper, J.C., 1994. A comparison of approaches to calculating confidence intervals for benefit measures from dichotomous choice contingent valuation surveys. Land Econ. 111–122.
Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.
Daly, H.E., 1992. Allocation, distribution, and scale: towards an economics that is efficient, just, and sustainable. Ecol. Econ. 6, 185–193.
Deco, G., Scarano, L., Soto-Faraco, S., 2007. Weber's law in decision making: integrating behavioral data in humans with a neurophysiological model. J. Neurosci. 27, 11192–11200.
Desaigues, B., Ami, D., Bartczak, A., Braun-Kohlová, M., Chilton, S., Czajkowski, M., Farreras, V., Hunt, A., Hutchison, M., Jeanrenaud, C., 2011. Economic valuation of air pollution mortality: a 9-country contingent valuation survey of value of a life year (VOLY). Ecol. Indic. 11, 902–910.
Duan, N., Manning, W.G., Morris, C.N., Newhouse, J.P., 1983. A comparison of alternative models for the demand for medical care. J. Bus. Econ. Stat. 1, 115–126.
Evans, G.W., Colome, S.D., Shearer, D.F., 1988. Psychological reactions to air pollution. Environ. Res. 45, 1–15.
Filippini, M., Martínez-Cruz, A.L., 2016. Impact of environmental and social attitudes, and family concerns on willingness to pay for improved air quality: a contingent valuation application in Mexico City. Lat. Am. Econ. Rev. 25, 7.
Fischhoff, B., Furby, L., 1988. Measuring values: a conceptual framework for interpreting transactions with special reference to contingent valuation of visibility. J. Risk Uncertain. 1, 147–184.
Fisher, R.J., 1993. Social desirability bias and the validity of indirect questioning. J. Consum. Res. 20, 303–315.
Freeman III, A.M., Herriges, J.A., Kling, C.L., 2014. The Measurement of Environmental and Resource Values: Theory and Methods. Routledge.
Grömping, U., 2009. Variable importance assessment in regression: linear regression versus random forest. Am. Stat. 63, 308–319.
Guan, D., Su, X., Zhang, Q., Peters, G.P., Liu, Z., Lei, Y., He, K., 2014. The socioeconomic drivers of China's primary PM2.5 emissions. Environ. Res. Lett. 9, 024010.
Guxens, M., Sunyer Deu, J., 2012. A review of epidemiological studies on neuropsychological effects of air pollution. Swiss Med. Wkly. 141, w13322 (2012).
Haab, T.C., McConnell, K.E., 1997. Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation. Edward Elgar Publishing Limited.
Haab, T.C., McConnell, K.E., 2002. Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation. Edward Elgar Publishing.
Hammitt, J.K., Zhou, Y., 2006. The economic value of air-pollution-related health risks in China: a contingent valuation study. Environ. Rec. Econ. 33, 399–423.
Hanemann, W.M., 1991. Willingness to pay and willingness to accept: how much can they differ? Am. Econ. Rev. 81, 635–647.
Hanley, N., Barbier, E.B., Barbier, E., 2009. Pricing Nature: Cost-Benefit Analysis and Environmental Policy. Edward Elgar Publishing.
Hoveidi, H., Aslemand, A., Vahidi, H., Limodehi, F., 2013. Cost emission of PM10 on human health due to the solid waste disposal scenarios, case study: Tehran, Iran. J. Earth Sci. Clim. Change 4, 139–143.
Howarth, R.B., Farber, S., 2002. Accounting for the value of ecosystem services. Ecol. Econ. 41, 421–429.
Istamto, T., Houthuijs, D., Lebret, E., 2014a. Willingness to pay to avoid health risks from road-traffic-related air pollution and noise across five countries. Sci. Total Environ. 497-498, 420–429.
Istamto, T., Houthuijs, D., Lebret, E., 2014b. Willingness to pay to avoid health risks from road-traffic-related air pollution and noise across five countries. Sci. Total Environ. 497, 420–429.

Ito, K., Zhang, S., 2016. Willingness to Pay for Clean Air: Evidence From Air Purifier Markets in China. https://doi.org/10.3386/w22367.
Jiang, P., Wu, H., Wang, W., Ma, W., Sun, X., Lu, Z., 2007. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. Nucleic Acids Res. 35, 339–344.
Johnston, R.J., Boyle, K., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R., 2017. Contemporary guidance for stated preference studies. J. Assoc. Environ. Econ. 4, 319–405.
Kan, H., Chen, B., 2004. Particulate air pollution in urban areas of Shanghai, China: health-based economic assessment. Sci. Total Environ. 322, 71–79.
Khalilia, M., Chakraborty, S., Popescu, M., 2011. Predicting disease risks from highly imbalanced data using random forest. BMC Med. Inform. Decis. Mak. 11, 51.
Krosnick, J., 1999. Maximising Measurement Quality: Principles of Good Questionnaire Design. Academic Press, New York.
Lamers, F., Van Oppen, P., Comijs, H.C., Smit, J.H., Spinhoven, P., Van Balkom, A.J., Nolen, W.A., Zitman, F.G., Beekman, A.T., Penninx, B.W., 2011. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). J. Clin. Psychiatry 72 (1), 478-348.
Lee, Y.J., Lim, Y.W., Yang, J.Y., Kim, C.S., Shin, Y.C., Shin, D.C., 2011. Evaluating the PM damage cost due to urban air pollution and vehicle emissions in Seoul, Korea. J. Environ. Manag. 92, 603–609.
Leggett, C.G., Kleckner, N.S., Boyle, K.J., Dufield, J.W., Mitchell, R.C., 2003. Social desirability bias in contingent valuation surveys administered through in-person interviews. Land Econ. 79, 561–575.
Leshowitz, B., Tavb, H.B., Raab, D.H., 1968. Visual detection of signals in the presence of continuous and pulsed backgrounds. Percept. Psychophys. 4, 207–213.
Liu, Y., Zhang, T., Liu, Q., Zhang, R., Sun, Z., Zhang, M., 2014. Seasonal variation of physical and chemical properties in TSP, PM10 and PM2. 5 at a roadside site in Beijing and their influence on atmospheric visibility. Aerosol Air Qual. Res. 14, 954–969.
Ma, X., 2005. Growth in mathematics achievement: analysis with classification and regression trees. J. Educ. Res. 99, 78–86.
Madaniyazi, L., Nagashima, T., Guo, Y., Yu, W., Tong, S., 2015. Projecting fine particulate matter-related mortality in East China. Environ. Sci. Technol. 49, 11141–11150.
Maler, K.-G., Vincent, J.R., 2005. Handbook of Environmental Economics: Valuing Environmental Changes. Elsevier.
Meerkerk, G.J., Van Den Eijnden, R.J., Vermulst, A.A., Garretsen, H.F., 2009. The compulsive internet use scale (CIUS): some psychometric properties. Cyberpsychol. Behav. 12, 1–6.
Mitchell, R., Carson, R., 1981. an experiment in measuring willingness to pay for intrinsic water pollution control benefits. https://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0011-01.pdf/$file/EE-0011-01.pdf.
Morrison, G.C., 1998. Understanding the disparity between WTP and WTA: endowment effect, substitutability, or imprecise preferences? Econ. Lett. 59, 189–194.
Morrison, M.D., Blamey, R.K., Bennett, J.W., 2000. Minimising payment vehicle bias in contingent valuation studies. Environ. Rec. Econ. 16, 407–422.
MusicSurvey, 2015, February 10. The music survey webpage. https://www.lediaocha.com/.
Mutanga, O., Adam, E., Cho, M.A., 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. Int. J. Appl. Earth Obs. Geoinf. 18, 399–406.
Nagashima, F., Kagawa, S., Suh, S., Nansai, K., Moran, D., 2017. Identifying critical supply chain paths and key sectors for mitigating primary carbonaceous PM2. 5 mortality in Asia. Econ.Syst. Res. 29, 105–123.
Nam, Y., Koh, S.-H., Jeon, S.-J., Youn, H.-J., Park, Y.-S., Choi, W.I., 2015. Hazard rating of coastal pine forests for a black pine bast scale using self-organizing map (SOM) and random forest approaches. Eco. Inform. 29, 206–213.
Navrud, S., 2001. Valuing health impacts from air pollution in Europe. Environ. Rec. Econ. 20, 305–329.
Nielsen, J.S., 2011. Use of the Internet for willingness-to-pay surveys: a comparison of face-to-face and web-based interviews. Resour. Energy Econ. 33, 119–129.
Panek, W., Stevens, S., 1966. Saturation of red: a prothetic continuum. Percept. Psychophys. 1, 59–66.
Pearce, D., 2002. The Role of 'Property Rights' in Determining Economic Values for Environmental Costs and Benefits, the Environment Agency Bristol (10.1.1.506.4709).
Peters, A., Veronesi, B., Calderón-Garcidueñas, L., Gehr, P., Chen, L.C., Geiser, M., Reed, W., Rothen-Rutishauser, B., Schürch, S., Schulz, H., 2006. Translocation and potential neurological effects of fine and ultrafine particles a critical update. Part. Fibre Toxicol. 3, 13.
Power, M.C., Kioumourtzoglou, M.-A., Hart, J.E., Okereke, O.I., Laden, F., Weisskopf, M.G., 2015. The relation between past exposure to fine particulate air pollution and prevalent anxiety: observational cohort study. BMJ 350, 11–16.
PRCMEP, 2012. Ambient Air Quality Standards GB 3095-2012. China Environmental Science Publisher.
PRCMEP, 2017. Air quality data services. http://datacenter.mep.gov.cn/index.
Rabl, A., Spadaro, J.V., 2000. Public health impact of air pollution and implications for the energy system. Annu. Rev. Energy Environ. 25, 601–627.
Ready, R.C., Navrud, S., Dubourg, W.R., 2001. How do respondents with uncertain willingness to pay answer contingent valuation questions? Land Econ. 77, 315–326.
Rowe, R.D., Schulze, W.D., Breffle, W.S., 1996. A test for payment card biases. J. Environ. Econ. Manag. 31, 178–185.
Samuelson, P.A., 1954. The pure theory of public expenditure. Rev. Econ. Stat. 36, 387–389.
Sanders, J.L., Brizzolara, M.S., 1982. Relationships between weather and mood. J. Gen. Psychol. 107, 155–156.
Sayman, S., Öncüler, A., 2005. Effects of study design characteristics on the WTA–WTP disparity: a meta analytical framework. J. Econ. Psychol. 26, 289–312.
Shah, S.A., Hoag, D.L.K., Loomis, J., 2017. Is willingness to pay for freshwater quality improvement in Pakistan affected by payment vehicle? Donations, mandatory government payments, or donations to NGO's. Environ. Econ. Policy Stud. 19, 807–818.
Shen, G.F., Yuan, S.Y., Xie, Y.N., Xia, S.J., Li, L., Yao, Y.K., Qiao, Y.Z., Zhang, J., Zhao, Q.Y., Ding, A.J., 2014. Ambient levels and temporal variations of PM2. 5 and PM10 at a residential site in the mega-city, Nanjing, in the western Yangtze River Delta, China. J. Environ. Sci. Health A 49, 171–178.
Sisler, J.F., Malm, W.C., 1994. The relative importance of soluble aerosols to spatial and seasonal trends of impaired visibility in the United States. Atmos. Environ. 28, 851–862.
Strazzera, E., Genius, M., Scarpa, R., Hutchinson, G., 2003. The effect of protest votes on the estimates of WTP for use values of recreational sites. Environ. Rec. Econ. 25, 461–476.
Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinforma. 8, 1.
Suades-González, E., Gascon, M., Guxens, M., Sunyer, J., 2015. Air pollution and neuropsychological development: a review of the latest evidence. Endocrinology 156, 3473–3482.
Sun, Y., 2012. Reductions of PM2. 5 in Beijing-Tianjin-Hebei urban agglomerations during the 2008 Olympic Games. Adv. Atmos. Sci. 29, 1330–1342.
Sun, Z., An, X., Tao, Y., Hou, Q., 2013. Assessment of population exposure to PM10 for respiratory disease in Lanzhou (China) and its health-related economic costs based on GIS. BMC Public Health 13, 891.
Sun, C., Yuan, X., Xu, M., 2016a. The public perceptions and willingness to pay: from the perspective of the smog crisis in China. J. Clean. Prod. 112, 1635–1644.
Sun, C., Yuan, X., Yao, X., 2016b. Social acceptance towards the air pollution in China: evidence from public's willingness to pay for smog mitigation. Energy Policy 92, 313–324.
Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random forest: a classification and regression tool for compound classification and QSAR modeling. J. Chem. Inf. Comput. Sci. 43, 1947–1958.
Szolnoki, G., Hoffmann, D., 2013. Online, face-to-face and telephone surveys—comparing different sampling methods in wine consumer research. Wine Econ. Policy 2, 57–66.
Van Exel, N., Brouwer, W., van den Berg, B., Koopmanschap, M., 2006. With a little help from an anchor: discussion and evidence of anchoring effects in contingent valuation. J. Socio-Econ. 35, 836–853.
Wang, H., Mullahy, J., 2006. Willingness to pay for reducing fatal risk by improving air quality: a contingent valuation study in Chongqing, China. Sci. Total Environ. 367, 50–57.
Wang, Y., Zhang, Y.-S., 2009. Air quality assessment by contingent valuation in Ji'nan, China. J. Environ. Manag. 90, 1022–1029.
Wang, X., Zhang, W., Li, Y., Yang, K., Bai, M., 2006. Air quality improvement estimation and assessment using contingent valuation method, a case study in Beijing. Environ. Monit. Assess. 120, 153–168.
Wang, K., Wu, J., Wang, R., Yang, Y., Chen, R., Maddock, J.E., Lu, Y., 2015. Analysis of residents' willingness to pay to reduce air pollution to improve children's health in community and hospital settings in Shanghai, China. Sci. Total Environ. 533, 283–289.
Watson, J.G., Chow, J.C., 1994. Clear sky visibility as a challenge for society. Annu. Rev. Energy Environ. 19, 241–266.
Wechat, 2015, April 22. Wechat webpage. https://wx.qq.com/.
Wei, W., Wu, Y., 2016. Willingness to pay to control PM2. 5 pollution in Jing-Jin-Ji Region, China. Appl. Econ. Lett. 1–9.
Wei, W., Wu, Y., 2017. Willingness to pay to control PM2. 5 pollution in Jing-Jin-Ji Region, China. Appl. Econ. Lett. 24, 753–761.
Weibo, 2015, April 20. Weibo webpage. http://weibo.com/.
Weidema, B.P., 2009. Using the budget constraint to monetarise impact assessment results. Ecol. Econ. 68, 1591–1598.
Wu, C.-F.J., 1986. Jackknife, bootstrap and other resampling methods in regression analysis. Ann. Stat. 1261–1295.
Wu, R., Dai, H., Geng, Y., Xie, Y., Masui, T., Liu, Z., Qian, Y., 2017. Economic Impacts from PM2. 5 Pollution-Related Health Effect: A Case Study in Shanghai. Environmental Science & Technology.
Xie, Y., Dai, H., Dong, H., Hanaoka, T., Masui, T., 2016a. Economic impacts from PM2.5 pollution-related health effects in China: a provincial-level analysis. Environ. Sci. Technol. 50, 4836–4843.
Xie, Y., Dai, H., Dong, H., Hanaoka, T., Masui, T., 2016b. Economic impacts from PM2. 5 pollution-related health effects in China: a provincial-level analysis. Environ. Sci. Technol. 50, 4836–4843.
Yin, H., Pizzol, M., Xu, L., 2017. External costs of PM2. 5 pollution in Beijing, China: Uncertainty analysis of multiple health impacts and costs. Environ. Pollut. 226, 356–369.
Zhang, X., Zhang, X., Chen, X., 2017. Valuing air quality using happiness data: the case of China. Ecol. Econ. 137, 29–36.
Zhao, P., Dong, F., He, D., Zhao, X., Zhang, X., Zhang, W., Yao, Q., Liu, H., 2013. Characteristics of concentrations and chemical compositions for PM2.5 in the region of Beijing, Tianjin, and Hebei, China. Atmos. Chem. Phys. 13, 4631–4644.
Zijlema, W., Wolf, K., Emeny, R., Ladwig, K., Peters, A., Kongsgård, H., Hveem, K., Kvaløy, K., Yli-Tuomi, T., Partonen, T., 2016. The association of air pollution and depressed mood in 70,928 individuals from four European cohorts. Int. J. Hyg. Environ. Health 219, 212–219.