University of Warsaw
Faculty of Economic Sciences

Working Papers
No. 33/2026 (527)

Does the Messenger Matter? Scenario Credibility and Information Source in Valuing Health Risk Reductions from Air Pollution

Anna Maria Boros
Anna Małgorzata Bartczak

Warsaw 2026
ISSN 2957-0506

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

Does the messenger matter? Scenario credibility and information source in valuing health risk reductions from air pollution

Anna Maria Boros[1, 2], Anna Małgorzata Bartczak[1]*

[1] *University of Warsaw, Faculty of Economic Sciences*
[2] *University of Warsaw, Robert Zajonc Institute for Social Studies*
[*] *Corresponding author: a.boros@uw.edu.pl*

Abstract: Perceived realism of the policy scenario presented in stated preference (SP) valuation depends on trust in the scientific soundness of the projected policy outcomes (epistemic credibility) and in policymakers' capability to deliver them (implementation credibility). Using a discrete choice experiment (DCE) with 1,100 Polish urban residents, we examine how policy scenario credibility shapes preferences for air pollution mitigation reducing risks from chronic obstructive pulmonary disease (COPD). We contrast domain experts and artificial intelligence (AI) with no-source control as information sources on policy outcome projections. We find that epistemic and implementation credibility are strongly correlated; thus, we treat credibility as a single variable affecting preferences. Using a hybrid mixed logit (HMXL) model, we demonstrate that higher perceived credibility increases programme support, lowers cost sensitivity, and thus elevates willingness to pay (WTP) for the programme and morbidity risk reduction. We find no evidence that the information source influences preferences or scenario credibility.

Keywords: air pollution, artificial intelligence, chronic obstructive pulmonary disease, credibility, discrete choice experiment, information source

JEL codes: D91, Q51, Q53

List of acronyms:

AI	artificial intelligence
ASC_{SQ}	alternative-specific constant (status quo)
COPD	chronic obstructive pulmonary disease
DCE	discrete choice experiment
EU	European Union
HMXL	hybrid mixed logit
LV	latent variable
PM	particulate matter
SP	stated preference
SQ	status quo
WTP	willingness to pay

1. Introduction

Public policies that protect environmental quality and human health often deliver benefits that are not traded in markets and therefore carry no observable prices. To express these benefits in monetary terms, researchers can apply SP valuation methods such as contingent valuation or DCE. In these survey-based approaches, researchers construct a hypothetical market, in which respondents are asked to assess proposed policy changes through structured valuation tasks (Boxall et al., 1996; Johnston et al., 2017). The reliability of SP valuation methods depends on the realism of the described policy scenario (Johnston et al., 2017). If respondents perceive the presented policy scenario as unrealistic, this can exacerbate the hypothetical bias, i.e., the divergence between declared preferences and preferences revealed in binding financial settings (Murphy et al., 2005). Hypothetical bias introduces systematic error into obtained WTP estimates and thus fuels methodological critiques of relying on SP valuation methods (Hausman, 2012; Kling et al., 2012).

Lay respondents rarely have the knowledge or experience needed to judge for themselves whether a scenario describing complex effects of environmental change is realistic. Respondents' acceptance of policy outcome projections described in the scenario utilised in SP valuation therefore rests on trust rather than independent verification. We argue that two distinct objects of trust are at work here, and we distinguish them accordingly. The first concerns the information itself: respondents must be willing to credit the scientific basis of the outcomes that the scenario projects, which we term epistemic credibility. The second concerns delivery: even a well-founded projection materialises only if the responsible authority is capable of implementing the policy effectively. We term trust in that capability implementation credibility. These two credibility dimensions can be interrelated. Trust in implementation affects scenario acceptance only to the extent that people credit the projected outcomes in the first place: successfully executing a policy built on doubted science yields no expected value. Additionally, some individuals may exhibit a generally eroded trust in institutions of both science and governance, stemming from perceived low morality or competence among representatives of those institutions (Lunz Trujillo & Trujillo, 2026). This can result in systematic discounting of both epistemic and implementation credibility dimensions.

To date, SP studies in environmental valuation have overwhelmingly focused on how trust in implementing institutions shapes preferences (see, e.g., Birol & Das, 2012; Glenk & Fischer, 2010; Hoffmann et al., 2022; Oh & Hong, 2012; Zambrano-Monserrate, 2024).

Conversely, general epistemic trust, defined as laypeople's reliance on expert knowledge (Hardwig, 1991; Hendriks et al., 2016), remains underrepresented in SP valuation research (King, 2025). Studies explicitly examining the influence of policy scenario credibility in environmental contexts do not separately examine evaluations of its underlying dimensions or implicitly target one of them (see, e.g., Kataria et al., 2012; Welling et al., 2022). For example, Welling et al. (2022) manipulate policy scenario credibility by informing some respondents that urban green measures, which constitute attributes of the described policy intervention, belong to an officially adopted municipal climate adaptation strategy. This manipulation signals institutional commitment rather than scientific evidence of ecological effectiveness. Although the authors do not state this explicitly, the nature of their experimental manipulation suggests they target what we term implementation credibility. This emphasis fits their empirical setting: whether measures such as additional street trees or green roofs will appear in the city is a matter of execution. Kataria et al. (2012) argue that when the promised outcomes rest on more complex projections, subjective beliefs can differ from scientific information. However, they apply an epistemic credibility frame only to the status quo (SQ) description by anchoring it in scientific water quality monitoring data, while their measure of policy-scenario credibility lacks a clear anchor, leaving the nature of respondent scepticism open to interpretation.

This study investigates the role of credibility in shaping preferences and WTP for public health risk reductions from urban air pollution mitigation. In examining this role, we focus on the scenario describing projected policy outcomes. Due to the complex character of the cause-and-effect relationship between the policy intervention and its effect on public health, many respondents are likely unable to verify the outcome projections presented to them. This makes the setting well suited to asking whether the epistemic dimension of policy scenario credibility matters for valuation, how it relates to perceptions of implementation credibility, and whether the source to which the scientific evidence is attributed shapes the credibility of a policy scenario. While best-practice guidelines in SP valuation design recommend relying on domain experts to construct a credible description of a policy scenario (Johnston et al., 2017), AI has emerged as a viable tool for generating and refining policy projections. Given that trust in algorithm-based and human-based judgements can differ significantly, reflecting both relative algorithm aversion and appreciation (Dietvorst et al., 2015; Logg et al., 2019; Logg & Schlund, 2024), we experimentally evaluate whether attributing policy projections to AI vs domain experts affects preferences and scenario rejection in the form of protest responses.

To address the objectives of this study, we employ a split-sample DCE conducted in the context of urban air pollution mitigation in Poland. Polish cities consistently rank among the most polluted in the European Union (EU), with elevated levels of particulate matter (PM) driving substantial excess mortality and morbidity, particularly from cardiovascular and respiratory diseases (Burchard-Dziubińska, 2019; EEA, 2023). Among the adverse health impacts, COPD represents one of the most serious long-term consequences of exposure to PM pollution (Adeloye et al., 2022; Li et al., 2016). In COPD, airborne PM acts as a double burden: it serves as a major risk factor for the initial development of the condition and increases the risk of symptom exacerbation (Bochenek et al., 2025; Sharma et al., 2020; Sin et al., 2023). In Poland, over two million people currently live with COPD, with approximately one-fifth of those cases directly attributable to poor air quality (Institute for Health Metrics and Evaluation, 2023; Ministerstwo Zdrowia, 2026). At the same time, public awareness of COPD remains limited, even among diagnosed patients (Śliwiński & Puchalski, 2015).

Respondents in our survey evaluated hypothetical air pollution mitigation programmes aimed at reducing COPD risk across a series of choice tasks. The epistemic and implementation facets of policy scenario credibility were each evaluated using two dedicated Likert-scale indicator questions. To isolate the effect of information source, participants were randomly assigned across two experimental conditions and a baseline control condition. In the expert condition, respondents were asked to assume that projected health outcomes were based on advice from respiratory disease experts; in the AI condition, the projections were attributed exclusively to an AI model; whereas in the control condition, no information regarding the information source was provided. Importantly, the underlying policy scenario, including policy outcome projections, was identical across all conditions.

To analyse the data, we use an HMXL model, which constitutes a flexible tool for integrating unobservable psychological constructs such as credibility into a choice modelling framework (Ben-Akiva et al., 2002). This specification accommodates unobserved preference heterogeneity across individuals, relaxes the restrictive Independence of Irrelevant Alternatives assumption, and controls for the panel structure of the repeated choice data (McFadden & Train, 2000; Train, 2009). By modelling Likert scores as imperfect proxies of an underlying latent variable (LV), the HMXL model filters out measurement noise and ordinal-scale distortions from the utility function (Vij & Walker, 2016).

Therefore, this study makes two main contributions. First, to the best of our knowledge, it is the first to examine the relationship between policy scenario credibility and the valuation of public health risk reductions delivered through environmental programmes. This context is demanding on both counts: respondents must trust claims about the relatively complex epidemiological mechanisms through which the intervention is said to operate, and their confidence that the projected outcomes would actually materialise may be undermined by a record of past programmes' failed execution (Fuller et al., 2022; NIK, 2025). Establishing how far credibility considerations move valuation results in such a setting is of practical relevance for the design of SP valuation studies that inform air pollution mitigation policies, and the design of communication related to environmental policies affecting health.

Second, we contribute to the emerging literature on acceptance of AI-assisted city governance (Lartey & Law, 2025; Oleksy et al., 2023). As public authorities increasingly integrate AI into policy modelling, risk forecasting, and environmental health projections (Fu et al., 2023; Singh, 2024), understanding how the public evaluates AI-generated scientific evidence is critical for fair governance. While healthcare contexts often evoke algorithm aversion because of perceived lack of empathy or personalisation (Longoni et al., 2019; Reis et al., 2024), quantitative forecasting and data-heavy tasks often trigger algorithm appreciation (Logg et al., 2019). To the best of our knowledge, by experimentally comparing AI against human domain experts and an undisclosed baseline, we provide the first empirical test of whether environmental valuation is sensitive to the type of information source backing projected policy outcomes.

The rest of the paper is structured as follows: Section 2 reviews the existing literature regarding the role of policy scenario credibility in SP valuation research and the influence of information source on message reception. Section 3 details the survey structure, experimental design, and sample characteristics. Section 4 outlines the empirical approach and describes the modelling framework utilised for data analysis. Section 5 presents the empirical results. Finally, Section 6 offers a discussion of the findings, practical implications for survey design and policymaking, and concluding remarks.

2. Literature review

2.1. The role of scenario acceptance

Ideally, respondents in SP valuation studies should perceive the valuation scenario information as sufficient, comprehensible, and realistic (Johnston et al., 2017). To achieve this, the valuation scenario must combine a well-defined description of the SQ baseline, the policy scenario, and the payment vehicle specifying the payment channel, schedule, and the entity bearing the cost. In turn, the policy scenario should include the causal mechanism linking the policy intervention to utility-relevant outcomes, the institutional context and implementing agency, and the intervention's spatial and temporal scope. When respondents perceive the policy scenario as lacking credibility, they can reject the valuation task. In that case, respondents are likely to opt out to signal their protest response. A debriefing section after the survey can facilitate identifying respondents who resort to this reaction (Bateman, 2002). However, when the described policy intervention outcomes are met with scepticism, individuals may also respond by constructing a subjective policy scenario that integrates the provided information with their own beliefs (see, e.g., Fischhoff et al., 1999; Kataria et al., 2012; Lundhede et al., 2015; Manski, 1999). Then, the derived WTP estimates can only be associated loosely with the policy outcomes described in the policy scenario (Boyle & Bergstrom, 1999). Therefore, it is advised to consider the role of subjective perceptions of attributes in the policy scenario if those beliefs are likely to strongly affect valuation (Adamowicz et al., 1997).

In DCEs, the outcomes of the evaluated policy are decomposed into a set of attributes, whose levels are systematically varied across choice tasks according to the experimental design. As a result, researchers often face a trade-off between ensuring that all attribute-level combinations remain plausible and preserving the design's statistical efficiency (Louviere, 2006; Louviere et al., 2000). Some respondents may therefore be presented with attribute-level combinations that are subjectively unrealistic, affecting the perceived credibility of the policy scenario. Environmental valuation studies examining the perceived credibility of the valuation scenario, including the policy scenario description, remain scarce (Kataria et al., 2012; Vásquez & De Rezende, 2019; Welling et al., 2022). In the DCE conducted in the context of water quality improvement, Kataria et al. (2012) show that most respondents find the described policy outcomes unlikely or rather unlikely, while most accept the SQ description. They find that respondents who trust the intervention can produce the described outcomes are less likely to opt out and exhibit higher WTP for the environmental improvement. Similarly, for urban green measures, Welling et al. (2022) show that perceived realism of intervention outcomes is associated with both higher valuation of all policy attributes and a lower propensity to opt for SQ. Taken together, studies suggest that scenario realism might be highly doubted in some contexts, while respondents who perceive the policy scenario as realistic are more interested in implementing described interventions and value their outcomes more.

Differences in credibility perceptions reported in some studies might stem from a focus on different parts of the valuation scenario. Credibility of the SQ (or baseline scenario) concerns the accuracy of a description of current conditions and does not entail trust in policymakers' capability to execute anything in the future (see, e.g., Kataria et al., 2012; Vásquez & De Rezende, 2019). Credibility of the payment vehicle concerns, for example, whether the stated cost differs from what will actually be collected (Flores & Strong, 2007). Therefore, it is a question of institutional conduct that does not entail trust in any causal scientific mechanism. Finally, the credibility of the policy scenario differs in this regard because it simultaneously entails a scientific claim about the mechanism (will the intervention produce the stated outcome) and a claim about the execution (will the responsible authority effectively deliver this intervention) (see, e.g., Kataria et al., 2012; Welling et al., 2022). Therefore, the subjective realism of the described policy scenario can be undermined by both respondents' scepticism toward epistemic and implementation credibility.

Valuing air pollution mitigation outcomes requires respondents to accept the complex mechanism linking exposure reduction to health benefits. However, public environmental health literacy regarding air pollution is often inadequate (Finn & O'Fallon, 2019; Hou et al., 2021; Pfleger et al., 2024). Consequently, information in SP valuation studies about long-term health risk reductions is likely novel to many respondents. Because novel information cannot easily be cross-checked against prior knowledge, its acceptance depends heavily on the perceived credibility of the policy scenario. It is important to distinguish scenario credibility perceptions from subjective health risk updating, in which respondents adjust presented risk reductions or baseline risk estimates to match their self-perceived health status (see, e.g., Cameron et al., 2011; DeShazo & Cameron, 2005; Dickie et al., 2022; Krupnick et al., 2002). Whereas studies focused on risk-updating explore how health information is personalised by respondents, scenario credibility defines whether respondents accept the scientific validity and implementation feasibility of the policy scenario itself.

In environmental contexts, which are often politicised, policy scenario acceptance might be particularly undermined. Strong political identities can prompt individuals to reject both scientific claims regarding environmental issues and distrust non-aligned institutional authorities (Clarke & Evensen, 2019; Hundemer et al., 2022). As a result, the interplay between epistemic and implementation credibility of the policy scenario can be complex and sensitive to the level of politicisation surrounding the topic.

2.2. The role of information source

It has long been established that providing respondents with novel or differently phrased information on the public good can influence their choices (Thaler, 1980). The content and phrasing of information scripts have been shown to affect environmental valuation (see, e.g., Bulte et al., 2005; Czajkowski et al., 2016; Hasselström & Håkansson, 2014; Yang & Hobbs, 2020). However, investigating choice behaviour solely through the lens of information content overlooks a fundamental aspect of human cognition. Acceptance and even the processing mode of provided information are heavily contingent upon the messengers' identity (Chaiken, 1980; Hovland & Weiss, 1951). From a theoretical standpoint, rational respondents evaluate the intentions of the source before determining whether to trust and integrate the provided information, particularly when the judgement based on this information is of high importance (Crawford & Sobel, 1982; Nowak et al., 2019).

In a DCE concerning microplastic restrictions, King (2025) finds that respondents with higher pre-existing trust in experts exhibit a lower propensity to choose SQ and declare higher WTP for unfamiliar attributes. However, studies examining the credibility of the information itself, where the source is provided in the background, demonstrate conflicting results. For example, Kataria et al. (2012) show that WTP for improving water quality is higher for respondents who distrust scientific information on SQ water quality from the local environmental centre and perceive it as too optimistic. Meanwhile, in a similar context, Vásquez and De Rezende (2019) find that the credibility of scientific information regarding river quality from researchers working at a local university was not a significant predictor of preferences. These contrasting results suggest that the effect of source attribution may depend on broader regional, cultural, and institutional contexts that govern public trust in scientific communication.

Traditionally, experts have been the main source of specialised health-related knowledge. More recently, the Internet has become a primary source of information for a large proportion of individuals facing medical problems (Baumann et al., 2020; Kummervold & Wynn, 2012; Swoboda et al., 2018). Additionally, AI has transformed how health-related information is generated and delivered in the digital world, for instance, through the appearance of online chatbots (Esmaeilzadeh et al., 2025; Vaishya et al., 2023). However, individuals can be averse to algorithmically produced content (Dietvorst et al., 2015). Specifically, people tend to systematically discount health-related information upon discovering it was generated by AI compared to medical experts (Wang et al., 2026; Yokoi et al., 2021). Reis et al. (2024) demonstrate that this trust deficit persists in hybrid decision-making settings, showing that individuals are less willing to follow medical advice attributed to AI supervised by a human physician than advice attributed to the physician alone. At the same time, when respondents face novel quantitative and predictive tasks, algorithmically produced estimates are often preferred over human judgment (Logg et al., 2019). Projections of health outcomes from an environmental policy fall into both of these categories at once: they concern health, where AI-generated information tends to be discounted, and they are quantitative as well as predictive in nature, for which algorithmic output tends to be favoured. Existing literature therefore provides no clear a priori expectation regarding the influence of AI information attribution in the context of environmental health valuation.

From a broad social cognition perspective, trust in any information source is anchored in two fundamental dimensions: warmth, reflecting perceived benevolence, and competence, reflecting actual capability (Fiske et al., 2007). Crucially, recent evidence confirms that humans evaluate AI across these exact dimensions (McKee et al., 2023). Specifically, AI algorithms are perceived as warmer when their objectives align with human interests, whereas they are judged as more competent when they operate independently of human oversight. More specifically to technology adoption, Qin et al. (2025) formulated the Capability-Personalisation Framework. This framework posits that the preference for human actors or AI stems from a situational assessment of two distinct factors: the agent's capability in a given domain and the perceived necessity for personalisation. According to this framework, information recipients still prefer humans when tasks require a high degree of individual tailoring. Whether individuals prefer human expertise for curating the design of SP valuation surveys remains understudied. This question has real stakes, since such studies later justify policy decisions affecting the distribution of benefits and costs of environmental policies aimed at improving public health.

3. Survey

3.1. Structure

The survey was developed in accordance with best-practice recommendations for SP valuation studies (Johnston et al., 2017). To enhance data quality and mitigate hypothetical bias, the questionnaire commenced with a truth-telling oath, serving as a formal commitment mechanism (Jacquemet et al., 2013; Stevens et al., 2013). The oath was included to encourage honest reporting and reduce potential cost-vector effects (Börger et al., 2025; De-Magistris et al., 2016). Following the oath, the survey was structured into three sections.

The first section established the contextual foundations for the study. Respondents were presented with an information script detailing the current air quality level in Polish cities, with a particular focus on PM pollution. The script identified the primary anthropogenic sources of PM in Poland, namely residential heating and vehicular traffic, and outlined the general repercussions for human health. We elicited respondents' perceptions regarding air pollution in their cities of residence, their associated level of annoyance, and the perceived effectiveness of public air pollution mitigation programmes implemented to date. We then provided a clinical overview of COPD, based on Jin et al. (2020) and refined in line with advice from respiratory disease experts at the Institute of Tuberculosis and Lung Diseases in Warsaw. This included information on primary risk factors (exposure to smoking and air pollution), prevalence, and mortality rates within the Polish context, distinguishing between total and air pollution-related cases, comparing them to other sources of mortality risk (i.e., road accidents), as well as outlining the typical progression of the disease. Respondents then indicated their prior familiarity with the condition.

In the second section, we informed respondents that they would be presented with a series of public programmes aimed at reducing air pollution and related COPD health risks, and they would be asked to choose between them. This part of the survey included a description of the hypothetical policy scenario and the DCE. Depending on the information-source condition respondents were assigned, some received additional signal about the source of information behind the projected policy outcomes (see Section 3.2).

The final section of the survey served as a comprehensive debriefing and validation designed to evaluate the drivers behind respondents' choices. To check whether respondents correctly identified their assigned information-source condition, the section began with a manipulation check asking them to identify the source of the presented projections. To formally distinguish valid zero-utility preferences from protest responses, individuals who selected the SQ alternative ("No additional programme") in all choice sets were directed to a dedicated follow-up debriefing block on potential protest responses (see Section 3.5). Additionally, we elicited perceptions of epistemic and implementation credibility, as well as general credibility perceptions of health risk reduction estimates from different information sources (see Section 3.4). The questionnaire concluded with questions about individual and household monthly income, followed by a debriefing that fully disclosed the survey construction process.

3.2. Hypothetical scenario and experimental conditions

The design of the policy scenario regarding potential health risk reductions resulting from air pollution mitigation was based on existing research (Bartczak et al., 2024). We provided a general description of actionable measures that could be undertaken to mitigate air pollution, such as replacing old heating stoves and electrifying public transport. As the study is focused on the feasibility of achieving the presented changes, i.e., mortality and morbidity risk reductions within a given time, we purposefully limited the description of policy implementation to a broad, high-level plan. This avoids introducing excessive procedural complexity that could induce cognitive overload or trigger objections to specific operational instruments. Respondents were informed that in the absence of new interventions, risk levels for COPD-related morbidity and mortality would remain constant at the described levels, serving as the SQ baseline. In line with contemporary SP valuation guidelines (Johnston et al., 2017), we informed respondents that the survey results would be shared with decision-makers and could thus influence the design and cost of future air pollution mitigation policies.

The hypothetical programmes were defined by four attributes: the annual reduction in COPD-related deaths per 100,000 urban residents and COPD incidence per 100,000 urban residents; time before the full health risk reductions manifest; and an annual cost in the form of a local household fee. We explained these attributes in detail to ensure clarity. Respondents were notified that the programme would be introduced next year, the fee would be levied for

10 years starting next year, and programme benefits were expected to last 10 years, after which further air quality interventions might be necessary to sustain the effects. Finally, participants were reminded of their budget constraint, noting that money spent on the programme would reduce their disposable income for other expenditures.

Respondents were randomly assigned to one of two experimental conditions or to a control condition. In the first experimental condition, respondents were asked to assume that the ranges of projected health risk reductions associated with improvements in air quality have been established in consultation with respiratory disease experts. In the second experimental condition, they were asked to assume that the same estimates have been developed with the support of an AI model. In the control condition, no information was provided on the source of the information behind the projected health benefits. In both experimental conditions, information about the source of the projections was displayed at the top of each choice card presented to the respondent. To isolate the effect of the information-source condition, all respondents evaluated choice cards drawn from the same experimental design (see Section 3.3).

Additionally, within each experimental condition and the control condition, respondents were randomly allocated to one of two consequentiality elicitation conditions. Half of the respondents assessed the perceived consequentiality of the survey before completing the choice tasks, whereas the remaining half did so immediately after completing the choice tasks.[1]

3.3. Choice attributes and experimental design

The survey was designed in accordance with subject literature and evidence concerning mortality and morbidity related to air pollution in Poland (Institute for Health Metrics and Evaluation, 2023). The health risk reduction ranges were generated with the assistance of a large language model (Google, 2026) and subsequently refined through consultations with respiratory disease experts at the Institute of Tuberculosis and Lung Diseases in Warsaw. Table 1 shows attributes and their respective levels used in the DCE.

[1] This experimental manipulation constitutes part of another independent research project and is outside the scope of the present paper.

Table 1. Attribute levels used to describe choice alternatives

Description	Levels
COPD-related mortality reduction per 100,000 city inhabitants per year	1, 2, 3, 5 fewer deaths per 100,000 inhabitants
COPD-related morbidity reduction per 100,000 city inhabitants per year	10, 15, 25, 35 fewer cases per 100,000 inhabitants
Time before the implemented policy achieves an effect	3 years, 5 years, 7 years
Annual cost per household	25 zł, 50 zł, 100 zł, 300 zł, 500 zł, 800 zł

Note: 1 Euro = 4.3 zł: nominal exchange rate (National Bank of Poland, 2026).

The 36 choice sets, organised into 6 blocks, were constructed using D-efficiency design in Ngene based on priors from a pilot study, resulting in a D-error of 0.35. Each respondent faced one block of six choice cards presented in a randomised order. Every choice set comprised two alternative programmes and a SQ option, the latter representing a business-as-usual scenario with no additional cost and no change in air pollution levels (see Figure 1). Crucially, the attribute levels for alternatives presented in choice tasks were the same irrespective of the experimental or control conditions that respondents were assigned.

Figure 1. Example of a choice card used in the DCE part of the survey

	Programme A	**Programme B**	**No additional programme**
COPD deaths prevented annually in your city	Annually: ↓ **2 fewer deaths per 100,000 people.**	Annually: ↓ **3 fewer deaths per 100,000 people.**	Annually: **The same number of deaths as currently.**
COPD cases prevented annually in your city	Annually: ↓ **10 fewer cases per 100,000 people.**	Annually: ↓ **25 fewer cases per 100,000 people.**	Annually: **The same number of cases as currently.**
Time required to achieve programme effects	**3 years**	**5 years**	-
Additional annual local fee per household	**100 zł**	**300 zł**	**0 zł**
My choice	○	○	○

3.4. Credibility perceptions

Policy scenario credibility was decomposed into two dimensions to capture potentially distinct facets of respondent scepticism. Epistemic credibility was measured with two questions:

1. "How would you rate the credibility of the projected reductions in COPD cases and deaths under the presented air quality programmes?" [Plausibility of health effects]
2. "How would you rate the credibility of the projected time to reductions in COPD cases and deaths under the presented air quality programmes?" [Plausibility of time horizon]

To measure implementation credibility, respondents were asked the following two questions:

1. "How likely do you think it is that your city authorities would be capable of implementing an air quality improvement programme that delivers the promised outcomes?" [Capability to deliver outcomes]
2. "How likely do you think it is that your city authorities would be capable of implementing the programme within the announced timeline?" [Capability to meet the timeline]

Before answering the questions about implementation credibility, respondents were explicitly instructed to assume the scientific policy outcome projections were accurate. This nstruction was given to prevent epistemic scepticism from spilling over into evaluations of implementation credibility, and to facilitate isolating perceived implementation credibility. For both epistemic and implementation credibility, participants responded to questions using a 7-point Likert scale (1: "Very low"/ "Not at all likely" to 7: "Very high"/ "Very likely").

Additionally, to examine general perceptions of information source credibility in similar policy contexts, respondents across all treatment groups rated the credibility of health outcome projections related to air pollution, attributed to three distinct sources. Specifically, they assessed the credibility of projections originating from: human expertise (scientists and medical experts); standalone AI (diagnostic algorithms and advanced chatbots); and a collaborative, augmented approach in which human domain experts are actively supported by AI (scientists and medical experts supported by AI). Credibility of projections from each potential information source was rated on a 7-point Likert scale (1: “Very low” to 7: “Very high” in each case).

3.5. Protest responses

Individuals who consistently chose the SQ alternative across all choice sets were asked to evaluate 13 statements capturing underlying motivation for such choices, presented in randomised order. Following the established literature in environmental economics, these statements were formulated to distinguish between valid economic preferences, encompassing zero valuations due to a lack of interest in the policy or binding budget constraints, and various potential motivations signalling protest behaviour. Respondents were asked to indicate the degree of agreement with each motivation on a 7-point Likert scale (1: “Strongly disagree” to 7: “Strongly agree” in each case). Respondents were classified as protesters unless they expressed strong agreement (a rating of 6 or 7) with at least one of the two statements representing valid economic preferences (“Reducing the health risks associated with COPD is not important to me”, “My household cannot afford an additional annual fee”).

To capture the potential influence of credibility-related protest motivations, we included statements signalling scepticism towards scientific foundations of the hypothetical scenario (“Estimates regarding the number of COPD cases and deaths that the programme could prevent were not credible to me”) and institutional capability to effectively execute the policy (“My city's authorities would not be able to implement the programmes, even if it is technically possible”). The full list of statements used to capture protest intentions can be found in the Appendix (Table A1).

3.6. Data collection and sample

The survey was administered online by a professional polling agency in June 2026 to a sample of the Polish urban population living in 37 cities with over 100,000 residents. The sample was drawn using a quota sampling approach, matching the Polish urban population in terms of age, gender, educational level, and strata based on city size (among cities with over 100,000 inhabitants). Prior to conducting the final study, we pre-tested the survey in several qualitative interviews and piloted it with 50 respondents from the target population to make final refinements. The study protocol was approved by the Research Ethics Committee of [anonymised for peer review] (Approval No. 6/26).

A total of 1,100 completed questionnaires were collected. The final sample reflects the characteristics of the population of Polish cities with over 100,000 residents in terms of age, gender, and city size. With respect to the highest attained level of education, respondents with secondary education were slightly overrepresented (48% of the full sample of 1,100 respondents compared with 40% in the reference population; GUS, 2023). This slight overrepresentation is unlikely to affect our conclusions: because the study does not aim to conduct a cost-benefit analysis, which would require a fully representative sample, a modest imbalance in educational attainment does not compromise its objectives.

After identifying and excluding protest responses (N = 105), the final sample comprised 995 respondents[2]. A chi-square test of independence revealed no statistically significant correlation between the information-source condition and the likelihood of exhibiting a protest response ($\chi^2 = 0.18$, df = 2, p = .91). Among protesters, 14% expressed strong or very strong agreement with epistemic-credibility-related motivations, and 13% with implementation-credibility-related motivations (see Appendix, Table A2 for distribution of responses to protest statements). There was a weak, positive, and statistically significant Spearman rank correlation between the ratings of the two credibility-related protest motivations (0.32, p < .001). The protesting individuals did not differ from the respondents in the final sample in terms of gender and age. Among protesters, there was a significantly higher proportion of individuals with higher education ($\chi^2 = 6.55$, p < .05) and from cities with over 500,000 residents ($\chi^2 = 4.60$, p < .05).

Table 2 presents the socio-demographic characteristics of the final sample, both overall and by information-source condition. A one-way analysis of variance (ANOVA) revealed a statistically significant difference in age across the two experimental conditions and one control condition (F = 5.20, p < .01). This outcome indicates that the randomisation process did not yield perfectly balanced groups; therefore, we include age as a control variable in further analysis. Chi-square tests for the remaining categorical variables revealed no statistically significant differences between the two experimental conditions and the control condition.

---

[2] Although we exclude protest responses in the final analysis to prevent distortion of the utility parameters, this screening may potentially introduce sample selection bias (Meyerhoff et al., 2014). However, the model estimated on the full sample suggests that protest-response screening does not affect interpretation of the main study results (Appendix, Tables A7-A9).

Table 2. Descriptive socio-economic statistics of the final sample (N= 995)

	Total	Experimental condition		Control condition
		Expert	AI	
Number of respondents	995	329	328	338
Female (%)	52.9	53.8	55.8	49.1
Age (years)	49.0	47.7	48.0	51.2
Highest educational attainment (%)				
Primary and vocational	10.4	9.7	10.4	11.3
Secondary	49.1	46.5	46.6	54.1
Higher	40.4	43.8	43.0	34.6
City population size (%)				
More than 500,000 residents	42.0	39.8	39.9	45.9
Net monthly individual income (zł)	5782	5599	5786	5940
Net monthly household income (zł)	9535	9273	9886	9432

Note: 1 Euro = 4.3 zł: nominal exchange rate (National Bank of Poland, 2026).

4. Econometric Approach

To incorporate respondents' credibility beliefs into the analysis of their preferences, we employ an HMXL framework based on the structural equation approach (Ben-Akiva et al., 2002). The model comprises three components: a choice component, a structural component, and a measurement component. The choice component is theoretically grounded in Random Utility Theory (RUT, McFadden, 1973) and Lancaster's Consumer Theory (Lancaster, 1966). Under Lancaster's framework (1966), individuals do not derive utility from a good in its entirety, but rather from its constituent attributes. RUT further posits that individuals aim to maximise utility, which can be decomposed into a deterministic part explained by observable characteristics (such as attributes of choice and the individual's socio-demographic characteristics) and a stochastic term representing unobserved factors. The mixed logit approach utilised in HMXL relaxes the standard multinomial logit model assumption of the Independence of Irrelevant Alternatives, allowing for random taste variation, non-proportional substitution patterns between the alternatives, and correlation in unobserved factors (Train, 2009). Additionally, the measurement equations in the HMXL model allow for capturing the effect of directly unobservable LVs on indicator questions, while structural equations explain the influence that directly observable variables have on LVs. By combining all three components, the HMXL approach allows estimation of the influence that observable variables and LVs have on respondent preferences.

To verify whether the epistemic and implementation credibility dimensions (see Section 3.4) should be treated as separate LVs in the model, we examine the correlations between responses to the indicator questions through which they are operationalised. Modelling two correlated LVs as independent constructs risks each partly absorbing variation that reflects the other. Because the indicator questions operationalising the two credibility dimensions are relatively highly correlated (see Section 5.1), we specify a single latent credibility construct that captures general policy scenario credibility.

The structural equation relates general policy scenario credibility to the information-source condition and age. The measurement equations link the latent credibility variable to the indicator questions capturing the underlying, unobservable credibility beliefs. All four indicators, covering both the epistemic and implementation dimensions, load on the single credibility construct.

The entire system of equations is estimated jointly using the Maximum Simulated Likelihood approach with 10,000 scrambled Sobol draws (Czajkowski & Budziński, 2019). Models are estimated in MATLAB using a dedicated package for DCE analysis[3]. To ensure numerical stability, mitigate simulation error, and avoid local maxima, we tested various starting values. Sections 4.1, 4.2, and 4.3 describe model components in detail.

4.1. Choice component

The expected utility U_{ijt} of individual i from choosing alternative j in the choice situation t can be expressed as

$$U_{ijt} = \mathbf{X}_{ijt}\boldsymbol{\beta}_i + e_{ijt}, \tag{1}$$

[3] Accessible on GitHub (https://github.com/czaj/DCE) under a CC BY 4.0 license.

where: $\boldsymbol{X}_{ijt}$ denotes a vector of attributes including health-risk related attributes, an attribute describing the number of years before policy has an effect, an alternative specific constant assigned to the SQ alternative (ASC_{SQ}), and the programme cost; $\boldsymbol{\beta}_i$ denotes a vector of individual-specific preference parameters (marginal utilities of non-monetary programme characteristics, and income, respectively); and the error term e_{ijt} captures unobserved factors influencing choices. This specification allows for heterogeneity in preferences across respondents. Rather than estimating the random parameters separately for each individual, we follow standard practice and assume they are drawn from an underlying multivariate distribution permitting non-zero correlation among the parameters (Train, 2009). We assume that e_{ijt} is independently and identically distributed (i.i.d.) according to a Type I Extreme Value distribution with constant variance $\frac{\pi^2}{6}$, which yields the conventional logit choice probabilities.

Additionally, we allow the taste parameters to vary systematically with the LV and observable variables. Consistent with standard practice, we assume the elements of $\boldsymbol{\beta}_i$ denoting marginal utilities from non-monetary attributes, follow a normal distribution, while we impose a log-normal distribution on marginal utility from income, which yields

$$\boldsymbol{\beta}_i^N = \boldsymbol{\alpha}_0^N + \boldsymbol{\alpha}_1^N LV_i + \boldsymbol{\alpha}_2^N \boldsymbol{Inf}_i + \boldsymbol{\alpha}_3^N Age_i + \boldsymbol{\xi}_i^N, \tag{2}$$

and

$$\boldsymbol{\beta}_i^{LN} = \exp(\boldsymbol{\alpha}_0^{LN} + \boldsymbol{\alpha}_1^{LN} LV_i + \boldsymbol{\alpha}_2^{LN} \boldsymbol{Inf}_i + \boldsymbol{\alpha}_3^{LN} Age_i + \boldsymbol{\xi}_i^{LN}), \tag{3}$$

where: $\boldsymbol{\alpha}^N$ and $\boldsymbol{\alpha}^{LN}$ are vectors of parameters to be estimated, LV_i denotes latent policy scenario credibility, $\boldsymbol{Inf}_i$ is a vector of dummy variables indicating the respondent's assignment to one of the experimental conditions, with the control condition as the reference category; Age_i denotes respondent's centred age (included due to imbalance in respondents' age between information-source conditions), and $\boldsymbol{\xi}_i^N$ and $\boldsymbol{\xi}_i^{LN}$ are stochastic components following a multivariate normal distribution with a vector of means and a covariance matrix to be estimated.

4.2. Measurement equation component

Since policy scenario credibility is not directly observable, we infer it from the four indicator questions described in Section 3.4: two that capture epistemic credibility (plausibility of health effects and plausibility of the time horizon) and two that capture implementation credibility (capability to deliver outcomes and capability to meet the timeline). Given the ordinal nature of the Likert-scale responses, we specify the measurement component as a set of ordered probit equations. The relationship between the responses to the indicator questions and latent credibility is given by

$$I_{i,n}^{*} = \zeta_n \, LV_i + \eta_{i,n}, \tag{4}$$

where: ζ_n is the coefficient to be estimated, which defines how strongly the LV affects the answer to the n-th indicator question on a k -point Likert scale, $\eta_{i,n}$ is a measurement error term following a standard normal distribution, and a set of $k+1$ coefficients t such that $I_{i,n} = l$ if $I_{i,n}^{*} \in (t_{l-1}^{n}, t_{l}^{n}]$ links the unobserved variable, $I_{i,n}^{*}$, to the observed indicator question response $I_{i,n}$. The probability of a response pattern across all indicator questions can be expressed as

$$P(\boldsymbol{I_i} \,|LV_i\,, \boldsymbol{\zeta}, \boldsymbol{t}\,) = \prod_{n} [\Phi(t_l^n - \zeta_n LV_i) - \Phi(t_{l-1}^n - \zeta_n LV_i)] \,, \tag{5}$$

where Φ is a cumulative distribution function of the normal distribution (with an assumption that $t_0^n = -\infty$ and $t_7^n = \infty$). The mean of the LV is normalised to zero, and its variance is normalised to one to ensure identification (Daly et al., 2012).

4.3. Structural component

The structural component of the HMXL model captures the causal effect of the experimental conditions, relative to the control condition, on respondents' unobservable perceptions of policy scenario credibility. We also control for age, the only socio-demographic characteristic that differs across the experimental conditions (see Table 2). The structural equation takes the following form

$$LV_i = \boldsymbol{Inf}_i \boldsymbol{\gamma}_1 + Age_i \boldsymbol{\gamma}_2 + \varepsilon_i, \tag{6}$$

where $\boldsymbol{\gamma}$ is the vector of coefficients to be estimated, and ε_i is a stochastic term following a standardised normal distribution.

The overall formula for the likelihood function can therefore be denoted as

$$L_i = \int P(\boldsymbol{y}_i \,|\, \boldsymbol{\alpha},\ LV_i,\ \boldsymbol{\xi}_i)\, P(\,\boldsymbol{I_i}\,|LV_i\,, \boldsymbol{\zeta}, \boldsymbol{t}\,) f(LV_i|\boldsymbol{\gamma}) g(\boldsymbol{\xi}_i|\boldsymbol{\Sigma}) dLV_i d\boldsymbol{\xi}_i, \tag{7}$$

where $P(\boldsymbol{y}_i \mid \boldsymbol{\alpha},\ LV_i,\ \boldsymbol{\xi}_i)$ denotes the probability of respondent i making a sequence of choices y, LV_i and stochastic components $\boldsymbol{\xi}_i$ are unobserved directly and need to be integrated out to derive the unconditional probability, the $f(LV_i|\boldsymbol{\gamma})$ is the probability distribution function of the LV, and $g(\boldsymbol{\xi}_i|\boldsymbol{\Sigma})$ is the probability distribution function of random parameters, where correlation matrix $\boldsymbol{\Sigma}$ needs to be estimated. Because the likelihood function has no closed form, we employ maximum simulated likelihood techniques.

5. Results

5.1. Dimensionality of scenario credibility

A Wilcoxon signed-rank test reveals that epistemic credibility of the policy scenario is rated significantly higher than implementation credibility (see Table 3). This holds for both the credibility of described health-related effects ($V = 112053$, $p < .001$) and the timeframe required to achieve them ($V = 110782$, $p < .001$). However, both dimensions of credibility share a common conceptual core, reflected in moderate to strong significant Spearman rank correlations between their indicators (ranging from 0.45 to 0.53, all $p < .001$). Because of these high correlations, we treat both credibility dimensions jointly and formulate a single latent policy scenario credibility construct based on all four indicator questions, as described in Section 3.4 and Section 4.

Table 3. Descriptive statistics of indicator questions related to epistemic and implementation credibility in the final sample (N = 995)

	Mean	Median	Min	Max
Epistemic credibility				
Plausibility of health effects	4.17	4	1	7
Plausibility of time horizon	4.06	4	1	7
Implementation credibility				
Capability to deliver outcomes	3.89	4	1	7
Capability to meet the timeline	3.76	4	1	7

Note: Responses were rated on a 7-point Likert scale ranging from 1 to 7, where higher values indicate higher perceived credibility.

5.2. HMXL model: preferences and WTP

The influence of choice attributes describing projected policy outcomes on utility from different alternatives is assumed to be linear, while the marginal utilities are modelled as fully correlated random parameters. Coefficients of all non-monetary attributes are assumed to follow a normal distribution. The cost coefficient is assumed to be distributed log-normally, which imposes a positive marginal utility of income. Table 4 presents the estimation results for the choice component of the model and the model diagnostics.

Table 4. HMXL model results: the choice component

	Dist.	Mean effect	Standard deviation effect
		Coef. (st.err)	Coef. (st.err)
Baseline model			
ASC_{SQ}	n	-5.83*** (0.72)	6.86*** (0.63)
Mortality	n	0.23*** (0.05)	0.38*** (0.05)
Morbidity/10	n	0.45*** (0.09)	0.75*** (0.08)
Delay	n	-0.08* (0.04)	0.15*** (0.04)
-Cost/100	ln	0.37* (0.18)	2.07*** (0.11)
Effect of LV: Scenario credibility			
ASC_{SQ}		-2.18*** (0.43)	
Mortality		-0.02 (0.03)	
Morbidity/10		0.03 (0.06)	
Delay		-0.01 (0.02)	
-Cost/100		-0.74*** (0.12)	
Effect of information-source condition: AI			
ASC_{SQ}		-0.59 (0.93)	
Mortality		-0.06 (0.07)	
Morbidity/10		-0.04 (0.11)	
Delay		0.03 (0.05)	

	Dist.	Mean effect	Standard deviation effect
		Coef. (st.err)	Coef. (st.err)
-Cost/100		-0.19 (0.23)	
Effect of information-source condition: Expert			
ASC_{SQ}		0.40 (0.91)	
Mortality		-0.06 (0.07)	
Morbidity/10		-0.08 (0.12)	
Delay		0.04 (0.05)	
-Cost/100		-0.23 (0.25)	
Effect of socio-demographic variable: Age			
ASC_{SQ}		0.04 (0.25)	
Mortality		0.03 (0.02)	
Morbidity/10		0.06 (0.03)	
Delay		0.00 (0.01)	
-Cost/100		0.22*** (0.06)	
Model diagnostics			
LL at convergence	-9701.21		
LL at constant(s) only	-13324.76		
McFadden's pseudo-R^2	0.27		
Ben-Akiva-Lerman's pseudo-R^2	0.27		
AIC/n	3.27		
BIC/n	3.35		
No. of respondents (n)	995		
No. of parameters	71		

*Note: ***, **, and * indicate 0.1%, 1%, and 5% significance levels, respectively; n stands for normal distribution, and ln stands for log-normal distribution. ASC_{SQ} denotes the alternative-specific constant introduced for SQ option. All model components are estimated jointly using the Maximum Simulated Likelihood approach with 10,000 scrambled Sobol draws. Annual cost per household is denoted in 100 zł.*

The negative and highly significant mean coefficient on the ASC_{SQ} indicates an overall preference for policy intervention, reflecting general support among respondents for the implementation of air pollution mitigation programmes. Turning to the health-related attributes, respondents derive positive marginal utility from reductions in COPD-related mortality and morbidity risks. The negative and significant coefficient on implementation delay indicates that respondents prefer immediate to delayed risk reductions. As specified in eq. (2) and eq. (3), the model includes several covariates intended to explain preference heterogeneity. Even with these covariates in place, the estimated standard deviations of the random parameters remain sizeable and statistically significant, indicating that a substantial share of preference heterogeneity remains unobserved.

Turning to the main objective of our study, we examine how perceived policy scenario credibility shapes preferences. Lower values of the LV reflect weaker perceived credibility of the policy scenario, whereas higher values indicate stronger beliefs in both the scientific foundations of the projected health effects and the institutional capability to deliver the programme. The interactions between the LV and the choice attributes reveal how these beliefs translate into preferences. The interaction with the cost attribute shows that respondents who find the projected outcomes more credible are less sensitive to programme cost. The interaction with the ASC_{SQ} is in the same direction: respondents who perceive the projected health risk reductions as more credible are less inclined to choose the business-as-usual option, i.e., they derive additional utility from the introduction of an air quality improvement programme. The interaction effects between the LV and non-monetary attributes appear to be insignificant.

With respect to the second aim of our study, the effects of the information-source condition on preferences are again captured through its interactions with the attributes in the choice component. None of the experimental conditions has a significant effect on preferences relative to the control condition. Since the age composition of the subsamples differs across the information-source conditions (see Table 2), we include centred age as a control to avoid confounding its effect with that of the treatment. The interactions with centred age show that older respondents are significantly more sensitive to programme cost, which suggests a higher marginal utility of income in this group.

The results of the measurement equation part of the HMXL model indicate that the LV captures unobserved beliefs about the plausibility of the projected health effects and the time frame in which the programme outcomes would be achieved (see Table 5). All four indicators are significantly and positively related to the LV. Higher values of the LV are therefore associated with stronger beliefs in the scientific plausibility of the projected health outcomes and of their time frame, and with stronger beliefs in the capability of local authorities to carry the programme through and to do so on schedule. The results support treating credibility as unidimensional in this setting.

Table 5. HMXL model results: the measurement component. Effect of LV (policy scenario credibility)

Indicator question	Coef.	St. error
Plausibility of health effects	0.87***	0.05
Plausibility of time horizon	0.76***	0.04
Capability to deliver outcomes	2.31***	0.12
Capability to meet the timeline	3.17***	0.31

*Note: ***, **, and * indicate 0.1%, 1%, and 5% significance levels, respectively.*

We tested whether exogenous factors explain latent credibility. The structural equation estimates the effect of the information-source conditions (expert vs AI, with the no-source condition as the control) on credibility. The results of the structural component (see Table 6) indicate that neither the information-source condition nor age has a significant effect on perceived credibility.

Table 6. HMXL model results: the structural equations component. Dependent variable: LV (policy scenario credibility)

Variable	Coef.	St. error
Information-source condition: AI	0.07	0.04
Information-source condition: Expert	0.04	0.04
Age	0.01	0.04

To facilitate the economic interpretation of our findings, we report the simulated median WTP estimates and the marginal effects of credibility on WTP in Table 7. We used the Krinsky-Robb method to simulate WTP distributions and account for parameter uncertainty (Krinsky & Robb, 1986). The morbidity attribute is scaled to represent the WTP for one avoided COPD case per 100,000 urban residents. Respondents are willing to pay approximately 5.40 zł to avoid one death due to COPD per 100,000 urban residents, and 0.95 zł to prevent one new case of COPD related to air pollution per 100,000 urban residents. The effect of a one-year policy implementation delay beyond the three-year baseline on WTP turned out to be statistically insignificant.

Table 7. Median WTP and marginal effects on WTP (zł)

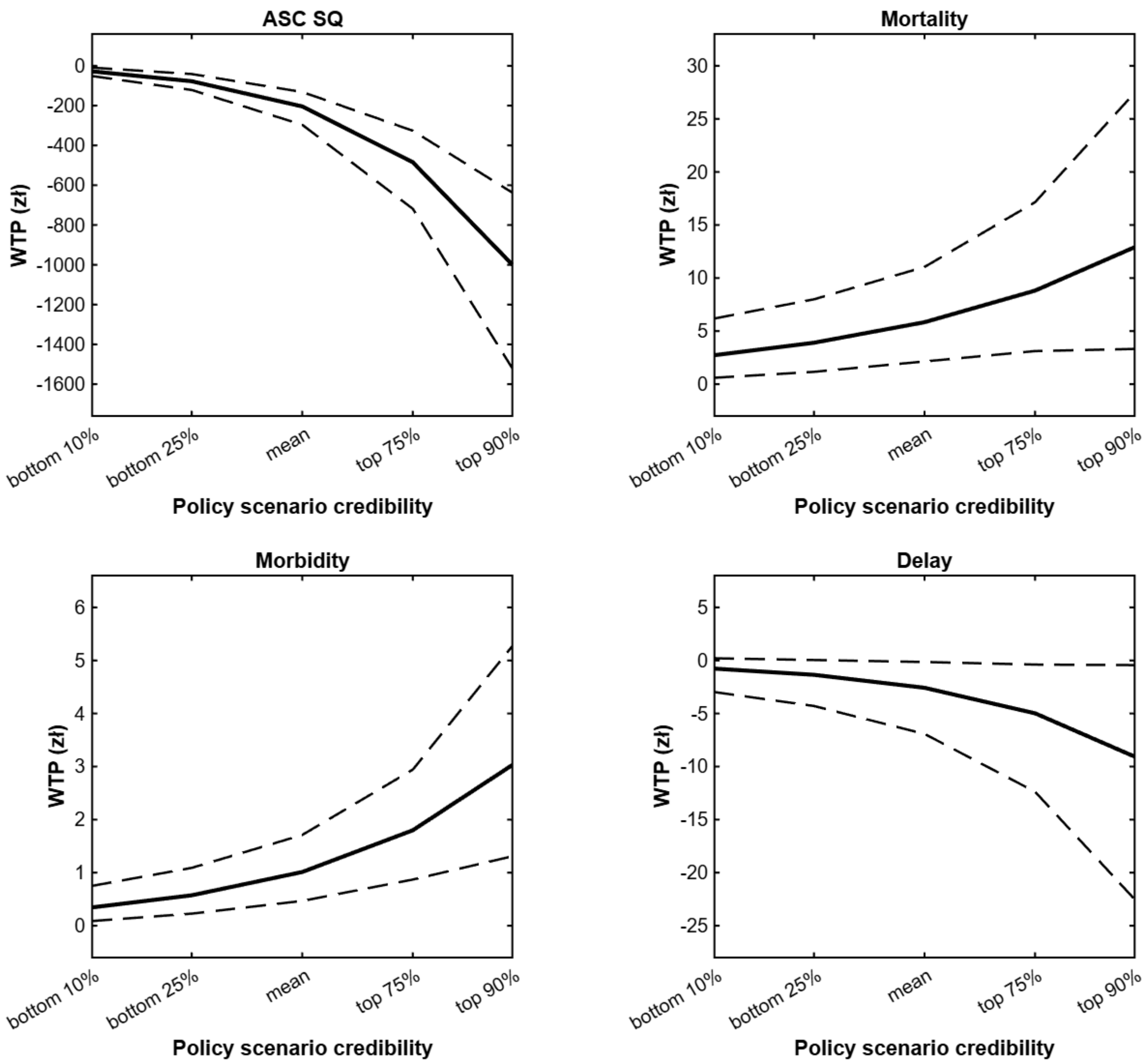
Attributes	Median WTP	Marginal effects on WTP			
		Policy scenario credibility	Information-source condition: Expert	Information-source condition: AI	Age
ASCSQ	-203.64***	-519.15***	-25.13	-84.53	42.18***
Mortality	5.40*	4.54	-1.31	-1.60	-0.13
Morbidity	0.95***	1.31*	-0.11	0.04	-0.01
Delay	-2.50	-4.14	1.07	0.71	0.45

*Note: ***, **, and * indicate 0.1%, 1%, and 5% significance levels, respectively; 1 Euro = 4.3 zł: nominal exchange rate (National Bank of Poland, 2026).*

As illustrated in Figure 2, WTP for specific attributes varies substantially depending on the credibility of the outcome projections described in the policy scenario. Specifically, higher credibility correlates with an elevated WTP for implementing additional municipal air pollution mitigation programmes relative to the SQ. Beyond general policy acceptance, higher policy scenario credibility is also linked to a higher marginal WTP for targeted health benefits related to reductions in COPD risks. Respondents who express higher trust in the provided projections of policy outcomes (1 SD above average) are willing to pay approximately 1.3 zł more to prevent one new COPD case annually per 100,000 urban residents. We do not observe such a relationship with mortality risk reductions. Additionally, older individuals exhibit significantly lower WTP for implementing additional municipal air pollution mitigation programmes relative to the SQ in their city of residence.

To better understand the HMXL model results regarding the influence of different information-source conditions, we conduct additional analysis (see Section 5.2). We verify respondents' ability to accurately recall the information source that policy projections were attributed to (or the fact that no source was disclosed) through the examination of manipulation check results. Additionally, we compare attitudes toward different information source types that can be used for health-related projections of air pollution mitigation policy outcomes; in particular, we compare respondents' trust toward projections generated by AI as well as scientists and medical experts. [4]This analysis is directly based on the examination of general credibility statements described in Section 3.4.

Figure 2. Marginal WTP of respondents with varying strengths of latent beliefs regarding scenario credibility. The solid black line denotes the estimated marginal WTP, while the dashed black lines represent the 95% confidence interval. Note: 1 Euro = 4.3 zł: nominal exchange rate (National Bank of Poland, 2026)



[4] For transparency, we additionally estimate a two-dimensional specification, which is included in the Appendix (Tables A3-A6). Despite its added flexibility, this alternative model achieves a lower log-likelihood and worse AIC and BIC values than the one-dimensional model, supporting our reliance on the pooled specification for the main analysis.

5.3. Recalled source and attitudes towards information sources

A chi-square test of independence reveals a highly significant relationship between the information-source condition assignment and the recalled source of information on policy projections (χ^2 = 172.10, df = 4, p < .001), indicating that the manipulation of information source is successfully noticed at an aggregate level (see Table 8).

Table 8. Information-source condition assignment vs recalled information source (N = 995)

	Recalled information source		
Information-source condition	Expert	AI	Other/do not know/undisclosed
Expert	53.1%	10.6%	36.2%
AI	30.8%	39.5%	29.7%
Control	41.3%	6.6%	52.2%

However, a closer inspection of the error patterns in responses to the manipulation check uncovers a heuristic of assuming expert involvement. In the AI condition, nearly one-third of the respondents (30.8%) erroneously attribute policy projections to respiratory disease experts. Additionally, in the control group, where no information about the source was provided, 41.3% of respondents spontaneously assume respiratory disease experts support the projections, whereas only 6.6% assume AI involvement. A chi-square goodness-of-fit test restricting the analysis to respiratory disease experts vs AI as the recalled information source confirms that respiratory disease experts were recalled significantly more frequently than AI as the information source (χ^2 = 92.2, df = 1, p < .001).

Table 9. Credibility of information on health-related projections of environmental policy outcomes from different information sources

Source	Mean
Scientists and medical experts	4.81
Scientists and medical experts supported by AI	4.34
AI (diagnostic algorithms and advanced chatbots)	3.69

Note: Responses were rated on a 7-point Likert scale ranging from 1 to 7, where higher values indicate higher perceived credibility.

In the post-choice survey part, all respondents were asked to rate the general credibility of projections on air quality health impacts in similar surveys if they were to be generated by three distinct information sources: domain experts (scientists and medical experts), AI, and domain experts supported by AI. The Friedman's test reveals a highly significant difference in credibility ratings (χ^2 = 607.62, df = 2, p < .001). To further examine these differences, we performed post-hoc pairwise comparisons using the Wilcoxon signed-rank test with a Bonferroni correction. The results indicate that credibility ratings for projections differ significantly across all three information sources (all p < .001), with projections from scientists and medical experts rated as the most credible, followed by those from the combined approach, and AI (see Table 9).

6. Discussion and conclusions

This study contributes to understanding the role of policy scenario acceptance in the valuation of environmental programmes. We examine how the credibility of projected policy outcomes affects preferences and WTP for COPD risk reductions from urban air pollution. We account for the potential roles of epistemic and implementation credibility, and we test whether the type of information source backing projections of policy outcomes significantly affects respondents' preferences. To address research objectives, we analyse results of a split-sample DCE conducted on respondents living in large Polish cities. Three main conclusions follow.

First, we demonstrate that policy scenario credibility can be an important determinant of valuation in complex environmental health settings. Respondents in our study who find the scenario more credible are less likely to opt for the SQ and less sensitive to programme cost, which translates into higher WTP for programme introduction and for reductions in annual COPD incidence risk. The magnitude of this effect has practical implications: because credibility acts mainly through the ASC_{SQ} and the cost parameter rather than through its influence on preferences regarding health attributes, a scenario-sceptical sample does not produce a differently shaped preference structure, but rather a uniformly deflated one. Aggregate benefit estimates from such a sample are therefore likely to largely understate public support for an intervention with little visible distortion in the relative valuation of its outcomes.

Second, epistemic and implementation credibility are conceptually separable but, in our empirical setting, the unidimensional credibility construct better explains respondents' preferences. Our results indicate that the average evaluations of epistemic credibility are significantly higher than evaluations of implementation credibility. This aligns with data suggesting that general trust in institutions of science and experts, who were the most frequently credited as an information source by respondents to our survey, is higher than trust in governing institutions (IPSOS, 2022; Wellcome, 2018). Yet both sets of indicator questions measuring epistemic and implementation credibility load on one construct in the estimated HMXL model, and this specification fits the data better. We treat this as evidence that in politicised, low-literacy domains such as environmental health policies, individuals are likely to form a holistic judgement of whether "the story adds up", rather than auditing its scientific and administrative premises separately. Existing research suggests that the authority of science has been decreasing in recent years, and that individuals' trust in scientific communication on some environmental issues depends on their political orientation and attitude toward governing authorities (Cologna et al., 2025; Hart et al., 2015; Schuster & Scheu, 2026). This has a design implication: elicitation instruments that ask respondents to hold one credibility dimension constant while evaluating the other, as our survey did, may not achieve the intended separation in some contexts. It also raises a substantive question for future work, namely, whether separability emerges in contexts where science is more contested than governance.

Third, we find that disclosing the source of policy outcome projections does not, by itself, move preferences or credibility. However, respondents hold clearly ordered attitudes toward different information sources operating in analogous settings, ranking human experts above experts using AI, and ranking both above standalone AI. This reflects a pattern similar to algorithm aversion demonstrated in health contexts (Reis et al., 2024). The contrast, where algorithm aversion appears to be uncovered by attitudinal questions but not in DCE settings, aligns well with the findings of Logg & Schlund (2024). They show that although individuals declare a preference for human-generated advice, they do not discount algorithmically generated advice in decision-making practice. However, we also find that a substantial share of respondents misattributed the source of policy outcome projections, assuming they originate from domain experts. This may be partially explained by information identifying researchers as survey authors, which we included in the survey preamble for transparency (Johnston et al., 2017). Assumptions of expert involvement might also attenuate the influence of the experimental condition in which AI was presented as the information source. We argue that practitioners considering AI-assisted scenario development in SP valuation should therefore be mindful that respondents may not process AI attribution unless it is made highly salient, and that general attitudes toward different information sources appear to favour involvement of human experts.

Some limitations warrant caution when interpreting our findings. Policy scenario credibility was measured after the choices, leaving open the possibility that respondent choices shaped subsequent credibility reports; source disclosure in experimental conditions was not highly salient; and the results pertain to one country and one environmental health outcome. We nonetheless demonstrate that policy scenario credibility is not a nuisance to be controlled but can be strongly related to the obtained welfare estimate, and that its two conceptually distinct dimensions may operate in an integrated manner. Finally, we show that explicit distrust in AI has a negligible impact on policy scenario acceptance and choices in this context. Even though respondents report relatively low trust in information generated by AI compared to domain experts, attributing environmental policy outcome projections to an AI model does not alter their preferences over health risk reductions.

Funding

The study has received funding from [anonymised for review]. The work of [anonymised for review] was supported by [anonymised for review]. All funding support is gratefully acknowledged.

References

Adamowicz, W., Swait, J., Boxall, P., Louviere, J., Williams, M., 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. Journal of Environmental Economics and Management, 32(1), 65–84. https://doi.org/10.1006/jeem.1996.0957

Adeloye, D., Song, P., Zhu, Y., Campbell, H., Sheikh, A., Rudan, I., 2022. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: A systematic review and modelling analysis. The Lancet Respiratory Medicine, 10(5), 447–458. https://doi.org/10.1016/S2213-2600(21)00511-7

Bartczak, A.M., Budziński, W., Jusypenko, B., Boros, P.W., 2024. The Impact of Health Status and Experienced Disutility on Air Quality Valuation. Ecological Economics, 217, 108071. https://doi.org/10.1016/j.ecolecon.2023.108071

Bateman, I., 2002. Economic valuation with stated preference techniques: A manual (Vol. 50). Edward Elgar Cheltenham. https://doi.org/10.4337/9781781009727

Baumann, E., Czerwinski, F., Rosset, M., Seelig, M., Suhr, R., 2020. How do people in Germany seek health information? Insights from the first wave of HINTS Germany. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz, 63(9), 1151–1160. https://doi.org/10.1007/s00103-020-03192-x

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D.S., 2002. Hybrid choice models: Progress and challenges. Marketing Letters, 13, 163–175. https://doi.org/10.1023/A:1020254301302

Birol, E., Das, S., 2012. Valuing the environment in developing countries: Modelling the impact of distrust in public authorities' ability to deliver public services on the citizens' willingness to pay for improved environmental quality. Urban Water Journal, 9(4), 249–258. https://doi.org/10.1080/1573062X.2012.660958

Bochenek, B., Jankowski, M., Wieczorek, J., Gruszczyńska, M., Sekuła, P., Jaczewski, A., Wyszogrodzki, A., Pinkas, J., Figurski, M., 2025. The impact of ambient air pollution and meteorological factors on emergency hospital admissions of COPD patients in Poland (2012–2019). Scientific Reports, 15(1), 21915. https://doi.org/10.1038/s41598-025-07684-1

Börger, T., Glenk, K., Meyerhoff, J., Rehdanz, K., 2025. Mitigating Cost-Vector Effects in Stated Choice Experiments Using Cheap Talk and Opt-Out Reminders. Journal of the Association of Environmental and Resource Economists, 12(3), 701–737. https://doi.org/10.1086/731886

Boxall, P.C., Adamowicz, W.L., Swait, J., Williams, M., Louviere, J., 1996. A comparison of stated preference methods for environmental valuation. Ecological Economics, 18(3), 243–253. https://doi.org/10.1016/0921-8009(96)00039-0

Boyle, K.J., Bergstrom, J.C., 1999. Doubt, Doubts, and Doubters: The genesis of a new research agenda?, in: Bateman, I., Willis, K. (Eds.), Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in The US, EU, and Developing Countries, 183–207. https://doi.org/10.1093/0199248915.003.0007

Bulte, E., Gerking, S., List, J.A., De Zeeuw, A., 2005. The effect of varying the causes of environmental problems on stated WTP values: Evidence from a field study. Journal of Environmental Economics and Management, 49(2), 330–342. https://doi.org/10.1016/j.jeem.2004.06.001

Burchard-Dziubińska, M., 2019. Air pollution and health in Poland: Anti-smog movement in the most polluted Polish cities. Economics and Environment, 69(2), 15. https://doi.org/https://ekonomiaisrodowisko.pl/journal/article/view/85

Cameron, T.A., DeShazo, J.R., Johnson, E.H., 2011. Scenario adjustment in stated preference research. Journal of Choice Modelling, 4(1), 9–43. https://doi.org/10.1016/S1755-5345(13)70017-4

Chaiken, S., 1980. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. Journal of Personality and Social Psychology, 39(5), 752. https://doi.org/10.1037/0022-3514.39.5.752

Clarke, C.E., Evensen, D.T., 2019. The politics of scientific consensus? Political divergence and partisanship in unconventional energy development in the United States. Energy Research & Social Science, 51, 156–167. https://doi.org/10.1016/j.erss.2019.01.005

Cologna, V., Mede, N.G., Berger, S., Besley, J., Brick, C., Joubert, M., Maibach, E.W., Mihelj, S., Oreskes, N., Schäfer, M.S., 2025. Trust in scientists and their role in society across 68 countries. Nature Human Behaviour, 9(4), 713–730. https://doi.org/10.1038/s41562-024-02090-5

Crawford, V.P., Sobel, J., 1982. Strategic information transmission. Econometrica: Journal of the Econometric Society, 1431–1451. https://doi.org/10.2307/1913390

Czajkowski, M., Budziński, W., 2019. Simulation error in maximum likelihood estimation of discrete choice models. Journal of Choice Modelling, 31, 73–85. https://doi.org/10.1016/j.jocm.2019.04.003

Czajkowski, M., Hanley, N., LaRiviere, J., 2016. Controlling for the Effects of Information in a Public Goods Discrete Choice Model. Environmental and Resource Economics, 63(3), 523–544. https://doi.org/10.1007/s10640-014-9847-z

Daly, A., Hess, S., Patruni, B., Potoglou, D., Rohr, C., 2012. Using ordered attitudinal indicators in a latent variable choice model: A study of the impact of security on rail travel behaviour. Transportation, 39(2), 267–297. https://doi.org/10.1007/s11116-011-9351-z

De-Magistris, T., Ben Youssef, K., Akaichi, F., 2016. Testing the effectiveness of the oath script in reducing the hypothetical bias in the Contingent Valuation Method. https://doi.org/10.17221/174/2015-AGRICECON

DeShazo, J.R., Cameron, T.A., 2005. The effect of health status on willingness to pay for morbidity and mortality risk reductions. https://escholarship.org/uc/item/9431841p

Dickie, M., Adamowicz, W., Gerking, S., Veronesi, M., 2022. Risk perception, learning, and willingness to pay to reduce heart disease risk. Journal of Benefit-Cost Analysis, 13(3), 363–382. https://doi.org/10.1017/bca.2022.14

Dietvorst, B.J., Simmons, J.P., Massey, C., 2015. Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114. https://doi.org/10.1037/xge0000033

EEA, (European Environment Agency), 2023. Europe’s air quality status 2023 [Briefing]. https://www.eea.europa.eu/publications/europes-air-quality-status-2023

Esmaeilzadeh, P., Maddah, M., Mirzaei, T., 2025. Using AI chatbots (eg, CHATGPT) in seeking health-related information online: The case of a common ailment. Computers in Human Behavior: Artificial Humans, 3, 100127. https://doi.org/10.1016/j.chbah.2025.100127

Finn, S., O’Fallon, L.R. (Eds), 2019. Environmental Health Literacy. Springer International Publishing. https://doi.org/10.1007/978-3-319-94108-0

Fischhoff, B., Welch, N., Frederick, S., 1999. Construal Processes in Preference Assessment. Journal of Risk and Uncertainty, 19(1–3), 139–164. https://doi.org/10.1023/A:1007823326511

Fiske, S.T., Cuddy, A.J., Glick, P., 2007. Universal dimensions of social cognition: Warmth and competence. Trends in Cognitive Sciences, 11(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Flores, N.E., Strong, A., 2007. Cost credibility and the stated preference analysis of public goods. Resource and Energy Economics, 29(3), 195–205. https://doi.org/10.1016/j.reseneeco.2006.09.001

Fu, L., Li, J., Chen, Y., 2023. An innovative decision making method for air quality monitoring based on big data-assisted artificial intelligence technique. Journal of Innovation & Knowledge, 8(2), 100294. https://doi.org/10.1016/j.jik.2022.100294

Fuller, R., Landrigan, P.J., Balakrishnan, K., Bathan, G., Bose-O'Reilly, S., Brauer, M., Caravanos, J., Chiles, T., Cohen, A., Corra, L., 2022. Pollution and health: A progress update. The Lancet Planetary Health. https://doi.org/10.1016/S2542-5196(22)00090-0

Glenk, K., Fischer, A., 2010. Insurance, prevention or just wait and see? Public preferences for water management strategies in the context of climate change. Ecological Economics, 69(11), 2279–2291. https://doi.org/10.1016/j.ecolecon.2010.06.022

Google, 2026, May 30. Quantifying Health Co-Benefits of Air Quality Policies: DCE Attribute Development for Polish Cities [Generative AI chat]. https://gemini.google.com/app/a91710d7b4af29d5?hl=pl

GUS, 2023. Local Data Bank (Bank Danych Lokalnych). https://bdl.stat.gov.pl/bdl/start

Hardwig, J., 1991. The role of trust in knowledge. The Journal of Philosophy, 88(12), 693–708. https://doi.org/10.2307/2027007

Hart, P.S., Nisbet, E.C., Myers, T.A., 2015. Public attention to science and political news and support for climate change mitigation. Nature Climate Change, 5(6), 541–545. https://doi.org/10.1038/nclimate2577

Hasselström, L., Håkansson, C., 2014. Detailed vs. fuzzy information in non-market valuation studies: The role of familiarity. Journal of Environmental Planning and Management, 57(1), 123–143. https://doi.org/10.1080/09640568.2012.736370

Hausman, J., 2012. Contingent valuation: From dubious to hopeless. Journal of Economic Perspectives, 26(4), 43–56. https://doi.org/10.1257/jep.26.4.43

Hendriks, F., Kienhues, D., Bromme, R., 2016. Trust in Science and the Science of Trust, in: Blöbaum, B. (Ed.), Trust and Communication in a Digitized World (pp. 143–159). Springer International Publishing. https://doi.org/10.1007/978-3-319-28059-2_8

Hoffmann, B.L., Scartascini, C., Cafferata, F.G., 2022. How can we improve air pollution? Try increasing trust first. Environment and Development Economics, 27(5), 393–413. https://doi.org/10.1017/S1355770X2100036X

Hou, W.-H., Huang, Y.-C., Lu, C.-Y., Chen, I.-C., Lee, P.-C., Lin, M.-Y., Wang, Y.-C., Sulistyorini, L., Li, C.-Y., 2021. A national survey of ambient air pollution health literacy among adult residents of Taiwan. BMC Public Health, 21(1), 1604. https://doi.org/10.1186/s12889-021-11658-z

Hovland, C.I., Weiss, W., 1951. The influence of source credibility on communication effectiveness. Public Opinion Quarterly, 15(4), 635–650. https://doi.org/10.1086/266350

Hundemer, S., Monroe, M.C., Kaplan, D., 2022. The water science communication problem: Water knowledge and the acceptance or rejection of water science. Journal of Hydrology, 604, 127230. https://doi.org/10.1016/j.jhydrol.2021.127230

Institute for Health Metrics and Evaluation, 2023. GBD Compare. Institute for Health Metrics and Evaluation. https://vizhub.healthdata.org/gbd-compare

IPSOS, 2022. Global Trustworthiness Monitor. https://www.ipsos.com/en/trust

Jacquemet, N., Joule, R.-V., Luchini, S., Shogren, J.F., 2013. Preference elicitation under oath. Journal of Environmental Economics and Management, 65(1), 110–132. https://doi.org/10.1016/j.jeem.2012.05.004

Jin, Y., Andersson, H., Zhang, S., 2020. Do preferences to reduce health risks related to air pollution depend on illness type? Evidence from a choice experiment in Beijing, China. Journal of Environmental Economics and Management, 103, 102355. https://doi.org/10.1016/j.jeem.2020.102355

Johnston, R.J., Boyle, K.J., 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. Journal of the Association of Environmental and Resource Economists, 4(2), 319–405. https://doi.org/10.1086/691697

Kataria, M., Bateman, I., Christensen, T., Dubgaard, A., Hasler, B., Hime, S., Ladenburg, J., Levin, G., Martinsen, L., Nissen, C., 2012. Scenario realism and welfare estimates in choice experiments–A non-market valuation study on the European water framework directive. Journal of Environmental Management, 94(1), 25–33. https://doi.org/10.1016/j.jenvman.2011.08.010

King, P., 2025. How confident are you in the ability of experts to provide reliable information? Evidence from a choice experiment on microplastics. Environment and Development Economics, 1–24. https://doi.org/10.1017/s1355770x25100077

Kling, C.L., Phaneuf, D.J., Zhao, J., 2012. From Exxon to BP: Has some number become better than no number? Journal of Economic Perspectives, 26(4), 3–26. https://doi.org/10.1257/jep.26.4.3

Krinsky, I., Robb, A.L., 1986. On approximating the statistical properties of elasticities. The Review of Economics and Statistics, 715–719. https://doi.org/10.2307/1924536

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R., Heintzelman, M., 2002. Age, health and the willingness to pay for mortality risk reductions: A contingent valuation survey of Ontario residents. Journal of Risk and Uncertainty, 24(2), 161–186. https://doi.org/10.1023/A:1014020027011

Kummervold, P.E., Wynn, R., 2012. Health Information Accessed on the Internet: The Development in 5 European Countries. International Journal of Telemedicine and Applications, 2012, 1–3. https://doi.org/10.1155/2012/297416

Lancaster, K.J., 1966. A new approach to consumer theory. Journal of Political Economy, 74(2), 132–157. https://doi.org/10.1086/259131

Lartey, D., Law, K.M., 2025. Artificial intelligence adoption in urban planning governance: A systematic review of advancements in decision-making, and policy making. Landscape and Urban Planning, 258, 105337. https://doi.org/10.1016/j.landurbplan.2025.105337

Li, J., Sun, S., Tang, R., Qiu, H., Huang, Q., Mason, T., Tian, L., 2016. Major air pollutants and risk of COPD exacerbations: A systematic review and meta-analysis. International Journal of Chronic Obstructive Pulmonary Disease, Volume 11, 3079–3091. https://doi.org/10.2147/COPD.S122282

Logg, J.M., Minson, J.A., Moore, D.A., 2019. Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Logg, J.M., Schlund, R., 2024. A Simple Explanation Reconciles "Algorithm Aversion" and "Algorithm Appreciation": Hypotheticals vs. Real Judgments (SSRN Scholarly Paper No. 4687557). Social Science Research Network. https://doi.org/10.2139/ssrn.4687557

Longoni, C., Bonezzi, A., Morewedge, C.K., 2019. Resistance to medical artificial intelligence. Journal of Consumer Research, 46(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Louviere, J.J., 2006. What you don't know might hurt you: Some unresolved issues in the design and analysis of discrete choice experiments. Environmental and Resource Economics, 34(1), 173–188. https://doi.org/10.1007/s10640-005-4817-0

Louviere, J.J., Hensher, D.A., Swait, J.D., 2000. Stated choice methods: Analysis and applications. Cambridge university press. https://doi.org/https://doi.org/10.1017/CBO9780511753831

Lundhede, T., Jacobsen, J.B., Hanley, N., Strange, N., Thorsen, B.J., 2015. Incorporating outcome uncertainty and prior outcome beliefs in stated preferences. Land Economics, 91(2), 296–316. https://doi.org/10.3368/le.91.2.296

Lunz Trujillo, K., Trujillo, M., 2026. What Drives a Contemporary Lack of Trust in Government and Science in the United States? American Politics Research, 54(4), 430–443. https://doi.org/10.1177/1532673X261443521

Manski, C.F., 1999. Analysis of Choice Expectations in Incomplete Scenarios. Journal of Risk and Uncertainty, 19(1–3), 49–66. https://doi.org/10.1023/A:1007806822876

McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior.

McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics, 15(5), 447–470. https://doi.org/10.1002/1099-1255(200009/10)15:5%3C447::AID-JAE570%3E3.0.CO;2-1

McKee, K.R., Bai, X., Fiske, S.T., 2023. Humans perceive warmth and competence in artificial intelligence. Iscience, 26(8). https://doi.org/10.1016/j.isci.2023.107256

Ministerstwo Zdrowia, 2026. Przewlekła obturacyjna choroba płuc – Mapy potrzeb zdrowotnych. https://basiw.mz.gov.pl/analizy/problemy-zdrowotne/przewlekla-obturacyjna-choroba-pluc/

Murphy, J.J., Allen, P.G., Stevens, T.H., Weatherhead, D., 2005. A meta-analysis of hypothetical bias in stated preference valuation. Environmental and Resource Economics, 30, 313–325. https://doi.org/10.1007/s10640-004-3332-z

National Bank of Poland, 2026. Exchange rates – table A. Narodowy Bank Polski - Internetowy Serwis Informacyjny. https://nbp.pl/en/statistic-and-financial-reporting/rates/table-a/

NIK, N.I.K., 2025. Kopciuchy wciąż zatruwają powietrze—Najwyższa Izba Kontroli. https://www.nik.gov.pl/najnowsze-informacje-o-wynikach-kontroli/realizacja-uchwal-antysmogowych.html

Nowak, A.K., Vallacher, R.R., Rychwalska, A., Roszczyńska-Kurasińska, M., Ziembowicz, K., Biesaga, M., Kacprzyk-Murawska, M., 2019. Target in Control: Social Influence as Distributed Information Processing. Springer International Publishing. https://doi.org/10.1007/978-3-030-30622-9

Oh, H., Hong, J.H., 2012. Citizens' trust in government and their willingness-to-pay. Economics Letters, 115(3), 345–347. https://doi.org/10.1016/j.econlet.2011.12.010

Oleksy, T., Wnuk, A., Domaradzka, A., Maison, D., 2023. What shapes our attitudes towards algorithms in urban governance? The role of perceived friendliness and controllability of the city, and human-algorithm cooperation. Computers in Human Behavior, 142, 107653. https://doi.org/10.1016/j.chb.2023.107653

Pfleger, E., Drexler, H., Lutz, R., 2024. Health literacy and environmental risks focusing air pollution: Results from a cross-sectional study in Germany. International Journal of Environmental Research and Public Health, 21(3), 366. https://doi.org/10.3390/ijerph21030366

Qin, X., Zhou, X., Chen, C., Wu, D., Zhou, H., Dong, X., Cao, L., Lu, J.G., 2025. AI aversion or appreciation? A capability–personalization framework and a meta-analytic review. Psychological Bulletin, 151(5), 580. https://doi.org/10.1037/bul0000477

Reis, M., Reis, F., Kunde, W., 2024. Influence of believed AI involvement on the perception of digital medical advice. Nature Medicine, 30(11), 3098–3100. https://doi.org/10.1038/s41591-024-03180-7

Schuster, C., Scheu, A.M., 2026. How Communication of Scientific Uncertainty Affects Trust in Science—A Systematic Review. Risk Analysis, 46(5), e70233. https://doi.org/10.1111/risa.70233

Sharma, S., Chandra, M., Kota, S.H., 2020. Health effects associated with PM 2.5: A systematic review. Current Pollution Reports, 6, 345–367. https://doi.org/10.1007/s40726-020-00155-3

Sin, D.D., Doiron, D., Agusti, A., Anzueto, A., Barnes, P.J., Celli, B.R., Criner, G.J., Halpin, D., Han, M.K., Martinez, F.J., 2023. Air pollution and COPD: GOLD 2023 committee report. European Respiratory Journal, 61(5), 2202469. https://doi.org/10.1183/13993003.02469-2022

Singh, G., 2024. Harnessing AI in Decision Support Systems for Comprehensive Pollution Management. In Spatial Intelligence for a Greener Planet (pp. 155–169). CRC Press. https://doi.org/10.1201/9781032718323-12

Śliwiński, P., Puchalski, K., 2015. Chronic obstructive pulmonary disease in the awareness of Polish society. Report from the public opinion survey by the Polish Respiratory Society and TNS Polska. Advances in Respiratory Medicine, 83(1), 1–13. https://doi.org/10.5603/PiAP.2015.0001

Stevens, T.H., Tabatabaei, M., Lass, D., 2013. Oaths and hypothetical bias. Journal of Environmental Management, 127, 135–141. https://doi.org/10.1016/j.jenvman.2013.04.038

Swoboda, C.M., Van Hulle, J.M., McAlearney, A.S., Huerta, T.R., 2018. Odds of talking to healthcare providers as the initial source of healthcare information: Updated cross-sectional results from the Health Information National Trends Survey (HINTS). BMC Family Practice, 19(1), 146. https://doi.org/10.1186/s12875-018-0805-7

Thaler, R., 1980. Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization, 1(1), 39–60. https://doi.org/10.1016/0167-2681(80)90051-7

Train, K.E., 2009. Discrete choice methods with simulation. Cambridge University Press. https://doi.org/10.1017/cbo9780511805271

Vaishya, R., Misra, A., Vaish, A., 2023. ChatGPT: Is this version good for healthcare and research? Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 17(4), 102744. https://doi.org/10.1016/j.dsx.2023.102744

Vásquez, W.F., De Rezende, C.E., 2019. Willingness to pay for the restoration of the Paraíba do Sul River: A contingent valuation study from Brazil. Ecohydrology & Hydrobiology, 19(4), 610–619. https://doi.org/10.1016/j.ecohyd.2018.01.001

Vij, A., Walker, J.L., 2016. How, when and why integrated choice and latent variable models are latently useful. Transportation Research Part B: Methodological, 90, 192–217. https://doi.org/10.1016/j.trb.2016.04.021

Wang, B., Chen, X., Yao, S., 2026. Birds of a feather? How perceived similarity shapes responses to AI versus human health messages. Communication Research Reports, 1–13. https://doi.org/10.1080/08824096.2026.2695125

Wellcome, 2018. Wellcome Global Monitor: How Does the World Feel about Science and Health? https://doi.org/https://cms.wellcome.org/sites/default/files/wellcome-global-monitor-2018.pdf

Welling, M., Zawojska, E., Sagebiel, J., 2022. Information, consequentiality and credibility in stated preference surveys: A choice experiment on climate adaptation. Environmental and Resource Economics, 82(1), 257–283. https://doi.org/10.1007/s10640-022-00675-0

Yang, Y., Hobbs, J.E., 2020. The Power of Stories: Narratives and Information Framing Effects in Science Communication. American Journal of Agricultural Economics, 102(4), 1271–1296. https://doi.org/10.1002/ajae.12078

Yokoi, R., Eguchi, Y., Fujita, T., Nakayachi, K., 2021. Artificial Intelligence Is Trusted Less than a Doctor in Medical Treatment Decisions: Influence of Perceived Care and Value Similarity. International Journal of Human–Computer Interaction, 37(10), 981–990. https://doi.org/10.1080/10447318.2020.1861763

Zambrano-Monserrate, M.A., 2024. Trust, corruption, and willingness to pay for improved environmental goods: An experimental analysis from a developing country. Social Science Quarterly, 105(3), 709–725. https://doi.org/10.1111/ssqu.13381

Appendixes

Table A1. Statements used to capture protest responses and valid economic motivations (true zeros)

No.	Statement
	True zero response
1	Reducing the health risks associated with COPD is not important to me.
2	My household cannot afford an additional annual fee.
	Protest response
1	I did not have sufficient information to make a decision.
2	Information regarding the health impacts of air pollution was not credible to me.
3	Estimates regarding the number of COPD cases and deaths that the programme could prevent were not credible to me.
4	My city's authorities would not be able to implement the programmes, even if it is technically possible.
5	The collected funds would, in reality, be used for a different purpose.
6	These activities should be funded from other sources instead of increasing local fees in my city.
7	I am opposed to increasing local fees regardless of the purpose.
8	People who have a higher risk of developing COPD (e.g., smokers) should fund these activities themselves.
9	People who contribute most to air pollution (e.g., use incorrect fuel in heating stoves) should pay for this programme themselves.
10	I have a right to clean air, and reducing the risk of developing COPD is not something I should pay extra for.
11	I refuse to value air quality in monetary terms.

Table A2. Descriptive statistics and distribution of responses for protest statements (N = 105)

No.	Protest motivation	Mean	Strongly disagree	Disagree	Rather disagree	Neither agree nor disagree	Rather agree	Agree	Strongly agree
1	I did not have sufficient information to make a decision.	4.09	7.6%	6.7%	11.4%	39.0%	21.0%	7.6%	6.7%
2	Information regarding the health impacts of air pollution was not credible to me.	3.98	6.7%	5.7%	12.4%	47.6%	17.1%	6.7%	3.8%
3	Estimates regarding the number of COPD cases and deaths that the programme could prevent were not credible to me.	4.31	3.8%	3.8%	8.6%	43.8%	25.7%	9.5%	4.8%
4	My city's authorities would not be able to implement the programmes, even if it is technically possible.	4.22	4.8%	3.8%	9.5%	46.7%	21.9%	8.6%	4.8%
5	The collected funds would, in reality, be used for a different purpose.	4.51	1.9%	2.9%	9.5%	42.9%	23.8%	8.6%	10.5%
6	These activities should be funded from other sources instead of increasing local fees in my city.	5.30	1.0%	1.9%	1.9%	31.4%	16.2%	20.0%	27.6%
7	I am opposed to increasing local fees regardless of the purpose.	5.10	1.0%	0.0%	10.5%	29.5%	21.0%	12.4%	25.7%

No.	Protest motivation	Mean	Strongly disagree	Disagree	Rather disagree	Neither agree nor disagree	Rather agree	Agree	Strongly agree
8	People who have a higher risk of developing COPD (e.g., smokers) should fund these activities themselves.	4.07	8.6%	6.7%	8.6%	43.8%	16.2%	10.5%	5.7%
9	People who contribute most to air pollution (e.g., use incorrect fuel in heating stoves) should pay for this programme themselves.	4.85	2.9%	4.8%	3.8%	38.1%	15.2%	14.3%	21.0%
10	I have a right to clean air, and reducing the risk of developing COPD is not something I should pay extra for.	5.32	1.0%	1.0%	2.9%	31.4%	17.1%	17.1%	29.5%
11	I refuse to value air quality in monetary terms.	4.43	5.7%	4.8%	7.6%	41.0%	18.1%	9.5%	13.3%

Note: Responses were rated on a 7-point Likert scale ranging from 1 to 7, where higher values indicate higher perceived credibility.

Table A3. The choice component and interactions of means for the two-dimensional HMXL model incorporating epistemic and implementation credibility of the policy scenario as separate dimensions

	Dist.	Mean effect	Standard deviation effect
		Coef. (st.err)	Coef. (st.err)
Baseline model			
ASC_{SQ}	n	-5.94*** (0.66)	4.83*** (0.53)
Mortality	n	0.24*** (0.05)	0.36*** (0.05)
Morbidity/10	n	0.41*** (0.09)	0.73*** (0.07)
Delay	n	-0.08* (0.04)	0.15*** (0.04)
-Cost/100	ln	0.61*** (0.16)	2.47*** (0.11)
Effect of LV: Epistemic credibility			
ASC_{SQ}		-1.39** (0.48)	
Mortality		0.04 (0.04)	
Morbidity/10		0.05 (0.07)	
Delay		0.01 (0.03)	
-Cost/100		-0.56*** (0.13)	
Effect of LV: Implementation credibility			
ASC_{SQ}		-0.59 (0.43)	
Mortality		-0.05 (0.04)	
Morbidity/10		-0.02 (0.07)	
Delay		-0.01 (0.03)	
-Cost/100		-0.66*** (0.11)	
Effect of information-source condition: AI			
ASC_{SQ}		-0.61 (0.83)	
Mortality		-0.06 (0.07)	
Morbidity/10		-0.02 (0.11)	

	Dist.	Mean effect	Standard deviation effect
		Coef. (st.err)	Coef. (st.err)
Delay		0.02 (0.05)	
-Cost/100		-0.16 (0.23)	
Effect of information-source condition: Expert			
ASC_{SQ}		-0.16 (0.81)	
Mortality		-0.04 (0.07)	
Morbidity/10		-0.03 (0.12)	
Delay		0.03 (0.05)	
-Cost/100		-0.02 (0.24)	
Effect of socio-demographic variable: Age			
ASC_{SQ}		-0.22 (0.22)	
Mortality		0.03 (0.02)	
Morbidity/10		0.06* (0.03)	
Delay		0.00 (0.01)	
-Cost/100		0.30*** (0.07)	

*Note: ***, **, and * indicate 0.1%, 1%, and 5% significance levels, respectively; n stands for normal distribution, and ln stands for log-normal distribution. Standard errors are reported in parentheses. All model components are estimated jointly using the Maximum Simulated Likelihood approach with 10,000 scrambled Sobol draws. Annual cost per household is denoted in 100 zł.*

Table A4. Comparison of model statistics for both one- and two-dimensional HMXL specifications

	Unidimensional model	Two-dimensional model
LL at convergence	-9701.21	-9765.33
LL at constant(s) only	-13324.76	-13324.76
McFadden's pseudo-R^2	0.27	0.27
Ben-Akiva-Lerman's pseudo-R^2	0.27	0.26
AIC/n	3.27	3.30
BIC/n	3.35	3.38
No. of respondents (n)	995	995
No. of parameters	71	77

Table A5. The measurement component of the two-dimensional HMXL model

Measurement equations – ordered probits		
Effect of LV: Epistemic credibility		
	Coef.	St. error
Plausibility of health effects	2.37 ***	0.66
Plausibility of time horizon	1.17 ***	0.12
Effect of LV: Implementation credibility		
Capability to deliver results	2.15 ***	0.31
Capability to deliver on time	4.36 *	2.64

*Note: ***, **, and * indicate 0.1%, 1%, and 5% significance levels, respectively.*

Table A6. The structural component of the two-dimensional HMXL model

Structural equations		
Dependent variable – LV: Epistemic credibility		
	Coef.	St. error
Information-source condition: AI	0.04	0.04
Information-source condition: Expert	0.03	0.04
Age	-0.02	0.04
Dependent variable – LV: Implementation credibility		
Information-source condition: AI	-	-
Information-source condition: Expert	-	-
Age	0.02	0.04

Table A7. The choice component and interactions of means for the one-dimensional HMXL model estimated on a full sample (including protest responses, N = 1,100)

	Dist.	Mean effect	Standard deviation effect
		Coef. (st.err)	Coef. (st.err)
Baseline model			
ASC_{SQ}	n	-5.47*** (0.79)	7.49*** (0.73)
Mortality	n	0.23*** (0.06)	0.38*** (0.05)
Morbidity/10	n	0.44*** (0.09)	0.74*** (0.08)
Delay	n	-0.08* (0.04)	0.15*** (0.04)
-Cost/100	ln	0.66*** (0.20)	2.24*** (0.13)
Effect of LV: Scenario credibility			
ASC_{SQ}		-2.49 *** (0.49)	
Mortality		-0.02 (0.03)	
Morbidity/10		0.03 (0.06)	
Delay		-0.01 (0.02)	
-Cost/100		-0.81*** (0.13)	
Effect of information-source condition: AI			
ASC_{SQ}		-0.29 (1.00)	
Mortality		-0.07 (0.07)	
Morbidity/10		-0.04 (0.11)	
Delay		0.03 (0.05)	
-Cost/100		-0.18 (0.24)	
Effect of information-source condition: Expert			
ASC_{SQ}		0.86 (0.97)	
Mortality		-0.07 (0.07)	
Morbidity/10		-0.11 (0.12)	

	Dist.	Mean effect	Standard deviation effect
		Coef. (st.err)	Coef. (st.err)
Delay		0.04 (0.05)	
-Cost/100		-0.34 (0.27)	
Effect of socio-demographic variable: Age			
ASC_{SQ}		-0.17 (0.27)	
Mortality		0.04 (0.02)	
Morbidity/10		0.08** (0.03)	
Delay		0.00 (0.01)	
-Cost/100		0.33*** (0.07)	
Model diagnostics			
LL at convergence	-10429.70		
LL at constant(s) only	-14585.79		
McFadden's pseudo-R^2:	0.28		
Ben-Akiva-Lerman's pseudo-R^2:	0.28		
AIC/n	3.18		
BIC/n	3.26		
No. of respondents (n)	1100		
No. of parameters	71		

*Note: ***, **, and * indicate 0.1%, 1%, and 5% significance levels, respectively; n stands for normal distribution, and ln stands for log-normal distribution. Standard errors are reported in parentheses. All model components are estimated jointly using the Maximum Simulated Likelihood approach with 10,000 scrambled Sobol draws. Annual cost per household is denoted in 100 zł.*

Table A8. The measurement component of the one-dimensional HMXL model estimated on a full sample (including protest responses, N = 1,100)

Indicator question	Coef.	St. error
Plausibility of health effects	0.84***	0.04
Plausibility of time horizon	0.74***	0.04
Capability to deliver results	2.27***	0.11
Capability to deliver on time	3.16***	0.30

*Note: ***, **, and * indicate 0.1%, 1%, and 5% significance levels, respectively.*

Table A9. The structural component of the one-dimensional HMXL model estimated on a full sample (including protest responses, N = 1,100)

Variable	Coef.	St. error
Information-source condition: AI	0.06	0.04
Information-source condition: Expert	0.04	0.04
Age	0.01	0.04

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl
ISSN 2957-0506