Valuing a reduction in the risk of very low birth weight

A large scale multi-country stated preference approach

OECD ENVIRONMENT WORKING PAPERS NO. 217

Milan Ščasný
Iva Zvěřinová
Damien Dussaux

Organisation for Economic Co-operation and Development

ENV/WKP(2023)9

Unclassified **English - Or. English**

1 June 2023

ENVIRONMENT DIRECTORATE

Valuing a reduction in the risk of very low birth weight

A large scale multi-country stated preference approach

Environment Working Paper No. 217

By Milan Ščasný (1), Iva Zvěřinová (1), Damien Dussaux (2)

1) Charles University
2) OECD Environment Directorate

Authorised for publication by Jo Tyndall, Director, Environment Directorate.

Keywords: low birth weight, health risk, economic valuation, health valuation, morbidity valuation, monetised benefits, chemicals regulation, non-market valuation, stated preferences, surveys, willingness-to-pay, value of a statistical case.

JEL codes: D61, I18, J17, K32, Q51, Q53, Q58

OECD Environment Working Papers are available at www.oecd.org/environment/workingpapers.htm

Damien Dussaux (Damien.Dussaux@oecd.org)

JT03520399

OECD ENVIRONMENT WORKING PAPERS

Comments on Working Papers are welcomed, and may be sent to:

OECD Environment Directorate

2 rue André-Pascal, 75775 Paris Cedex 16, France

or by email: env.contact@oecd.org

OECD Environment Working Papers are published on
www.oecd.org/environment/workingpapers.htm as well as
on the OECD iLibrary (www.oecdilibrary.org)

Abstract

There is ample evidence that exposure to various chemicals can increase the probability of children to be born with low or very low birth weight. Infants born with very low birth weight have a higher risk of suffering from neurosensory problems, issues related to behavioural and social competencies, and learning disabilities than infants born with normal birth weight. Authorities face challenges in regulating chemical substances through actions such as bans and prohibitions, because of the difficulty in explicitly considering the economic benefits and costs of such regulations. Moreover, existing Values of a Statistical Case (VSC) of very low birth weight are rare and cannot be directly applied to the cost benefit analysis of chemical management options for a wide range of countries.

This paper is part of the series of large scale WTP studies resulting from the Surveys to elicit Willingness to pay to Avoid Chemicals related negative Health Effects (SWACHE) project that intends to improve the basis for doing cost benefit analyses of chemicals management options and environmental policies in general. The present paper details a stated preference survey estimating WTP to reduce the risk of very low birth weight, filling an important gap in the valuation literature and addressing a need for applied benefits analysis for chemicals regulation. The SWACHE very low birth weight survey was fielded in 9 countries: Canada, the Czech Republic, Italy, Mexico, the Netherlands, Switzerland, Türkiye, the United Kingdom, and the United States. In each country, a sample of 1 200 adults, representative of the general population and of childbearing age who are in a relationship and plan for a(nother) child within the next five years, was collected and empirically analysed.

The mean Value of a Statistical Case (VSC) of very low birth weight estimated in this study equals USD_{2022} Purchasing Power Parity (PPP) 1 194 000 and the median VSC equals USD_{2022} PPP 614 000. Country-specific mean VSC of very low birth weight vary between USD_{2022} PPP 805 000 for the United Kingdom and USD_{2022} PPP 1 744 000 for Italy.

Keywords: low birth weight, health risk, economic valuation, health valuation, morbidity valuation, monetised benefits, chemicals regulation, non-market valuation, stated preferences, surveys, willingness-to-pay, value of a statistical case.

JEL Codes: D61, I18, J17, K32, Q51, Q53, Q58

Résumé

Il existe de nombreuses preuves scientifiques que l'exposition à diverses substances chimiques peut augmenter la probabilité que les enfants aient un faible ou très faible poids à la naissance. Les enfants nés avec un très faible poids ont un risque plus élevé de souffrir de problèmes neurosensoriels, de problèmes liés aux compétences comportementales et sociales et de troubles de l'apprentissage que les enfants nés avec un poids normal. Les autorités sont confrontées à des défis lorsqu'elles souhaitent réglementer des substances chimiques par des mesures telles que des interdictions, car il est difficile de prendre en compte de manière explicite les bénéfices et les coûts économiques de telles mesures réglementaires. En outre, les valeurs existantes d'un cas statistique (VCS) de très faible poids à la naissance sont rares et ne peuvent pas être directement appliquées à l'analyse coûts-bénéfices des options de gestion des produits chimiques pour un large éventail de pays.

Ce document fait partie d'une série d'études portant sur le consentement à payer et réalisées à grande échelle dans le cadre du projet SWACHE (Surveys to elicit Willingness to pay to Avoid Chemicals related negative Health Effects). Ce projet vise à améliorer la réalisation des analyses coûts-bénéfices des options de gestion des produits et composés chimiques et des politiques environnementales en général. Le présent document détaille une enquête sur les préférences déclarées estimant le consentement à payer pour réduire le risque de très faible poids à la naissance, comblant ainsi une lacune importante dans la littérature portant sur la valorisation et répondant à un besoin dans la quantification des bénéfices lors de l'évaluation des options de gestions des produits et composés chimiques. L'enquête SWACHE sur le très faible poids à la naissance a été menée dans 9 pays : Canada, République tchèque, Italie, Mexique, Pays-Bas, Suisse, Türkiye, Royaume-Uni et États-Unis. Dans chaque pays, un échantillon de 1 200 adultes, représentatifs de la population générale et en âge de procréer, qui vivent en couple et envisagent d'avoir un (autre) enfant au cours des cinq prochaines années, a été recueilli et analysé empiriquement.

La VCS de très faible poids de naissance moyenne estimée dans cette étude est de USD2022 1 194 000 USD en parité de pouvoir d'achat (PPA) et la VCS médiane est de USD2022 PPA 614 000. Les VCS de très faible poids de naissance moyennes spécifiques à chaque pays varient entre USD2022 PPA 805 000 pour le Royaume-Uni et USD2022 PPA 1 744 000 pour l'Italie.

Mots-clés : faible poids à la naissance, risque pour la santé humaine, valorisation économique, valorisation de la santé, valorisation de la morbidité, bénéfices monétisés, réglementation des composés chimiques, valorisation non marchande, préférences déclarées, enquêtes, consentement à payer, valeur d'un cas statistique.

Classification JEL : D61, I18, J17, K32, Q51, Q53, Q58

.

Acknowledgements

The paper was prepared in the context of the SWACHE project by Milan Ščasný, Iva Zvěřinová and Damien Dussaux. This work was undertaken under the guidance of Olof Bystrom, Team Lead at the Environment and Economy Integration Division and Eeva Leinala, Principal Administrator at the Environment, Health and Safety Division at the OECD Environment Directorate. The work benefited from the overall supervision of Shardul Agrawala, Head of the Environment and Economy Integration Division at the OECD Environment Directorate.

The SWACHE project is organised in co-operation between the OECD Working Party on Integrating Environmental and Economic Policies (WPIEEP) and the Working Party on Risk Management (WPRM) and with the support or the SWACHE advisory group including experts in chemicals regulation, toxicology and economic valuation that was set up for this project.

The authors are grateful to delegates of the WPIEEP, delegates of the WPRM and members of the expert advisory group for their helpful comments and valuable feedback. The authors are also thankful to Nicolina Lamhauge of the OECD Secretariat for review of the outputs and communication support. The authors would also like to thank Nils Axel Braathen for the initial coordination of the SWACHE project. Illias Mousse Iye and Ivan Babiy provided editorial assistance and Elizabeth Del Bourgo, Lea Stapper and Clara GaÏda provided communications support.

This paper was co-funded by the European Union. The views expressed herein do not necessarily reflect the official views of the OECD or of the governments of its member countries and can in no way be taken to reflect the official opinion of the European Union.

This paper has also been produced with the financial assistance of Canada, France, the Netherlands, Norway, Poland, Sweden, Switzerland, the United Kingdom and the United States. The authors wish to thank the European Chemicals Agency for its contributions throughout the SWACHE project. The views expressed herein do not necessarily reflect the official views of the OECD or of the governments of its member countries and can in no way be taken to reflect the official opinion of OECD member countries.

Table of contents

Tables

Figures

Boxes

Executive summary

Infants born with very low birth weight have a higher risk of suffering from neurosensory problems, issues with behavioural and social competencies and learning disabilities than infants born with normal birth weight. The prevalence of very low birth weight is high in many countries (6 to 13 in 1 000 new-borns depending on the country). Moreover, there is ample evidence that exposure to various chemicals can increase the probability of children to be born with low or very low birth weight.

One key challenge for chemicals risk management relates to the monetisation of health benefits expected from actions to curb emissions of and exposure to such substances. Balancing the expected benefits against the costs of regulation is typically done using willingness-to-pay (WTP) values as inputs to cost-benefit analysis. Several studies have used stated preference methods to evaluate the benefits of reducing the risk of low birth weight and very low birth weight, but these studies date back to 2014 and 2016, focus on a limited number of countries and employ an elicitation scenario based on the use of a novel combination of vitamins that is not directly linked to chemical exposure and these studies therefore capture values for other health co-benefits.

This paper reports the results of a new stated preference study which estimates a policy relevant value per statistical case (VSC) of very low birth weight that is part of the series of large scale WTP studies resulting from the Surveys to elicit Willingness to pay to Avoid Chemicals related Health Effects (SWACHE) project. An online valuation survey was administered to 12 000 respondents from nine OECD countries, asking respondents whether they would be willing to pay a monthly fee over a period of 8 months to ensure a reduction in very low birth weight risk from substances in every-day products. Out of the 12 000 respondents, 6 940 passed a later quality screening based on time needed to complete the survey and whether respondents adequately understood the concept of risk. Respondents are representative of the respective general populations of each country with the additional constraints that they and their spouses are of childbearing age, in a stable relationship and wished for a(nother) child within the next five years. In other words, they would be direct beneficiaries of the risk reductions offered in the survey.

The WTP values provided in this study are uniquely valuable for socio-economic analysis practitioners and policy makers since they are derived for different countries using the same methodology and are therefore internationally comparable. Furthermore, because the present study is part of the SWACHE project that provides an economic valuation of 10 health effects using the same general approach, the values provided by the present report are also comparable across health effects. This large scale and comprehensive valuation effort, that to our knowledge has not been attempted previously, will facilitate quantitative analyses of chemicals management options and be helpful in formulating national and regional policy affecting health outcomes.

Across the countries surveyed, the resulting mean WTP is USD 597 per month for an average reduction of 4 in 1 000 in the risk of very low birth weight for a total period of 8 months, corresponding to a mean VSC of USD_{2022} PPP 1 194 000 (all bids converted to USD and adjusted for purchasing power). The study also derives country-specific VSC of very low birth weight and mean values vary between USD_{2022} PPP 805 000 for the United Kingdom and USD_{2022} PPP 1 744 000 for Italy.

Various checks indicate that both the mean and the country-specific VSC estimates are fairly robust towards different modelling, data cleaning and screening choices. A comparison to previous studies that have estimated VSC of very low birth weight using a new combination of vitamins as payment vehicles show that the estimates of the present study are 1.2 to 2.8 times higher, depending on the country, using a similar estimation strategy. This divergence might be due to an acceptability for safer products used in the present survey that is higher than the ingestion of novel combination of vitamins and minerals used in previous studies.

The baseline risk level does not affect the WTP for a reduction in the risk of very low birth weight. Consequently, no further adjustment with respect to the differences in the baseline incidence or prevalence of very low birth weight is needed in benefit transfer or cost benefit analyses.

While respondents were presented with the health benefits of using safer products only in terms of reduced risk of very low birth weight, some of them were also thinking about other effects, potentially increasing their stated WTP. Consequently, the baseline estimates of WTP and VSC of very low birth weight may implicitly include values for these other effects, notably a reduced probability of miscarriage and infant mortality as well as a reduction in other health impacts. Therefore, the present paper also provides estimate of VSC that exclude the values for these other effects and can be used in specific cost-benefit analysis applications that consider miscarriage and infant death to avoid double counting.

1 The valuation of very low birth weight

1.1. Motivation

There is substantial evidence on the adverse effects of chemicals on birth weight (Bell et al., 2010[1]; Miao et al., 2011[2]; Birks and Casas, 2016[3]; de Cock et al., 2016[4]; Govarts et al., 2016[5]; Lenters et al., 2016[6]; Sun et al., 2016[7]; Woods et al., 2017[8]; Wikström et al., 2020[9]; Hu et al., 2021[10]). These studies analyse the effects of various chemicals, including heavy metals, particulate matter (PM) fractions and co-exposure of chemicals on various birth outcomes, including birth weight, size at gestational age, length of gestation, preterm birth (studies addressing the effects of chemicals on low birth weight or at least on birth weight were primarily searched).[1] "Low birth weight can be a consequence of pre-term birth (i.e. before 37 completed weeks of gestation), or due to small size for gestational age, defined as weight for gestation < 10th percentile, or both" (Edmond and Bahl, 2007[11]). A causal link between outdoor air pollution and lowered birth weight is also supported by a review of the epidemiological and toxicological literature conducted by WHO (2004[12]), see (Hunt, 2011[13]).

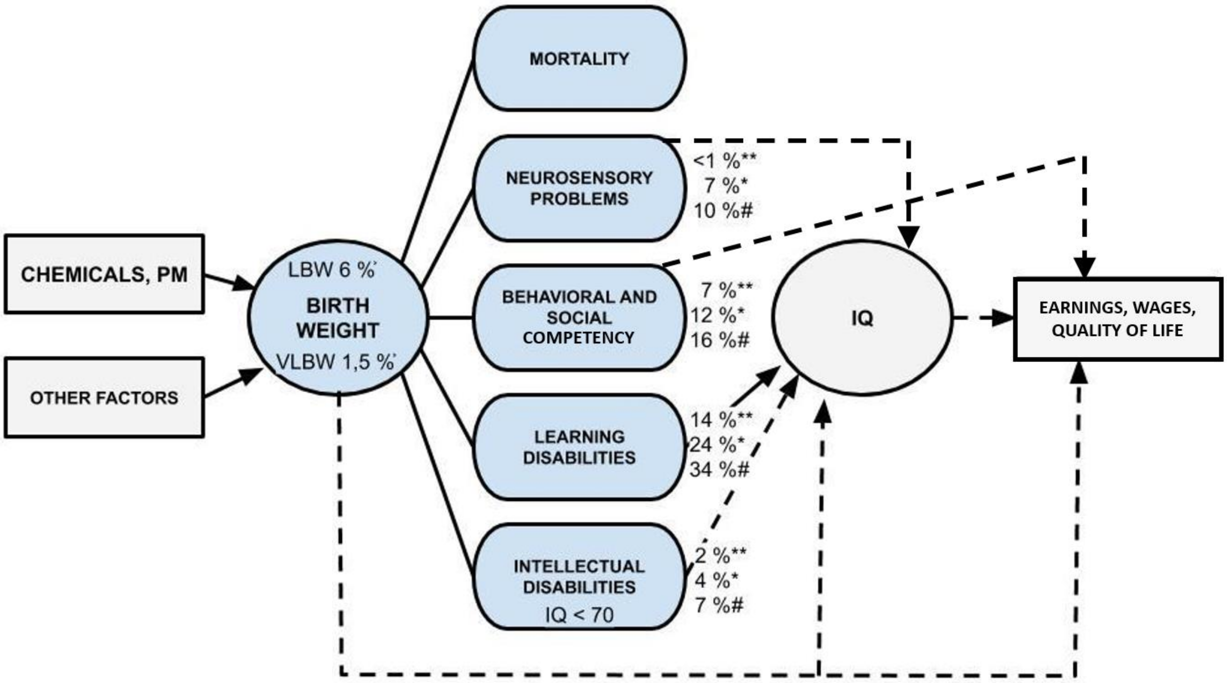
There is also a considerable body of literature on the effects of adverse birth outcomes on health condition of infants, young children, and continuing to young adults (Hack, Klein and Taylor, 1995[14]). These studies cover the effects on *morbidity* (Barker, 1995[15]; Almond and Mazumber, 2005[16]; Johnson and Schoeni, 2011[17]), premature *mortality* (Van der Bergh, Lindeboom and Portrait, 2006[18]), and *psychological outcomes* (Fletcher, 2011[19]).

Some studies examined the effects of birth weight or low birth weight on *the IQ score* (Kormos et al., 2013[20]), *early-life education* (Lin and Liu, 2009[21]; Fletcher, 2011[19]; Torche and Echevarría, 2011[22]; Figlio et al., 2014[23]; Bharadwaj, Eberhard and Neilson, 2018[24]), *completed education* (Royer, 2009[25]; Currie and Moretti, 2007[26]; Conley, 2000[27]), or on *the long-run returns to birth weight measured by earning or wages* (Behrman and Rosenzweig, 2004[28]; Cook and Fletcher, 2015[29]; Currie and Hyson, 1999[30]; Johnson and Schoeni, 2011[31]; Rosenzweig and Zhang, 2013[32]; Black, Devereux and Salvanes, 2007[33]).

Figure 1.1 summarises the effects of low birth weight and very low birth weight on health and socioeconomic impacts. Compared to children born with normal birth weight, children born with very low birth weight are 10% more likely to suffer from neurosensory problems, 9% more likely to experience behavioural and social competence issues, 20% more likely to have learning disabilities and 5% more likely to have intellectual disabilities. These higher health risks for children born with very low birth weights affect IQ as well as earning, wages and their quality of life over the long run.

[1] A low birth weight is a birth weight of less than 2 500 g (or 5.5 pounds); A very low birth weight is a birth weight of less than 1 500 g (or 3.25 pounds); SGA = Small for gestational age.

Figure 1.1. Generalised links between chemicals, birth weight, health and socioeconomic impacts



Notes: LBW stands for low birth weight (< 2 500 g) and VLBW stands for very low birth weight (< 1 500 g). **Share of children with normal birth weight who have the health issue. *Share of children with LBW who have the health issue. #Share of children with VLBW who have the health issue.
Source: own literature review of the authors.

One key challenge for chemicals risk management relates to the monetisation of health benefits expected from actions to curb emissions of and exposure to substances linked to (very) low birth weight. Balancing the expected benefits against the costs of regulation is typically done using WTP values as inputs for cost-benefit analyses. However, despite its significant consequences on health, there is a very limited number of studies that elicit WTP for a reduction in the risk of low birth weight and very low birth weight. This is even more problematic for very low birth weight as it can lead to more serious and more frequent health impacts.

1.2. Previous work

Only three studies that aimed at valuing lowered birth weight were identified. A study by Clarke, Oreffice and Quintana-Domeque (2017[34]; 2019[35]) used a discrete choice experiment with attributes on the costs, the season in which the baby is born, the sex of the child, and the weight of the baby at birth, with the weight levels randomised within the normal range. Restriction on normal weight in this particular study was motivated by two reasons. First, a continuous measure of birth weight has a greater explanatory power for a large range of variables than a low birth weight indicator (Black, Devereux and Salvanes, 2007[33]); and, second, following recent evidence, marginal increases in birth weight within the normal weight range are found be particularly important for well-being (Royer, 2009[25]; Clarke, Oreffice and Quintana-Domeque, 2017[34]). However, this study valued birth weight with a primary focus on well-being and demand for season of birth, which is more relevant for developing countries where malnutrition is the leading cause of low birth weight.

Two other studies relied on the non-attribute stated preference approach also called "contingent valuation method" using the dichotomous choice elicitation format. The first of the two studies were commissioned by the European Chemicals Agency with an original survey conducted during February and June 2014 in four European Countries (the Czech Republic, Italy, the Netherlands, and the United Kingdom), see Ščasný and Zvěřinová (2014[36]). The second study was prepared by the same research team for Health Canada with a slightly upgraded survey instrument through which respondents from Canada were interviewed during February – March 2016 (Ščasný and Zvěřinová, 2016[37]).

Specifically, Ščasný and Zvěřinová (2014[36]; 2016[37]) used the double-bounded dichotomous choice question to elicit WTP for reducing the probability of a respondent's child (or all children) to be born with very low birth weight using both a private good as well as a public good contingent scenario. The private contingent product was linked to a combination of vitamins and minerals that would deliver a reduction in the risk of low birth weight. In the public good scenario, risk reduction was a consequence of using "chemical-free products" due to the introduction of a stricter policy at the EU level in Ščasný and Zvěřinová (2014[36]) or due to "new regulations" in Canada in Ščasný and Zvěřinová (2016[37]). Additionally, preferences for reducing the probability of a respondent's child (or all children) to be born with low birth weight were also elicited in the 2016 study for Canada (Ščasný and Zvěřinová, 2016[37]).

In both studies of Ščasný and Zvěřinová, VSC of low birth weight and very low birth weight were derived from the estimate of WTP for avoiding 1 in 1 000 risk of (very) low birth weight using data from the first single-bounded binary choice.[2] For the private good scenario, mean VSC of very low birth weight was estimated for the four European countries in a range of EUR PPP 80 000 to 245 000, with a value of EUR PPP 126 200 for EU28, based on a benefit transfer method. VSC of very low birth weight is about EUR PPP 202 000 in Canada. VSC of very low birth weight is higher for the public good scenario, also because it includes all children as beneficiaries, and a central estimate of VSC for very low birth weight at EU28 derived for the public good scenario is EUR PPP 548 000 (for the general population). VSC of low birth weight is only available for Canada and it is EUR PPP 136 000 for the private good and EUR PPP 221 000 for the public good scenario (see Table B.1).

It is important to note that surveys to elicit WTP to reduce the risk of negative health impacts can be impacted by the overall health situation in the population at the time of implementation. Ščasný and Zvěřinová (2016[37]) was carried out in Canada during the world-wide spread of the Zika virus that affected mainly the health of the new-borns.[3] Some of the respondents who considered Zika and its effects when completing the survey reported a higher WTP value, which in turns impacted the resulting mean WTP.

1.3. Current effort: SWACHE project and selection of very low birth weight

The present valuation study is carried out within the first round of the OECD Surveys on Willingness-to-Pay to Avoid Negative Chemicals-Related Health Effects (SWACHE) project described in Box 1.1. Given the absence of internationally comparable WTP estimates for the risk of very low birth weight and its association with many chemicals, it was identified as one of five priority health endpoints for valuation through SWACHE, along with chronic kidney disease, asthma, infertility, and IQ loss.

[2] In Ščasný and Zvěřinová (2014[36]; 2016[37]), a double bounded dichotomous choice was used to elicit WTP. However, Ščasný and Zvěřinová (2014[36]; 2016[37]) computed

[3] In May 2015, the Pan American Health Organisation issued an alert regarding the first confirmed Zika virus infection in Brazil. Later, the World Health Organisation declared the epidemic of Zika virus a public health emergency of international concern on February 1st, 2016 that was followed by a strong media warning of Zika virus infection that may have adverse effect on new-borns. Zika was diagnosed mainly in South and Latin America, but there were several cases also reported in other parts of the world, including Canada.

The OECD recruited a panel of prominent experts to develop a common general approach to valuing endpoints through stated preference methods, while still allowing the survey instruments for each endpoint to be specifically tailored. Draft survey instruments were reviewed by the expert panel as well as delegates from member countries in September 2019 and April 2020 and surveys were revised each time based on comments received. As the surveys evolved through focus group and one-on-one interview testing, as well as reviews by health professionals and other experts, additional less-formal discussions among the expert panel were held to help ensure the survey instruments elicit the WTP of respondents using adequate and appropriate stated preferences methods.

The main aim of the SWACHE project is to provide health benefit estimates for use in applied benefit analysis. This study extends and enriches the previous two studies commissioned by ECHA and Health Canada by increasing the geographical coverage by of WTP to nine OECD countries. In comparison to the previous two studies, this study provides WTP and VSC values derived from a private good scenario linked to regulation of chemicals in products, which makes the results more relevant for regulators than the scenario linked to a new novel combination of vitamins used in the previous studies.

The main results provided in this study are the estimates of VSC of very low birth weight for nine OECD countries, including provision of a range of estimates based on various methodological assumptions. Also included is a sensitivity analysis to the level of baseline risk and consideration by the respondent of effects other than very low birth weight in the valuation. The association of the WTP with sociodemographic (including income elasticity) and health characteristics are examined.

This paper is organised as follows. First, the survey design and methods are described in Section 2. Section 3 provides information on data gathering and descriptive statistics. Section 4 gives details on the empirical strategy used to analyse survey data, the approach to estimate WTP and to derive values of statistical case. The survey results are presented in Section 5. Recommended values of statistical case of very low birth weight are summarised in Section 6. The conclusions of the paper, summary and limits of the study are presented in Section 7.

Box 1.1. The OECD SWACHE Project

Chemicals are part of our daily life and must be soundly managed to limit risks to human health and the environment. While countries around the world are setting up legal frameworks to address these risks, the cost of policy inaction is still poorly understood. Assessment of chemicals management options and environmental policies can be considerably improved by better estimating their costs and benefits. The resourcing of national chemicals management programmes also often requires economic justification of the benefits of such investment. However, current socio-economic analyses of chemical regulations use values for morbidity impacts that are often incomplete. In most cases, these values cover only lost productivity, lost earning or cost-of-illness and disregard the disutility costs of pain and suffering from the illnesses (Navrud, 2018[38]).

The OECD project Surveys on Willingness to Pay to Avoid Negative Chemicals-Related Health Impacts (SWACHE) brings together expertise on chemical safety and economic analysis to fill this gap. The project aims to establish internationally comparable values for the willingness-to-pay (WTP) to avoid negative health effects due to exposure to chemicals. Such values can be used to demonstrate and measure the economic benefits of minimising the impacts of chemicals on human health.

The only way to capture the full WTP to avoid illness is to conduct a stated-preference study, i.e., surveys where individuals are asked to report their WTP to reduce their risk of negative health impacts due to chemicals exposure. Contingent valuation methods and discrete choice experiments do just that, and WTP figures based on these methods have been used in assessment efforts (Alberini, 2017[39]). To derive WTP values, surveys of a large number of citizens of countries have therefore been conducted under the SWACHE project. Particularly, these stated preference surveys provide data that can shed light on the disutility in terms of symptoms and lower quality of life of a given disease or health effect, which is not captured by existing metrics such as those based on the cost of illness.

The SWACHE project is organised in two rounds focusing on 5 health effects each. The first round of health effects includes asthma, fertility loss, IQ loss, chronic kidney disease and very low birth weight. The first round of surveys was implemented in 2022 in at least five countries each where representative samples of at least 1 200 respondents each were collected. Overall, one to five of the surveys were implemented in 22 countries, totalling 46 surveys conducted. Survey responses are empirically analysed to estimate mean WTP for a given reduction in health risk for each country surveyed.

The results of this first round are presented in five working papers, one for each health effect. The research described in individual working papers makes a variety of empirical contributions to health valuation in the context of chemicals exposure, although, by design, the approach was not to break new conceptual, theoretical, or econometric ground. Moreover, the comparison of the estimated WTP across health effects and across countries will be carried out in a separate summary paper, which will also provide guidance for the transfer of WTP value over time and to non-surveyed countries

2 Survey design

2.1. General SWACHE approach to survey design

The survey on very low birth weight followed the general SWACHE Approach to Survey Design described in Box 2.1. This harmonised approach comprised development of the survey in all phases, such as description of the health effect to be valued, a risk reduction mechanism, a payment vehicle and an elicitation method, risk explanation, harmonised sociodemographic and debriefing questions, the survey adaptation to different countries, pretesting, and data gathering.

2.2. Definition and description of very low birth weight

Very low birth weight was defined in the survey as birth weight less than 1 500 g. The prevalence of very low birth weight varies from 0.6% to 1.3% depending on country (see Figure A.1 of Annex A).[4] Among countries for which data is available, Hungary has the highest prevalence (1.3%) and Finland, Iceland, and Mongolia have the lowest (0.6%). Based on prevalence data from the United Nations' database and from Eurostat, three base levels of prevalence were chosen for the survey: 0.6%, 1%, or 1.2%. One of these baseline levels was assigned to each respondent at random. Respondents were introduced to the concept of probability, explained the prevalence and shown a figure illustrating the probability of a child to be born with a very low birth weight (see Figure A.5 in Annex A).

Very low birth weight infants experience many more health and developmental difficulties than infants with normal birth weight. This study focuses on three types of health problems that may occur if a child is born with very low birth weight: (1) neurosensory problems, (2) behavioural and social competency problems and (3) learning disabilities. These health problems were described to respondents of the questionnaire survey before the valuation questions (see Figure A.2, Figure A.3 and Figure A.4 of Annex A). For each type of health problem, the survey provided information on different related health issues, treatment, quality of life impact, and prevalence of the health problems in comparison to infants with normal birth weight (Hack, Klein and Taylor, 1995[14]).

[4] The figures show prevalence for all countries for which data were available either in the United Nations' or Eurostat' database.

Box 2.1. Development of SWACHE survey questionnaires and application of best practices

Each SWACHE survey questionnaire was drafted by a team of authors that includes recognised experts in the field of stated preference surveys related to health impacts as well as practitioners in the socio-economic analysis (SEA) of chemicals management options.

Each survey questionnaire was developed in several steps. First, a description of the health effect (endpoint) was drafted including information about the related quality-of-life health impact, a review of any prior stated preference studies on the same health effect and suggestions for how to characterise the endpoint in a new study. Second, various valuation scenarios were developed describing the target population, the risk reduction mechanism, the payment vehicle and the elicitation method. Third, a complete draft survey questionnaire was developed including the most appropriate valuation scenario.

A steering group of experts including internationally renowned academics, SEA practitioners, regulators and health professionals provided regular feedback throughout the process. The final working papers were reviewed by the expert group as well as by country delegations as per the OECD review process.

All SWACHE survey instruments featured a harmonised introduction that contains language to minimise non-response bias and comply with ethics principles:

> *Welcome!*
>
> *This survey is part of an international initiative coordinated by the Organisation for Economic Co-operation and Development (OECD) that aims to help design better policies.*
>
> *The survey asks for your views about a proposal to reduce the risk of [health effect] due to the exposure to chemicals and chemical products.*
>
> *Please read all the information and answer the questions carefully.* ***There are no right or wrong answers to the questions asked in this survey. It is your honest opinion that matters to us.*** *The survey can be completed on a mobile device, but we recommend doing it on a larger device, such as a tablet, laptop or desktop.*
>
> *We will ask some questions related to your health, habits and attitudes. Rest assured that a "Prefer not to answer" option will be available for you to select, at your discretion.*
>
> *Your answers throughout this survey will be kept* ***confidential****. Participation in the survey is* ***voluntary*** *and you may withdraw consent at any time by writing to support. Before agreeing, please also read this information sheet [hyperlink to information sheet screen].*

The informed consent of all participants to the surveys was collected by the internet panel provider. All survey response data are anonymised and participation in the survey was voluntary. In addition, best practices in terms of safe data storage are applied.

A description of the SWACHE project and the first five draft questionnaires were submitted to an institutional review board, the Inserm Ethics Evaluation Committee (CEEI), for an external, independent ethics review. The submission process included a detailed description of the research project including type of data collected, measures to protect personal data, research objectives, research hypotheses and methodology. CEEI gave a favourable opinion on the project and had no significant concerns.

All survey questionnaires also include language to minimise non-response bias within the questionnaire. For example, the following language reduces the risk of "yea"-sayers:

Please keep these things in mind

In surveys such as this one, people sometimes say that they would pay for a reduction in risk even if they cannot afford it.

Please treat the following questions as if they were a real-life situation, so that your answers are as accurate as possible.

Don't agree to pay an amount that you cannot afford to pay or if you feel that there are more important ways to spend your money.

When answering the next questions, please consider:

your personal income and savings

that the payment would reduce your spending on other things you may value.

All surveys included harmonised debriefing questions to collect data on predictors of WTP such as income and age but also questions to control for non-response bias in empirical analysis. For instance, respondents were asked how much they agree with the following statements:

- I responded to the survey as I would have done in real life.
- The survey provided me with enough information to make informed choices.
- Did you agree or disagree with the description of [health effect] provided in this survey?

All survey questionnaires included a series of debriefing questions specific to the health effect valued in order to capture potential co-benefits or protests linked to the risk reduction mechanism. These survey specific questions are described in individual working papers.

Finally, all draft surveys questionnaires were tested in at least ten one-on-one interviews with people of various background and characteristics in an English-speaking country and in a non-English speaking country. The survey questionnaires were programmed and extensively tested. The translation into languages of target countries was verified by native speakers. Some surveys benefited from a pre-pilot to further revise the survey questionnaires.

Each survey questionnaire was piloted in all target countries with 50 survey responses per country. The pilots allowed for calibration of the bid levels that were presented to respondents to maximise the even distribution of responses across the four possible outcomes of the double bounded dichotomous choice.

2.3. Risk reduction mechanism

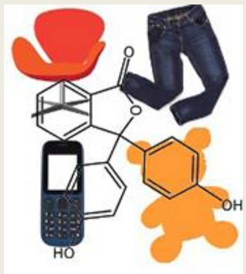
The contingent product used in the survey is a private good specified as new safer products that do not contain toxic substances that could increase the probability of very low birth weight (for further details see Box 2.2). Thus, the use of these products reduces the probability of children being born with a very low birth weight and hence the probability of suffering from the three types of adverse health effects presented in Section 2.2.

As the study aimed to work with a realistic risk reduction mechanism, respondents were informed that in order to have a positive impact on the development of their child the "safer products" had to be used during pregnancy for 8 consecutive months. Such a contingent product is, however, relevant only for people who intend to have a biological child in the near future. Thus, the target population was limited to population of a reproductive age (18-44 for women and 18-65 for men)[5] who would like to have a biological child within the next 5 years.

Box 2.2. Description of risk reduction mechanism provided in the questionnaire

"New safer products"

Scientific studies have found that a variety of chemicals that are currently used in your country can increase the probability of children to be born with very low birth weight.



Various products such as clothes, laundry detergents, household cleaning products, textiles, furniture, and electronics can contain such chemicals which can be absorbed into the body through skin and by breathing. Unborn children are particularly sensitive because such chemicals may have an impact on their development.

Please imagine that new safer products are introduced in your country. These products do not contain toxic substances that increase the probability to have a child born with very low birth weight and do not harm the development of children. Usage of these products would not have any other positive health and environmental effects. Using these products has no impact on the probability of miscarriage and no impact on the probability that your infant will die. Public authorities would certify the safer products.

However, companies would incur additional costs in order to comply with the requirements.

These costs would lead to higher product prices and may reduce your spending on other things.

The safer alternative will be available for all the conventional products that you buy, such as laundry detergents, household cleaning products, toiletries, cosmetics, clothes, textiles, furniture, and electronics.

If you buy and use the "safer products", the probability of your child having a very low birth weight will be reduced by FILL per 1,000 new-born children (from TT per 1,000 to YY per 1,000, as shown in the figure on the next page).

The "safer products" are used during pregnancy (for 8 months) to have the positive impact on child development. Therefore, they also lower the probabilities of the adverse health effects described previously. These effects are expected when you choose the safer alternative for almost all the products that you buy.

Note: FILL, TT, and YY were replaced with numbers according to design (see Section 2.3.)

[5] The target population is women of age 18-44 and men of age 18-65. Including respondents younger than 18 would be problematic due to several reasons. First, the legal age for giving consent for participation in a survey differs for countries (it seems to be 14 in Italy and 16 in the Netherlands, for example), which will lead to different samples according to age. Second, very few, if any, young people of age 15-16 plan to have a baby within the next 3-5 years. Even if young people planned to have a baby before their age of 18, in many countries they cannot be legally responsible for the baby (for example, in the Czech Republic, although the legal age for having a baby is 15, if a girl gets pregnant under 18 years, she cannot be legally responsible for the baby and the same holds for a father of the same age, and in that case parents of the young pregnant girl will be usually legally responsible). Third, respondents should state their WTP considering explicitly their financial resources they have a full control over, while a majority of young people below 18 do not work yet, and hence they do not have own financial resources to allocate.

In the experiment, the risk reduction by 2, 3, 5, 6, and 7 in 1 000 from the baseline level that is 6, 10, or 12 in 1 000 was proposed. Naturally, large magnitudes of the risk reduction are not possible for the lower baseline risks. For that reason, the risk reduction levels are dependent on the baseline level in the design. It implies that the survey shows only the 2 and 3 in 1 000 risk reduction for the baseline at 6 in 1 000, the reductions by 2, 3, 5, and 6 in 1 000 for the baseline at 10 in 1 000 and all risk reduction levels for the baseline at 12 in 1 000. It means that only the risk reductions of 2 and 3 in 1 000 are applied for all three baselines levels. Using different magnitudes of the baseline risk across countries might potentially affect stated preferences. On the other hand, if WTP is affected by the baseline risk, then using the same baseline risk in the valuation experiment might make future benefit transfer problematic. To investigate this issue, a split-sample design that attributed one of the three baseline risks to each respondent at random was used.[6]

The following eleven combinations of baseline risk levels and risk reduction levels were used in the design:

- baseline risk level = 6, then risk reduction levels are 2 and 3,
- baseline risk level = 10, then risk reduction levels are 2, 3, 5, and 6,
- baseline risk level = 12, then risk reduction levels are 2, 3, 5, 6, and 7.

Since the concept of probability is used to describe health characteristics of the population as well as to elicit preferences for reducing the risk of very low birth weight, special attention is paid to how information is provided. Before the valuation question, information about the risk (incidence) was provided in terms of the number of children per 1 000 who are born with a very low birth weight in respondents' country as well as corresponding percentage. Risks were also graphically depicted using a group of 1 000 children, instead of a grid, as used in premature mortality valuation to convey randomness and the magnitude of the risk (Ancker et al., 2006[40]; Alberini and Šcasný, 2011[41]; 2018[42]; 2021[43]).

As in all SWACHE surveys, different colours that could be seen by people with colour blindness were used to show children with very low birth weight and children with normal birth weight. At the end of the tutorial, respondents took a simple quiz, asking which child has the higher probability of being born with a very low birth weight, after information about different probabilities for the two children (see Figure A.6 in Appendix) were provided. Respondents who incorrectly answered this question were considered to fail the probability quiz and these respondents were excluded from the main data analysis.

2.4. Valuation questions

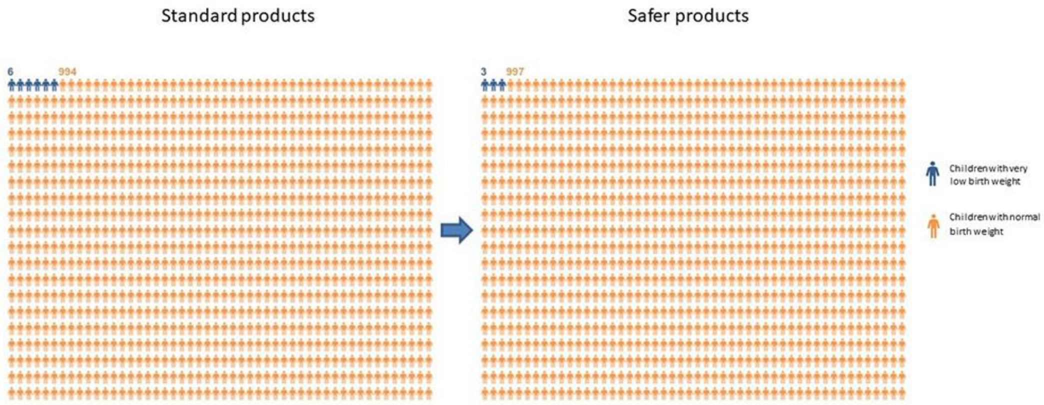
The double-bounded dichotomous choice (DBDC) question with an open-ended follow-up question was used to elicit WTP for reducing the probability of a respondent's child to be born with very low birth weight. An example of the dichotomous choice question is shown in Figure 2.1. Each respondent chose between two product bundles, one with standard products and one with safer products. The standard products were described by the baseline level of the risk the next child would be born having very low birth weight with no extra cost. In case of the safer products, the level of the risk was reduced and the extra amount of money respondents would need to pay for the offered risk reduction was displayed.

In the valuation questions, the risk reduction of very low birth weight was visualized using grids for a group of 1 000 children. While children with very low birth weight were depicted in blue, children with normal birth weight were in orange (see Figure 2.1).

[6] In the ECHA study (Ščasný and Zvěřinová, 2014[36]) the baseline level for VLBW was 15 in 1 000 children and 10 in 1 000 children born with VLBW was used in the Canadian study (Ščasný and Zvěřinová, 2016[37]).

Figure 2.1. An example of willingness-to-pay question

We offer you two product bundles, one with standard products and one with safer products. Please consider both the risk reduction and the additional cost when choosing between them. **Remember your household budget!**



	Buy standard products	Buy safer products
Probability of your child having a very low birth weight	6 in 1,000	**3** in 1,000
Extra expenditure during the 8 months of pregnancy	**0 GBP**	**130 GBP** per month (**1,040 GBP** in total)
Which of these two options would you choose?	○	○

The additional expenditure for the safer products was randomly determined from a vector of six starting bids varying from USD PPP 740 to USD PPP 350 (see Table 2.1). All bids were converted from USD to local currencies using Purchasing Power Parities for actual individual consumption of 2019 from the OECD.[7] Following the conventional double bounded dichotomous choice (DBDC) contingent valuation design (Carson and Hanemann, 2005[44]), respondents were offered the same bundle at a cost twice or half as high as the starting bid, depending on their answer to the first purchase decision. As a follow-up, the survey inquired about the maximum additional expenditure per month respondents were willing to make to obtain the safer products bundle.

Table 2.1. Starting bids as monthly additional expenditure for the United States

United States (USD)			Italy (EUR)		
Starting bid	Follow-up bid if starting bid rejected	Follow-up bid if starting bid approved	Starting bid	Follow-up bid if starting bid rejected	Follow-up bid if starting bid approved
40	20	80	30	15	60
80	40	160	60	30	120
130	65	260	90	45	180
170	85	340	120	60	240
280	140	560	200	100	400
350	175	700	250	125	500

Note: All bids were converted from USD to local currencies using Purchasing Power Parities for actual individual consumption of 2019. Data are provided by the OECD. https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm.

[7] Available on https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm.

2.5. Approach to minimise biases

2.5.1. Limit yea-saying

The survey used language to minimise the bias in favour of automatically responding yes to the valuation question.[8] Specifically, the following text was shown just after introducing the contingent scenario and just before the valuation question:

> *"Please treat the following questions as if they were a real-life situation, so that your answers are as accurate as possible. Don't agree to pay an amount that you cannot afford to pay or if you feel that there are more important things to spend your money on."*

After the pilot, the visualisation of the magnitude of the risk reduction and the payment was also improved by adding two boxes, one for each issue, under the risk visualisation in the grids. Although the experiment relies on a non-attribute valuation task, the survey shows the risk reduction and the costs as an attribute-based task with two clearly visible attributes to increase readability of provided information, as shown in Figure 2.1.

2.5.2. Limit co-benefits

Ščasný and Zvěřinová (2014[36]) and Ščasný and Zvěřinová (2016[37]) highlighted that a non-negligible portion of respondents considered other effects while valuing the reduction in risk of lowered birth weight. To encourage respondents to ignore other effects resulting from using the new safer products and to control for potential co-benefit in the survey responses, several statements were used. First, the survey highlighted in the scenario that:

> *"Usage of these products would not have any other positive health and environmental effects. Using these products has no impact on the probability of miscarriage and no impact on the probability that your infant will die. Public authorities would certify the safer products."*

Despite this statement, respondents might still be considering the other effects and therefore the survey asks in the debriefing section, placed just after the valuation task, the following question:

> *"We noted earlier that the new safer products would not have any other positive or adverse health and environmental effects. However, different people might think differently about possible side effects.*
>
> *When you were thinking about paying for the safer products, did you consider any other effects, positive or negative, aside from the decrease of the probability of very low birth weight?*
>
> - *No, I didn't consider any other effects.*
> - *Yes, I considered mostly other effects.*
> - *Yes, I considered some other effects."*

In the case respondents considered some or mostly other effects, they were asked another question about what the specific effects:

> *"When you agreed to pay higher price for the safer products, what other effects did you consider?*
>
> *Tick all that apply.*
>
> - *Improving the state of the environment (water, air, soil, plants and animal species)*

[8] In the literature, this is referred as using "cheap talk" language to minimise "yeah saying".

> - *Improving my health*
> - *Improving the health of other family members*
> - *Adverse impacts on the economy and employment*
> - *Reducing the probability of miscarriage*
> - *Reducing the probability of infant death*
> - *Other. Please specify: [OPEN TEXT BOX]*
> - *I actually did not consider any other effects"*

Respondents who considered other effects were asked how those effects impacted their willingness to pay:

> *"These other effects, which you took into consideration …*
> - *increased your willingness to pay for the safer products*
> - *decreased your willingness to pay for the safer products*
> - *did not have any effect on the willingness to pay for the safer products"*

2.5.3. Limit protesting

Respondents could agree not to pay for the offered risk reduction due to three very different reasons; 1) they are willing to pay for the offered good, but less than the offered price (0 < WTP > bid), or 2) they are protesting against the contingent scenario; or 3) they are not protesting against the scenario but are truly willing to pay nothing (true zero WTP, i.e. the Spike at zero). These three cases were identified using two questions included in the survey instrument.

First, respondents who decline to pay for two successively lower bids were asked what the maximum is they would be willing to pay for the safer products. There are 131 (1.21%) respondents who are willing to pay just zero (93 or 1.34% after excluding speeders and those who failed the probability quiz). These respondents constitute what is called the Spike at zero.

Second, the same respondents were asked why they responded as they did, with not protesting options numbered 1, 4, 5, 6, and 8, or protesting options numbered by 2 and 3:

"*Q123. Why did you choose the way you did?*

1. *The safer products were too expensive or my spending is too high.*
2. *I do not trust the information I have received.*
3. *The price increase of products should be covered by the government.*
4. *The effects associated with very low birth weight are not severe enough to pay to avoid them.*
5. *Reduction in the probability is too low.*
6. *I don't believe that the safer products would reduce the probability of very low birth weight.*
7. *I chose the safer products, but in a real situation I wouldn't pay.*
8. *There are more effective ways to attain the same goal (for example lifestyle changes)."*

ANSWER OPTIONS: 1. Definitely not the reason for my choice, …, 5. Definitely the reason for my choice

2.5.4. Incentive compatibility

Lastly, the analysis paid attention to:

- incentive compatible responses, that is when respondents indicated that they responded in the survey as they would have done in real life (Q22_1), chose the safer products, but in a real situation they wouldn't pay (Q123_7; see bullet 7 in the question on zero WTP vs. protesting just above),
- whether respondents thought that the survey provided enough information to make informed choices (Q22_2), and how confident respondents were that the information that has been provided in this survey was correct (Q22x).

3 Survey data

3.1. Pretesting and data collection

The survey was implemented in nine OECD countries from January to February 2022 (for the surveys carried out in eight countries) and from May to June 2022 (for the survey conducted in Canada). Inhabitants of these countries aged between 18 and 65 were surveyed. For this survey, the target population was males (aged 18-65) and females (aged 18-45) who were planning to have a biological child within the next 5 years, including those who were currently expecting a child and wished to have another child within the next 5 years. A total of 10 800 high quality interviews were completed, 1 200 in each country surveyed (Table 3.1). The selection of respondents was based on quotas matching key demographic characteristics (gender, age group, level of education and geographic region) to ensure representativeness.

Table 3.1. Description of the sample, speeders, and who failed the probability quiz

	Final sample	Pilot	Speeders	Failed Prob quiz	no speeders, no failures
Canada	1 200	50	18.1%	22.0%	795
Czech Republic	1 200	198	11.8%	21.9%	838
Italy	1 200	200	14.2%	21.2%	829
Mexico	1 200	50	16.8%	26.4%	754
Netherlands	1 200	200	20.2%	27.4%	732
Switzerland	1 200	50	19.9%	22.3%	765
Türkiye	1 200	50	23.2%	32.2%	627
United Kingdom	1 200	206	17.0%	18.6%	833
United States	1 200	50	16.4%	25.7%	767
All	10 800	1 054	17.5%	24.2%	6 940

The fieldwork was conducted in two phases: pilot phase, and main stage. A first pilot (50 completed interviews) was conducted in Mexico, followed by the second pilot (n = 50 in each) conducted in all countries but Mexico and Canada. After review of the pilot data, which resulted in a high share of Yes-Yes response, the bid values for all countries were increased. These new bid values were used for a third pilot providing an additional n = 148, and 150, respectively, completes from the Czech Republic and Italy. A fourth pilot for n = 150 completes was then conducted in the UK and the Netherlands. Based on these results, the final bid values were set, and mainstage fieldwork began on 14 January 2022 and end on 1 March 2022. A final pilot with n = 50 respondents was conducted in Canada in April 2022 prior to the main wave of data collection there, in May and June 2022.

The fieldwork, covering both the pilot and the main stage, was carried out by Ipsos European Public Affairs (hereafter Ipsos) in all surveyed countries. The surveys were conducted via Computer-Assisted Web Interviewing (CAWI) (for more details on the internet panels see Box 3.1).

Box 3.1. Quality of the internet panels used in SWACHE

The field implementation of the SWACHE surveys was carried out in all surveyed countries by Ipsos European Public Affairs (hereafter Ipsos), selected after a careful call for tender process. Ipsos has significant experience in multi-country projects and maintains panels of respondents in many countries. Fieldwork, pilot and main stage, took place between June 2021 and June 2022 for the first round of surveys. The surveys were conducted via Computer-Assisted Web Interviewing (CAWI). Random samples of at least 1 200 respondents matching the target population were drawn for each country from a high-quality network of online access (non-probability) panels. Some surveys had specific requirements regarding the target population due to the endpoint under consideration. This is elaborated in survey-specific information.

Online panels are databases of potential participants who declare that they will cooperate for future data collection if selected, generally in exchange for a reward or incentive. Loyalty card and subscription databases are included here if there is a continuous relationship with members who understand the commitment asked of them. Ipsos has its own supply of sample through its globally managed i-Say (IIS) panels and some locally owned Ipsos panels. In addition, Ipsos partners with many different types of external suppliers to source sample when needed to fulfil project requirements. This includes other traditional research panels, reward or loyalty communities, intercept or offer wall providers, and sample exchanges. Ipsos can also leverage its Direct-to-Survey channel which accesses respondents directly through social media platforms. To reach respondents, Ipsos has a proprietary project management and workflow system that controls access to their panel assets and where necessary, external respondent sources.

Importantly, Ipsos implements procedures to make sure that respondents to surveys are real, unique, engaged and fresh. To ensure that their respondents are real, i.e. they are who they claim to be, Ipsos uses country geo-IP validation and digital fingerprinting to check if the respondent used a device that is truly located or if it is evading detection and also if the respondent's device has any past history of fraud. These tools used in combination with cookies can make sure that each respondent is unique and has not already accessed the survey. To guarantee respondents are engaged, their survey taking behaviour is evaluated in real time, through standard self-adjusting algorithms involving speeding and straight-lining detection (i.e., always choosing the first (or nth) answer in multiple choice). The worst offenders are automatically removed from the data deliverables and are not counted against quotas. Finally, Ipsos invited members of their panels that were fresh, i.e., that have not taken part in any of the other SWACHE surveys and were not overburdened with surveys in general.

After the main stage was completed, the online survey data were evaluated by Ipsos using several quality markers that feed into an overall quality score for each respondent: survey length and speeding, straight lining and proportion of "don't know" answers.

When sampling from access panels, quota sampling is used. Via quota sampling, respondents are selected in such a way that the achieved sample is representative of the target population. For this SWACHE survey, sampling quota were set based on gender, age (18-24, 25-34, 35-39, 40-44, and 45-65 year-olds), level of education (low or medium vs high level of education) and geographic region. As is common practice across all major panel providers, Ipsos uses a survey router, a software system that allocates willing respondents to surveys for which they are likely to qualify.

The questionnaire started with a set of screening questions to assess eligibility in terms of age and plans to have a child. Respondents who were not planning to have a child within the next 5 years were not eligible for this survey. Demographic questions were also asked to assess to which sampling quota a respondent belongs – in terms of gender, age, level of education and geographic region. Only respondents who passed the screening questions and belonged to a sampling quota cell that was not yet full, could continue to questionnaire.

The break-off rate, i.e. a proportion of survey participants who partially completed questionnaires, was 14.4%; the Czech Republic, Mexico and the Netherlands had relatively high break-off rates, between 22% and 27%. About 2.6% of respondents did not pass a lower threshold for the quality score and were removed from the final dataset.

The median survey duration varies between 14.8 minutes (in the Netherlands) to 27.8 minutes (in Mexico). One quality marker of the low-quality check was survey length. A valid complete is one where the time spent by a respondent on the questionnaire is not lower than one third of the median survey duration. Later, following *"SWACHE Core principles of survey analysis"*, speeders that are defined by the time spent by a respondent on both the questionnaire and the valuation section. Specifically, respondents were screened out if their time spent on the survey was lower than 48% of the country-specific median duration, according to the recommendation of Survey Sampling International (Mitchell, 2014). This approach helped identify 17.5% speeders, 12% in the Czech Republic, 20% in the Netherlands and Switzerland, and 23% in Türkiye.

For each of the quota variables, a comparison between target population and (unweighted) samples in identical categories is provided in Table C.1 in Annex C. It shows that lower educated respondents are underrepresented in all countries among the final respondents, while medium educated respondents are overrepresented, especially in Italy, Mexico, and Türkiye. Regarding gender, males are underrepresented in four countries, but not in Czech Republic, Italy, Mexico, Switzerland, and Türkiye. There are relatively younger respondents below age of 25, while respondents between 25 and 34 years are underrepresented in all countries.

Therefore, after data collection, a post-stratification weighting procedure is carried out to adjust the sample to selected population totals. Data for this survey are weighted to match official population statistics on gender crossed by age, educational level and geographic region. Specifically, the raking procedure is performed by iterative proportional fitting using contingency table analysis, and any weights larger than 3.0 are automatically set to equal to 3.0 at the end of each iteration of the algorithm.

Overall, out of 10 800 finalised interviews included in the sample provided by Ipsos, there are 17.5% speeders (11.8% in Czech Republic to 23.2% in Türkiye) and 24.2% respondents who failed the probability quiz (18.6% in United States to 32.2% in Türkiye). Excluding both speeders and those who failed the probability quiz gives a total of 6 940 observations for the main modelling (excluding 35.7%, 30.2% in Czech Republic to 48% in Türkiye), see Table 3.1.

3.2. Summary statistics on the main regressors

Table 3.2 displays statistics for the key variables used in the analysis (on the left), compared with the sample statistics before excluding speeders and respondents who failed the probability quiz (on the right). In the sample (including speeders and those who did not pass the quiz), there are more females (55%), about 39% respondents had a university degree and 40% completed upper secondary education. On average, household monthly income is USD PPP 4 679 or USD MER 4 236 expressed in market exchange rate (MER).

About 53% of the survey participants were married at the time of the survey. Most respondents living with a partner have been in their relationship for more than 6 years, about 15% have been in a relationship between 6 months and 1 year and only 5.3% have been in the current relationship less than 6 months.

About 41% of respondents do not have a child yet, for 34% the next baby will be their second child, and 25% had more than one child at the time of the survey.

All survey participants plan to have a baby within the next 5 years. The more distant plan time, the less realistic answers might be. Most of the survey participants are however planning to have a baby within the next three years (27% the next year, 38% within the next two years, and 20% within the next three years), with only 7% and 9% planning to have a baby later, in four, and five years, respectively.

With respect to health status, 44% think their health is above average relative to others of the same age and gender, while only 11% think their health is below this average. If they need some form of medical treatment, 12% would have to cover most of these costs out of their own pocket and 48% can't rely solely on public health insurance. About 28% have y had COVID-19, and 54% respondents stated that their relatives have had COVID-19.

Table 3.2. Sample descriptive statistics

	All (N = 10 800)	no speeders, no quiz failures (N = 6 940)
Socio-demographic		
female	0.550	0.565
university	0.390	0.410
upper secondary	0.395	0.385
AGE 18-24	0.231	0.221
AGE 25-34	0.461	0.469
AGE 35-39	0.213	0.215
AGE 40-44	0.073	0.073
AGE65 (only males)	0.022	0.022
married	0.534	0.510
couple in a relationship (years)	6.167	6.521
… less than 6 months	0.053	0.037
… between 6 months and 1 year	0.148	0.115
h.income (USD PPP, a month)	4 679	4 760
h.income (USD exch. rate, a month)	4 236	4 322
info about income not provided	0.058	0.061
About children		
planning to have a baby next year	0.265	0.255
… within two years	0.382	0.380
… within three years	0.197	0.200
… within four years	0.071	0.068
… within next five years	0.085	0.097
childless couples	0.411	0.431
having one child	0.344	0.334
having more than one child	0.245	0.235
Health status		
Had covid	0.277	0.291
Relatives had covid	0.537	0.574
Your health: below average	0.112	0.103
Your health: above average	0.436	0.427
Medical costs: out of own pocket	0.115	0.127
Medical costs: can't rely on public insurance	0.478	0.453

Monthly household income (after tax) across the nine countries ranges from USD PPP 1 378 in Mexico to USD PPP 7 269 in Switzerland. Differences in household income are even larger when incomes are expressed by market exchange rate; in that case mean income by country ranges from USD MER 673 in Mexico to USD MER 8 824 in Switzerland. Household income does not differ for the whole sample (on the left) or the sample from that speeders and who failed probability quiz are excluded (on the right in Table 3.3).

Table 3.3. Household income by country

	All obs.			Speeders and those who failed the quiz excluded		
	National currency	USD PPP	USD MER	National currency	USD PPP	USD MER
Canada	7 211	5 604	5 751	7 137	5 547	5 692
Czech Republic	49 121	3 856	2 266	49 602	3 894	2 288
Italy	2 350	3 647	2 779	2 393	3 714	2 831
Mexico	13 649	1 378	673	14 000	1 413	691
Netherlands	4 056	5 352	4 797	4 038	5 328	4 776
Switzerland	8 064	7 269	8 824	8 272	7 457	9 052
Türkiye	6 724	2 580	760	7 207	2 765	814
United Kingdom	7 452	7 452	7 452	7 441	7 441	7 441
United States	3 493	5 230	5 230	3 508	5 252	5 252
All		4 679	4 236		4 760	4 322

3.3. Risk reduction

In the valuation experiment, risk to be born with very low birth weight is reduced by 2, 3, 5, 6, and 7 in 1 000 and the risk reduction level was randomly attributed to each respondent. However, due to the split-sample treatment on the baseline risk (6, 10, and 12 in 1 000), the offered risk reduction levels were not applied evenly in the valuation tasks as detailed in Section 2. However, the offer risk reduction is applied evenly for each of the baseline risk level taken separately (see Table 3.4). In the design there is about the same number of observations for each combination given by risk reduction and baseline risk level. It implies that while the risk reduction by 2 and 3 were applied each in about 27% tasks, the reductions by 7 in 1 000 were applied in 9% tasks only.

Table 3.4. Number of observations by risk reduction levels used in the valuation tasks, all observations

	Baseline risk levels			
∆RISK	6	10	12	Total
2	982	984	982	2 948
3	987	987	977	2 951
5	NA	983	982	1 965
6	NA	979	979	1 958
7	NA	NA	978	978
Total	1 969	3 933	4 898	10 800

Table 3.5 shows that about 64% respondents agreed to pay the offered price for the program in the first binary choice question and this proportion increases with the risk reduction levels, from 61% (with ∆RISK = 2) to 65-68% (for ∆RISK = 6 and 7) (speeders and quiz failures excluded). In the second binary questions, 46% of the respondents always agreed to pay and 22% always refused to pay, with remaining 18% and 14% for ‘yes-no’, and ‘no-yes’ responses, respectively.

Table 3.5. Proportion of responses by risk reduction level (speeders and quiz failures excluded)

∆RISK	‘yes-yes'	‘yes-no'	‘no-yes'	‘no-no'
2	43.2%	18.2%	13.6%	25.1%
3	45.1%	17.5%	15.2%	22.2%
5	46.9%	18.6%	14.6%	19.9%
6	50.3%	17.3%	13.6%	18.7%
7	47.1%	18.3%	14.6%	19.9%
Total	46.1%	17.9%	14.3%	21.7%

4 Empirical strategy

4.1. General SWACHE empirical strategy

The data were cleaned and analysed based on the core principles for empirical analysis that were formulated by the SWACHE researchers (see Box 4.1).

Box 4.1. Consistent analysis of survey responses across SWACHE health effects

Each focused on a specific health effect, the SWACHE working papers will ultimately feed into an OECD summary paper that will gather the recommended estimates for WTP values and Value of a Statistical Case (VSC) for all endpoints, compare them across countries, and offer comprehensive guidance for benefit transfer. Consequently, the different teams involved in the SWACHE project adopted a similar core strategy on how the data would be cleaned and analysed empirically to allow the proper comparison of WTP values across countries and endpoints. A series of consensus meetings with the teams of survey authors led to the adoption of a set of Core Principles of Survey Analysis that are applied but adapted, when necessary, to survey specificities and data. As indicated in Box 1.1, the idea is not to break new conceptual, theoretical or econometric ground, but to set up core principles that are consistent with and widely recognised in the economic valuation literature. These shared principles ensure that all the working papers apply the same empirical strategy in terms of data cleaning, screening of respondents, specification, estimators, robustness checks and guidance on which central WTP or VSC value should be used in regulatory impact analysis. The final version of these Core Principles of Survey Analysis is presented in Annex D.

4.2. Screening strategy

4.2.1. Speeders

To ensure informed preference elicitation, the study applied a two-staged screening process based on a set of core principles for the empirical analysis agreed upon by the SWACHE researchers (see Box 4.1): (1) initial screening by the service provider based on several quality markers that feed into an overall quality score; and (2) additional screening based on indicators of uninformed preferences. In the initial screening, Ipsos excluded finished interviews with low quality (see Section 2). Following the “SWACHE core principles of survey analysis” presented in Annex D, speeders, defined as people completing the entire questionnaire and the valuation section faster than 48% of the median in their respective country, were also excluded from the analysis, see Table 3.1.

4.2.2. Probability quiz failures

Since the concept of probability is used to describe health characteristics of the population as well as to elicit preferences for reducing the risk of very low birth weight, special attention is paid to communicating information about risk. Prior to asking the valuation question, a tutorial on risk was provided to respondents. Information about the risk (incidence) was provided in terms of the number of children per 1 000 who are born with a very low birth weight in the respondent's country as well as the corresponding percentage.

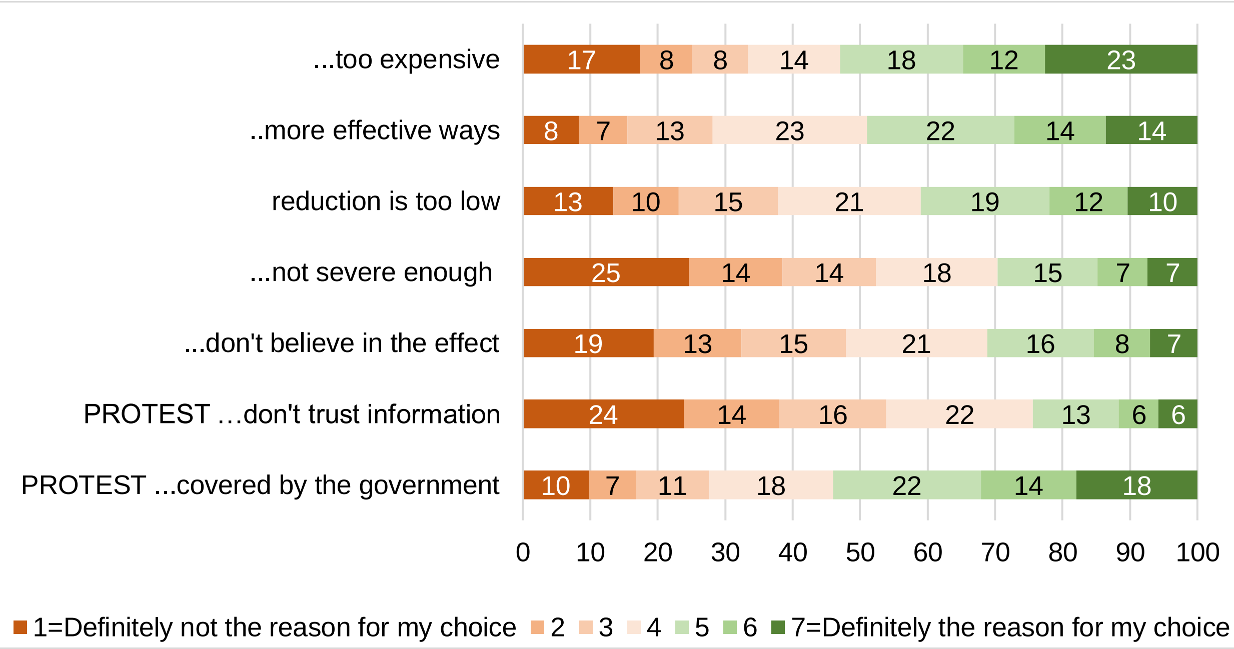
Risks were also graphically depicted using a group of 1 000 children, instead of a grid, as used in previous premature mortality valuation studies, to convey randomness and the magnitude of the risk (Ancker et al., 2006[40]; Alberini and Ščasný, 2011[41]; 2021[43]; 2018[42]). At the end of the tutorial, respondents took a simple quiz that presented risk profiles of two children and asked which child has the higher probability of being born with a very low birth weight (see Figure A.6).

Respondents who incorrectly answered were considered to fail the probability quiz. The proportion of survey participants who failed the probability quiz ranged from 18.6% in the United Kingdom to 32% in Türkiye, with sample average at 24.2%. These respondents were excluded from the main data analysis.

4.2.3. Protesters

The debriefing section elicited the reasons why respondents chose to pay or not for the new safer products, allowing to tick all that apply. Most of the respondents selected a reason that was classified as 'not protesting' (Figure 4.1).[9]

Figure 4.1. Reasons why respondents chose to pay or not for the new safer products (%)



Out of eight options offered in the questionnaire, a respondent is identified as a protester if one of the following two options was selected: "I do not trust the information I have received", "The price increase of

[9] The safer products were too expensive or my spending is too high; The effects associated with very low birth weight are not severe enough to pay to avoid them; Reduction in the probability is too low; I don't believe that the safer products would reduce the probability of very low birth weight; or There are more effective ways to attain the same goal (for example lifestyle changes).

products should be covered by the government". There are 12% and 32% respondents who ticked 6 or 7 on the 7-point Likert scale, where 7 means "Definitely the reason for my choice". There are 42.5% protesters among respondents who responded 'no-no' to the DBDC, and 38.6% protesters in total in the complete sample. When speeders and those who failed the probability quiz are excluded, the proportion of protestors (37%) does not change significantly (see Table 4.1). However, the proportion of protestors gets higher if only respondents who always disagreed to pay are examined.

Table 4.1. Protesters, by country

	All (N=10800)	Speeders, quiz failures excluded (N = 6 940)	'No-No' (N = 1 506)
Canada	27.6%	25.3%	38.1%
Czech Republic	40.4%	40.2%	52.7%
Italy	36.2%	35.7%	42.5%
Mexico	32.6%	32.7%	38.0%
Netherlands	50.3%	52.8%	44.4%
Switzerland	43.2%	41.4%	44.0%
Türkiye	35.9%	32.7%	37.5%
United Kingdom	42.0%	38.5%	44.7%
United States	39.6%	40.0%	46.8%
All	38.6%	37.2%	42.5%

4.2.4. Incentive compatibility

The majority of respondents agreed with provided information (75%) and was somewhat or very confident that the information was correct (67%), see Table 4.2. Most of the respondents also responded to the survey as they would have done in real life (87%), with only 13% of respondents who chose the product but in real situation they wouldn't pay. Excluding speeders and those who failed the probability quiz increased incentive compatibility of the survey.

4.2.5. Co-benefits and other effects

When thinking about paying for the safer products, almost a half (47%) of respondents considered other, positive or negative, effects aside from the decrease of the probability of very low birth weight. About 22% respondents considered other effects mostly and 25% considered some of them.

Out of those who considered other effects, 69% (or 31% from the cleaned sample) reported that these other effects increased their WTP for the safer products, 14% (6% from the cleaned sample) selected that decreased their WTP, and 17% (8%) did not know whether these effects affected their WTP either way. Whether this tendency is reflected in respondents' responses when paying for the risk reduction will be investigated later in this paper.

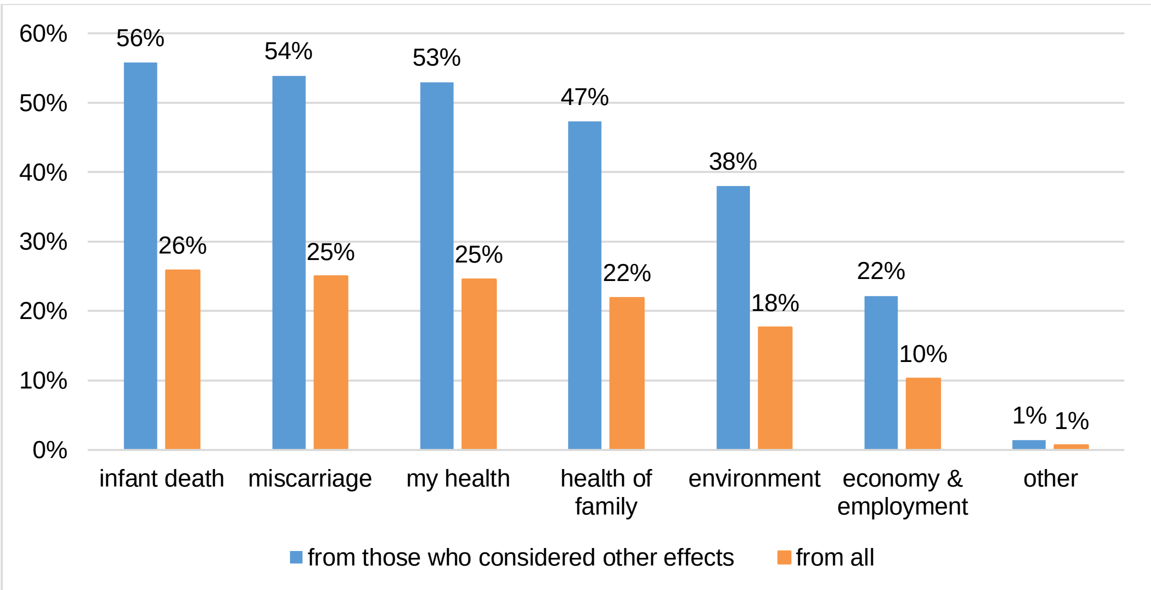
Among the effects explicitly listed in the survey instrument, most of the respondents were considering reducing the probability of infant death or of miscarriage (26% and 25% from all valid observations, 56% and 54% from those who considered other effects) and improving their own health (25% and 53%). Other effects followed: improving the health of other family members (22% and 47%), improving the state of the environment (18% and 38%) and adverse impacts on the economy and employment (10% and 22%) (see Figure 4.2). Still, 2% from those who said they considered other effects in the binary question also said later that they considered no effect when asked to choose the specific effects considered.

Table 4.2. Evaluation of the scenario information

	Definition	All	No speeders, no quiz failures	'yes-yes', 'yes-no', or 'no-yes'*
I chose the safer products, but in a real situation I wouldn't pay	4 and 5 on the 5-point Likert scale (with 5=Definitely the reason of my choice)	13%	11%	13%
I responded to the survey as I would have done in real life	Strongly agree Somewhat + strongly agree	56% 80%	63% 87%	64% 88%
The survey provided enough information to make informed choices	Strongly agree Somewhat + strongly agree	40% 75%	43% 79%	44% 80%
Did you agree with the description of health issues related to low birth weight provided in this survey?	Strongly agree Somewhat + strongly agree	32% 71%	33% 75%	36% 77%
How confident are you that the information that has been provided in this survey is correct?	Very confident Somewhat + very confident	21% 63%	20% 67%	22% 70%

Note: * if agreed at least one time to pay for the reducing risk to be born with very low birth weight.

Figure 4.2. Other effects than the decrease of the probability of very low birth weight that respondents considered when thinking about the payment (speeders and quiz failures excluded)



4.3. Baseline estimation strategy

4.3.1. WTP estimation

We posit that responses to contingent valuation questions about reducing the risk of very low birth weight, $\Delta RISK$, are driven by a random utility model. Each respondent is choosing between the status quo, i.e. the baseline level of the risk the next child will be born having very low birth weight and no cost, and the level of the risk that is reduced by $\Delta RISK$ and the amount of money they would need to pay for the offered risk reduction.

The willingness to pay is the maximum amount of money a person is ready to pay to have at least the same utility level as would be obtained with the baseline risk and disposable income the person had. To estimate the WTP, one can ask a sample of the (target) population if they would pay for a certain amount of money to reduce the risk of very low birth weight. This elicitation method is called a single-bounded dichotomous choice.

The single-bounded dichotomous choice question is the most straightforward contingent valuation elicitation format and it is also the only incentive compatible method to incentivise truthful preference disclosure. The other elicitation formats, like DBDC, may provide more information about individuals' underlying WTP. As shown by Vossler and Holladay (2018) and Vossler and Zawojska (2020), single-bounded and double-bounded elicitation formats can deliver statistically indistinguishable WTP estimates "when they are designed such that consistent economics incentives are provided, making the truthful preference statement the best response strategy for a respondent" (Czajkowski et al., 2022[45]). Hanemann et al. (1991[46]) also showed that the DBDC approach is asymptotically more efficient than the single-bounded dichotomous choice approach.

Therefore, following the SWACHE core principle of survey analysis presented in Annex D, the present study analyses the responses from the double-bounded dichotomous choice questions. At first, the first bid, BID1, is proposed to a respondent to be paid. If respondent replied yes to the first valuation question, doubled amount, $BID_2^U = 2BID_1$, is proposed in the follow-up, second, question. If respondent answered no to the first valuation question, a half of the first amount, $BID_2^L = \frac{1}{2}BID_1$, is proposed to a person in the second valuation question.

This elicitation method yields four outcomes for each respondent: yes-yes, yes-no, no-yes, and no-no. Responses provided by respondents in the survey are translated into bounds of their WTP that is not directly observable. A 'yes-yes' response to a first bid and a second bid gives that the lower bound of the WTP is BID_2^U, and the upper bound is unknown, and $WTP > BID_2^U > BID_1$. In the case of 'no-no' response, the lower bound is unknown and the upper bound is BID_2^L, so $WTP < BID_2^L$. For 'yes-no' (and 'no-yes') the lower bound WTP is defined by BID_1 (BID_2^L) and the upper bound WTP is given by BID_2^U (BID_1), hence $BID_1 < WTP < BID_2^U (BID_2^L < WTP < BID_1)$.

Then, the probability of a particular response is equal to the probability that WTP is smaller than the upper bound less the probability that WTP is smaller than the lower bound. WTP can be then estimated non-parametrically, following Turnbull lower-bound approach (using e.g. Kaplan-Meier estimator), or parametrically.

WTP models are based on the random utility framework (McFadden, 1974). It implies that the WTP estimates are inferred indirectly from estimation of the underlying utility function unless the WTP is directly derived from discrete choice model estimated in WTP-space rather than in preference-space. When the goal of a study is to provide a single estimate of WTP, a direct estimation of WTP from discrete choice model estimated in WTP-space may be preferred (see, Cameron, 1988). As noted by Czajkowski et al. (2022[45]), the direct WTP functions are more flexible and allow for avoiding in some cases implausible (distributional) consequences of indirect estimation of WTP based on preference functions estimation.

It is assumed that the responses to the choice questions are driven by the WTP for the risk-reducing alternative, which is not directly observed. The unobserved WTP depends on the magnitude of the risk reduction as follows:

$$WTP_i = \alpha + \beta \cdot \Delta RISK_i + COUNTRY \times \delta + \varepsilon_i \quad (1)$$

where $\Delta RISK$ is the reduction in the risk that the planned child will be born with very low birth weight, $COUNTRY$ is a vector of country-specific dummy variables (or "dummies"), and ε is error term distributed by assumed distribution, see above. Subscript i denote the respondent, and the coefficients α, β, and the vector of $\delta's$ are estimated. Country dummies are also interacted with the demographic post-stratification weights defined at the country level since some segments of people were slightly under- or over-represented in the sample, see above.

Estimation of the WTP distribution parameters is conditional on the assumption made on the parametric distribution for which cumulative distribution functions of distribution evaluated at the upper and at the lower bound are calculated. To check the robustness of the results, the present study compares several distributions of the errors, specifically log-normal, log-logistic, and Weibull. The parameters of WTP distribution are estimated using the maximum likelihood method in which the sum of the probabilities for all individuals is maximised.

Following the SWACHE core principles, the mean WTP are derived by recovering the individual level WTP estimates for each observation. The individual mean WTP is computed by integrating the probability of an individual responding yes to the valuation question over the interval from 0 to the maximum bid. However, after calculation of individual WTP, some portion of the respondents might have a WTP greater than their income that violates standard economic theory because no one can pay more than their income. To avoid this issue, the truncated mean is used as an alternative. As shown by Boyle et al. (1988[47]), approximating the upper limit of integration by the highest bid used in the survey is not a correct indication of the expected WTP and a normalisation procedure should be applied, as suggested in Boyle et al. (1988[47]).

Following the SWACHE core principles, the mean WTP is derived from post-estimation of individual WTP values that are i) truncated at the maximum bid, ii) truncated at the maximum bid with adjustment method of Boyle et al. (1988[47]), and iii) considering the median of the individual WTPs since the median is generally a more robust measure of central tendency (Hanemann, 1984[48]). Since the risk reduction is expressed as a unit per 1 000, the resulting WTP estimates is multiplied by 1 000 to get the estimate of the VSC. The central estimate of very low birth weight benefits corresponds to the mean VSC that is based on assumption of log-lognormal distribution of the errors in the WTP model, with exclusion of speeders and respondents who failed the probability quiz, and relying on the individual mean WTP truncated at the maximum bid with adjustment.

All monetary values are expressed in USD PPP.[10]

4.3.2. Spike configuration

Contingent valuation studies usually suffer from a large share of respondents who are not willing to pay anything. Still, these people may choose the reduced risk alternative when it costs nothing. Alternatively, some of them might choose the status quo otherwise even if it costs them nothing because, for instance, they do not care about very low birth weight risk, or the effects associated with very low birth weight are perceived not severe enough, or the proposed reduction is too low). These reasons of saying 'no' indicate

[10] The conversions are done using Purchasing Power Parities for actual individual consumption of 2019. Data are provided by the OECD. https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm. The following rates to express monetary values in USD PPS are: CZK 12.74, EUR 0.64 (Italy), MXN 9.90, EUR 0.76 (Netherlands), CHF 1.11, TRY 2.61, GBP 0.67, and CAD 1.29.

true zero options. However, these cases create a spike near zero in the WTP distribution, and if a spike is important and not included in the model, WTP may be overestimated (Carson and Hanemann, 2005[44]). In the contingent valuation studies, this issue has been typically treated by including a jump discontinuity in a probability density function (a spike) of selected parametric distribution. The econometric approach used here is based on censoring the parametric distribution (Kriström 1997) following the SWACHE core principles.

We measure the spike near zero by using the respondents' responses to the follow-up open-ended question followed the second valuation question ('no-no' response in the double-bounded dichotomous choice) by asking *"How much would* you *be willing to pay at maximum for these safer products to reduce the probability of your child having a very low birth weight by FILL in 1 000?"* This new information enriches the likelihood function by knowing that $0 < WTP < BID_2^L$ for those who responded 'no-no' and then stated some positive value in the open-ended follow up, while $WTP \leq 0$ for the persons who responded 'no-no' and stated zero to the follow-up. In this report, we further assume log-normal with Spike and Weibull with Spike for the distribution of the errors.

4.3.3. Other controls

The intercept α in the WTP equation can include the characteristics of the choice task and or other socio-demographic variables, risk attitudes or health status of a person that will allow us to see if WTP depends on them.

To find whether the risk preferences depend on the status quo risk that is the baseline incidence, two strategies are followed. First, the baseline risk is added to the intercept. Second, the model includes additively the interaction term between the risk reduction and the baseline risk, $\Delta RISK \cdot BASELINE$, that allows to investigate whether preference for the risk reduction differs across the baseline risks (used as the split-sample treatment).

Following the same strategy as in the case of the baseline risk, special attention is paid to other effects that respondents might consider when thinking about the payment for the safer products that would reduce the risk of very low birth weight. In several separate models, additional controls were included: a dummy equal to 1 for respondents who indicated they thought about other effects (OTHER), a dummy equal to 1 for respondents who said the other effects either increased her WTP (OTHERincr) or decrease her WTP (OTHERdecr), and a dummy equal to 1 if respondents indicated that the safer products would also reduce the probability of miscarriage or infant death (OTHERdeath).

Last, the effect of a lack of incentive compatibility on WTP and hence VSC is also examined. Lack of incentive compatibility is defined by a dummy (IC) that is equal to one if respondents definitely agreed or agreed (level 6 and more at the 7-point Likert scale) with the statement that although they chose the safer products, they wouldn't pay for it in a real situation.

4.4. Robustness checks

Robustness checks include exploring the influence of distributional choices on WTP (log-logistic, log-normal, Weibull, log-normal with Spike, and Weibull with Spike) and different ways of screening out respondents (excluding speeders, quiz failures, both of them, none of them). We also examine what is the effect of using the post-stratification weights on WTP compared to the results without weighting.

Country specific WTP estimates, and hence VSC of very low birth weight, are derived from the pooled baseline model and country-specific data models.

5 Results

5.1. Main results

The WTP models are estimated following eq. (1), assuming different distributions of the errors, including log-normal, log-logistic, Weibull, and spike configuration of log-Normal and Weibull. The preferred model is the WTP model with log-logistic distribution as it fits the data best, i.e. as the log-logistic model has the lowest Bayesian information criterion in comparison to all models with other distributions considered.

The proportion of 'no-no' responses on the double-bounded dichotomous choice questions is about 22% and this proportion varies quite a lot among countries; it is 11 and 13% in Mexico and Italy and 29–31% in Canada and Türkiye. The share of (true) zero WTP – as stated in the follow-up open-ended question – is however very small in all countries, between 0.2–0.3% (Türkiye, Mexico) and 1.8–2.1% (Switzerland, the UK, the Czech Republic), see Table 5.1. This is also the reason why WTP and hence VSC estimates do not differ much between the models assuming the same distribution with or without spike.[11]

Table 5.1. Proportion of 'no-no' responses and zero WTP

	All observations			Speeders & who failed the quiz excluded		
	yes-yes'	no-no'	spike WTP(open ended)=0	yes-yes'	no-no'	spike WTP(open ended)=0
Canada	39.0%	27.5%	1.4%	38.2%	29.3%	1.4%
Czech Republic	44.3%	20.1%	1.5%	41.9%	18.5%	1.8%
Italy	55.5%	15.1%	0.9%	56.1%	13.3%	0.8%
Mexico	52.9%	11.7%	0.3%	54.0%	11.1%	0.3%
Netherlands	40.3%	23.2%	1.1%	41.7%	24.5%	1.4%
Switzerland	39.9%	26.6%	1.7%	40.0%	25.1%	2.1%
Türkiye	54.8%	31.0%	0.3%	60.0%	23.0%	0.2%
United Kingdom	33.8%	30.7%	1.8%	33.3%	31.1%	1.9%
United States	48.7%	20.1%	0.6%	52.8%	19.6%	0.8%
All countries	45.5%	22.9%	1.1%	46.1%	21.7%	1.2%

The estimation results of the preferred double-bounded dichotomous choice model are presented in Table 5.2. Column (1) displays the baseline estimation results, providing the basis for deriving the central estimates of VSC of very low birth weight. The coefficient for $\Delta RISK$ is positive and statistically significant,

[11] In this report, the estimates based on Weibull distribution with "spike" configuration are mainly reported to compare the results from this study with the results from other studies conducted within SWACHE project. The spike is defined when a respondent responded 'no-no' and also was willing to pay nothing (zero) in the open-ended question, followed after the second binary choice question (Kriström, 1997[53]).

meaning that the higher the risk reduction, the higher the probability that respondents choose the safer products and hence the higher the WTP. The statistically significant and positive coefficient also provides evidence for scope sensitivity. The additional costs of the safer products that decrease the risks have a negative and significant effect on the probability that the safer products are chosen.

The results indicate a mean WTP of USD 597 per month for an average reduction of 4 in 1 000 in the risk of very low birth weight for a total period of 8 months. This corresponds to a mean VSC of very low birth weight of USD_{2022} PPP 1 194 000, which constitutes the central (baseline) estimate (see column (1) of Table 5.2). The median VSC equals USD_{2022} PPP 614 000. Typically, the WTP distribution is skewed and that is also the case in this study. It implies the median values of WTP are much smaller than the mean counterparts and that the median provides a more conservative estimate.

The VSC of very low birth weight results are robust to various data cleaning strategies and the use of weights to correct for difference between achieved quota and target quota. Column (2) in Table 5.2 shows the estimation results when the post-stratification weights (and their interactions with country dummies) are not included in the model. The mean and median VSC of very low birth weight is only about 1% higher than the baseline estimates.

Finally, the VSC of very low birth weight results are robust to various data cleaning strategies. Column (3) to (5) in Table 5.2 shows the results when only speeders are excluded, those who fail the probability quiz are excluded, and when no respondent is excluded from the model, respectively. Including those who failed increases the mean VSC by 7%, including speeders decreases the mean VSC by 4%, and including everyone increases the mean VSC by 2%. The data cleaning has a slightly larger effect on the median VSC, increasing it by 8% compared to the baseline estimate.

5.2. Country-specific estimates

The country-specific dummies in Table 5.2 indicate that respondents from Italy and Mexico are willing to pay the most, followed by respondents from the United States (that is the reference category) and Türkiye, whilst respondents from Canada and the United Kingdom are willing to pay the least. This tendency is supported in all models, assuming different WTP distributions, weighting, or screening the data. Country-specific estimates of VSC of very low birth weight rely on are the baseline model assuming log-logistic distribution of the errors and mean truncation at the maximum bid with adjustment. Country-specific mean VSC are in a range of USD_{2022} PPP 805 000 (United Kingdom) to USD_{2022} PPP 1 744 000 (Italy). The median VSCs are in a range of USD_{2022} PPP 338 000 (United Kingdom) to USD_{2022} PPP 1 050 000 (Italy), see Table 5.3.

Cross-country differences in VSC values are similar between the log-logistic model and Weibull with spike, except Türkiye where log-logistic model leads to lower VSC estimates than Weibull with “spike” due to relatively more positively skewed WTP distribution in Türkiye. In the majority of countries, however, the log-logistic model results in slightly larger VSC values than Weibull with “spike” for truncated means, whilst VSC is a bit larger for Weibull with spike if VSC is derived from individual median WTP estimates, see Figure C.1 and Table C.2-Table C.7 in Annex C. Overall, these differences are very small in magnitude.

It seems that weighting does not affect country-specific VSC estimates, see Figure C.3 in Annex C. Actually, the differences between weighted and unweighted VSC values are much smaller than differences due to assuming different distributions. Using the pooled data, the weighted VSC is only 0.14% larger than the unweighted VSC assuming a log-logistic distribution of the errors (see Table C.12).

Table 5.2. Estimates of WTP for reducing very low birth weight, log-logistic distribution

	Baseline		Without weights		Speeders excluded		Quiz failed excluded		None excluded	
	(1)		(2)		(3)		(4)		(5)	
Intercept	6.958	***	7.118	***	6.958	***	6.690	***	6.581	***
	(0.131)		(0.15)		(0.131)		(0.133)		(0.114)	
Czech Republic	-0.285	***	-0.310	**	-0.285	***	-0.164	.	-0.144	.
	(0.084)		(0.095)		(0.084)		(0.088)		(0.077)	
Italy	0.326	***	0.316	**	0.326	***	0.390	***	0.367	***
	(0.087)		(0.099)		(0.087)		(0.09)		(0.079)	
Mexico	0.139	.	0.143	.	0.139	.	0.267	**	0.228	**
	(0.087)		(0.1)		(0.087)		(0.091)		(0.078)	
Netherlands	-0.426	***	-0.423	***	-0.426	***	-0.257	**	-0.256	***
	(0.087)		(0.099)		(0.087)		(0.09)		(0.077)	
Switzerland	-0.256	**	-0.269	**	-0.256	**	-0.139	.	-0.160	*
	(0.086)		(0.098)		(0.086)		(0.088)		(0.077)	
Türkiye	0.048	.	-0.141	.	0.048	.	-0.481	***	-0.322	***
	(0.093)		(0.11)		(0.093)		(0.096)		(0.081)	
United Kingdom	-0.717	***	-0.750	***	-0.717	***	-0.563	***	-0.552	***
	(0.085)		(0.096)		(0.085)		(0.087)		(0.076)	
Canada	-0.553	***	-0.613	***	-0.553	***	-0.451	***	-0.418	***
	(0.086)		(0.097)		(0.086)		(0.088)		(0.077)	
$\Delta RISK$	0.055	***	0.063	***	0.055	***	0.055	***	0.044	***
	(0.012)		(0.013)		(0.012)		(0.012)		(0.01)	
log(bid)	-0.911	***	-0.939	***	-0.911	***	-0.900	***	-0.874	***
	(0.015)		(0.017)		(0.015)		(0.015)		(0.013)	
Distribution	Log-logistic		Log-logistic		Log-logistic		Log-logistic		Log-logistic	
Post-stratification weights interacted with country dummies	yes		no		no		no		no	
LL	-8 798.9		-8 775.28		-11 172.8		-10 540.7		-13 893.9	
AIC	17 621.8		17 572.6		22 367.5		21 103.5		27 809.7	
BIC	17 703.9		17 647.9		22 445.6		21 180.6		27 889.9	
Observations	6 940		6 940		8 910		8 188		10 800	
Excluded	speeders, quiz failed		speeders, quiz failed		speeders		quiz failed		none	
VSC of very low birth weight, USD2022 PPP[a]										
Mean VSC[b]	1 194 223		1 201 945		1 284 213		1 155 448		1 223 771	
Mean VSC[c]	894 810		896 453		933 950		869 922		903 698	
Median VSC	613 816		622 760		673 745		571 134		609 182	

Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; standard errors are given in parentheses. [a] For clarity, VSC are expressed in USD2022 because the field implementation of the surveys was completed in 2022. [b] Mean derived from individual mean WTP truncated at the maximum bid with adjustment; [c] mean derived from individual mean WTP truncated at the maximum bid.

Using the pooled baseline model, country-level estimates of the mean (median) WTP are the average of the individual mean (median) WTP estimate for all observations from that country. Following this approach implies that country-level WTP estimates are derived from a single model and hence using the same estimated coefficients. The pooled model also includes country dummies interacted with the post-stratification weights which corrects for the under-/over- representation of individuals in the sample as well as captures country-specific, cultural, differences among the countries. Because using the pooled data also increases the statistical power of the country-level estimates this approach is expected to aid in transferring WTP estimates to countries other than those surveyed" (Dockins et al., 2023[49])). Following the SWACHE Core principles, the results reported in Table 5.3 are used as the recommended values for each country (based on the model in Table 5.2, column 1).

Table 5.3. Country-specific VSC of very low birth weight, pooled baseline model, thousands USD_{2022} PPP

	Czech Republic	Canada	Italy	Mexico	Netherlands	Switzerland	Türkiye	United Kingdom	United States
Mean VSC	1 120	872	1 744	1 528	1 013	1 107	1 274	805	1 389
Median VSC	546	384	1 050	869	476	549	658	338	753

Note: Based on the baseline model presented in column (1) of Table 5.2. For clarity, VSC are expressed in USD_{2022} because the field implementation of the surveys was completed in 2022.

The estimation results of modelling each country separately are presented in Table C.8. For all countries, the coefficients have signs that are consistent with the results from the baseline model. The coefficients for the risk reduction are however positive and statistically significant in four countries only due to the small sizes of the country samples. The mean (median) VSC varies from USD_{2022} PPP 735 000 (USD_{2022} PPP 330 000) in the United Kingdom to USD_{2022} PPP 1 679 000 in Italy (USD_{2022} PPP 1 074 000 for Türkiye). The mean VSC is about 15% and 17% lower for Mexico and Czech Republic and 15% higher for Türkiye when WTP is estimated from the country-specific models. The mean VSC resulting from these two different estimation strategies are highly similar (see Figure C.4).

5.3. Baseline risk

Two models are estimated to investigate the effect of the baseline level of the risk that varied across respondents at random, with a magnitude of the prevalence 6, 10, and 12 in 1000. If the effect of baseline risk on WTP is significant one would need to consider this fact in a benefit transfer to a country with lower or higher prevalence of very low birth weight. Overall, the effect of baseline risk on WTP for reducing risk of very low birth weight is not statistically significant. Therefore, it is not necessary to transfer WTP values and VSC for very low birth weight by baseline risk level.

In the first modelling strategy, baseline risk enters the model as a continuous variable (column (1)) or as two dummies equal to one if baseline risk level was either 10 or 12 (column (2)). These results are reported in Table 5.4. In these two specifications the baseline risk levels are not statistically significant at any confidence level.

Table 5.4. Controlling for baseline risk

	(1)		(2)	
Intercept	6.784	***	6.743	***
	(0.17)		(0.144)	
Canada	-0.178	**	-0.177	**
	(0.068)		(0.068)	
Czech Republic	0.001		0.002	
	(0.08)		(0.08)	
Italy	0.599	***	0.599	***
	(0.086)		(0.086)	
Mexico	0.409	***	0.409	***
	(0.085)		(0.085)	
Netherlands	-0.065		-0.062	
	(0.074)		(0.074)	
Switzerland	0.045		0.043	
	(0.08)		(0.08)	
Türkiye	0.053		0.052	
	(0.089)		(0.089)	
United Kingdom	-0.305	***	-0.305	***
	(0.07)		(0.07)	
United States	0.282	***	0.283	***
	(0.084)		(0.084)	
$\Delta RISK$	0.064	***	0.064	***
	(0.014)		(0.014)	
log(bid)	-0.936	***	-0.937	***
	(0.017)		(0.017)	
baseline risk	-0.002			
	(0.012)			
baseline 10:1000			0.068	
			(0.069)	
baseline 12:1000			-0.004	
			(0.07)	
Distribution	Log-logistic		Log-logistic	
Post-strata weights	yes		yes	
LL	-8 798.88		-8 797.77	
AIC	17 623.77		17 623.54	
BIC	17 712.75		17 719.37	
No	6 940		6 940	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; standard errors are given in parentheses.
Log-logistic distribution, weighted, speeders and who failed the probability quiz excluded.

As an alternative modelling strategy, the baseline risk level enters the model via interaction with the risk reduction, expecting that respondents' preferences for reducing the risk depend on the baseline level. The rationale of this approach is that the more likely a baby is born with very low birth weight, the more likely a person is accepting a program that reduces given risk. If this is the case, then the coefficients on the three interaction terms will be statistically different from one another. The results from this model are reported in Table C.9. Column (A) reports the results if responses on all choice questions are used, that means the risk reduction by 2, 3, 5, 6, and 7 in 1 000. It is found that the preference for the reduced risk option differs between baseline risk level 10 and 12 ($\chi^2 = 4.131$, $p = 0.042$). However, due to the survey design, larger levels of risk reduction can be applied only to larger baseline risk levels. Consequently, the effects of the baseline risk levels and of the risk reduction levels cannot be disentangled from one another.

Therefore, the model was re-estimated using only responses from people who faced small changes in risk reduction that are 2 and 3 in 1000 that were attributed to all baseline risk levels. These results are reported in column (A23) in Table C.9. None of the three coefficients is statistically different from zero or statistically different from one another. A similar conclusion is drawn when a normal distribution is assumed for the error term.

5.4. Other effects and incentive compatibility

The impacts of considering other effects from the use of safer products aside from the decrease of the probability of very low birth weight on WTP are analysed in Table 5.5. Column (1) to (3) reports the results for the data cleaning, keeping only respondents who did not consider other effects in column (1), who did not think that other effects increased or decreased their WTP in column (2), and who did not think the other effects increased their WTP in column (3). The next two columns of Table 5.5 report the results when respondents lacking incentive compatibility are excluded from the sample, in column (4), together with those who considered other effects, column (5). The last three columns report results from models that use the data from respondents who considered other effects, column (6), who thought they increased their WTP, column (7), and who thought the safer products may also reduce the probability of miscarriage or infant death, column (8).

WTP for those who did not consider other effects is much smaller than WTP for those who considered these effects. Mean VSC is in a range of USD_{2022} PPP 1 008 000 – 1 029 000 for the former, and it is USD_{2022} PPP 1 450 000 – 1 661 000 for the latter group. The mean and median VSC from model (1), only relying on responses that did not consider other effects, are equal to USD_{2022} PPP 1 008 000 and USD_{2022} PPP 499 000 respectively. These values are 16% and 19% smaller than the baseline VSC estimates and could be used as a conservative value of the estimates.

After excluding those who would not pay in a real situation, the mean (median) VSC is USD_{2022} PPP 1 165 000 (577 000), see column (4), compared to USD_{2022} PPP 1 194 000 (614 000) from the baseline model. Excluding incentive not compatible responses thus reduces the mean (median) WTP by 2.4% (6.1%). If responses that took into account other effects are also excluded in column (5), the mean (median) VSC is USD_{2022} PPP 992 000 (478 000), that is 17% (22%) lower than the baseline estimates. However, this stringent exclusion strategy results in a quite small sample that contains only 3 458 valid observations.

The estimation results from an alternative model that includes the interaction terms between other effects and the risk reduction are presented in Table C.10 in Annex C. The coefficient for the interaction term with increased effect on WTP is positive and significant, whilst the effect of decreased effect on WTP is negative but not statistically significant at any convenient level, see column (2). Looking at specific effects in column (1), the positive effect of the change in the risk reduction on the probability to choose the reduced risk option is statistically larger for people who thought that the safer products would reduce the probability of infant death and/or improve the health of other family members.

Table 5.5. WTP with other effects and incentive compatibility

	other effects not considered		if other effects increased or decreased WTP excluded		if other effects increased WTP excluded		Incentive compatible		Incentive compatible & who did not consider other effects		who considered other effects		other effects increased their WTP		reduction of miscarriage and infant death	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Intercept	6.949	***	6.809	***	6.624	***	6.407	***	6.728	***	6.680	***	7.445	***	7.223	***
	(0.193)		(0.178)		(0)		(0.146)		(0)		(0.208)		(0.274)		(0.26)	
Canada	-0.259	**	-0.310	***	-0.264	**	-0.205	**	-0.297	**	-0.156	.	-0.023		-0.267	*
	(0.101)		(0.089)		(0.083)		(0.072)		(0.105)		(0.094)		(0.123)		(0.108)	
Czech Republic	-0.093		-0.064		-0.005		0.023		-0.101		0.267	*	0.249		0.237	
	(0.103)		(0.096)		(0.093)		(0.083)		(0.107)		(0.132)		(0.165)		(0.152)	
Italy	0.631	***	0.555	***	0.605	***	0.600	***	0.594	***	0.557	***	0.744	***	0.685	***
	(0.117)		(0.107)		(0.101)		(0.09)		(0.12)		(0.129)		(0.171)		(0.164)	
Mexico	0.339	**	0.436	***	0.497	***	0.411	***	0.300	*	0.387	**	0.213		0.301	*
	(0.122)		(0.111)		(0.103)		(0.094)		(0.131)		(0.119)		(0.152)		(0.141)	
Netherlands	-0.084		-0.066		-0.021		-0.098		-0.133		0.050		0.210		-0.053	
	(0.095)		(0.089)		(0.083)		(0.078)		(0.097)		(0.122)		(0.176)		(0.142)	
Switzerland	0.127		0.144		0.200	*	0.040		0.082		-0.003		-0.207		-0.148	
	(0.104)		(0.097)		(0.093)		(0.082)		(0.106)		(0.13)		(0.166)		(0.155)	
Türkiye	0.129		0.043		-0.287	**	-0.039		0.006		-0.077		0.472	**	0.283	.
	(0.138)		(0.126)		(0.111)		(0.095)		(0.15)		(0.118)		(0.163)		(0.15)	
United Kingdom	-0.347	***	-0.402	***	-0.322	***	-0.346	***	-0.380	***	-0.252	*	-0.177		-0.260	*
	(0.097)		(0.089)		(0.084)		(0.074)		(0.101)		(0.102)		(0.132)		(0.12)	
United States	0.354	**	0.308	**	0.356	***	0.239	**	0.382	**	0.160		0.079		0.192	
	(0.119)		(0.11)		(0.105)		(0.09)		(0.125)		(0.119)		(0.142)		(0.141)	
ΔRisk	0.072	***	0.069	***	0.067	***	0.066	***	0.074	***	0.054	**	0.064	*	0.036	
	(0.018)		(0.016)		(0.015)		(0.014)		(0.018)		(0.019)		(0.025)		(0.023)	

log(bid)	-0.989 ***	-0.968 ***	-0.947 ***	-0.895 ***	-0.961 ***	-0.892 ***	-0.975 ***	-0.939 ***
	(0.023)	(0.021)	(0.02)	(0.017)	(0.024)	(0.024)	(0.032)	(0.03)
Distribution	Log-logistic	Log-logistic	Log-logistic	Log-logistic	Log-logistic	Log-logistic	Log-logistic	Log-logistic
Post-strata weights	yes	yes	yes	yes	yes	yes	yes	yes
LL	-4 816.8	-5 593.1	-6 183.6	-7 846.7	-4 458.5	-3 935.5	-2 504.9	-2 711.9
AIC	9 657.6	11 210.2	12 391.2	15 717.5	8 940.9	7 895.0	5 033.7	5 447.8
BIC	9 732.3	11 286.7	12 468.8	15 798.3	9 014.7	7 967.9	5 101.9	5 516.7
Observations	3 723	4 334	4 781	6 207	3 458	3 217	2 159	2 296
Excluded	speeders, quiz failed	speeders, quiz failed	speeders, quiz failed	speeders, quiz failed	speeders, quiz failed	speeders, quiz failed	speeders, quiz failed	speeders, quiz failed
VSC of very low birth weight, thousand USD PPP								
Mean VSC	1 008	1 015	1 029	1 165	992	1 450	1 661	1 572
Mean VSC*	800	802	805	874	788	1 008	1 099	1 059
Median VSC	499	497	500	577	478	797	1 001	917
Difference from the baseline estimate								
Mean VSC	-16%	-15%	-14%	-2%	-17%	21%	39%	32%
Mean VSC*	-11%	-10%	-10%	-2%	-12%	13%	23%	18%
Median VSC	-19%	-19%	-19%	-6%	-22%	30%	63%	49%

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses. Log-logistic distribution, with post-stratification weights, speeders and who failed the probability quiz excluded. Incentive incompatible are respondents who choose the safer product but also said that they wouldn't pay in real situation.

5.5. Sociodemographic characteristics and health

The effects of sociodemographic characteristics, child planning*,* purchasing behaviour and health status on WTP are reported in Table 5.6. The model in column (A) includes regressor on socio-demographic variables and some on child planning. Purchasing behaviour is added in column (B), and health variables and length of partner relationship are added to the model reported in column (C). All regressors, except employment status, are included in the last model, reported in column (D). Even when these characteristics are controlled for, the change in the risk reduction of very low birth weight and the bid level stay highly statistically significant with coefficients of similar size as the baseline estimates.

Further, in all models in column (A-D), it is found that WTP for the safer products increase with household monthly income, and the implied income elasticity is 0.30-0.34, depending on the model specification. WTP is also higher for people who chose not to disclose information on their income. People living in urban areas (settlement with a high population density and infrastructure of built environment) are more willing to pay for the safer products than people from rural areas. University educated respondents are more likely to choose the safer products than people with lower secondary education and less educated. Employed or part time employed respondents do not differ in their preferences from students, unemployed, long-term sick or disabled, homemakers or retired respondents.

The more biological children respondents think they will have *(morekid1,_2,_3)* the more likely they are to choose the safer products compared to those who have not thought about it or do not know. This tendency reaches its maximum with people who intend to have three children. People who think they will have four or more children *(morekid4)* are not more likely to pay for the safer products. When people would like to have a child *(when_Xyrs)* or whether they have already a child *(kids1, kids2)* do not have a significant effect on the decision to pay for the safer products.

Married respondents are slightly less likely to choose the safer products than those who have a partner but are not married. Respondents identifying as female tend to be willing to pay less than respondents identifying as male. Respondents aged 25-34 are less likely to pay for the safer products than respondents below 24 years. When the couple has been in relationship five to six years (*five-six-years* dummy), is the respondent is more willing to pay for the safer products than if they are in a couple in a longer relationship (reference category).

Respondents from households that usually or every time (90% to 100% of the purchases) have bought toiletries *(fre_toiletries)* and furniture *(fre_furniture)* that contain fewer toxic substances than conventional products during the last year are willing to pay more for the safer products. The association is less strong, but still positive and significant at 10% significance level, in the case of households that have bought usually or every time cleaning products *(fre_cleaningp)* and clothing *(fre_clothing)* that contains fewer toxic substances. This indicates that responses to the contingent valuation questions were consistent with the past purchase behaviour of respondents. Thus, their responses to the very low birth weight survey represent their preferences quite well. Since 'safer' household purchase is measured by four dummies, the more respondents bought products containing fewer toxic substances, the more likely they are to buy the new safer products that will reduce the probability of very low birth weight. Overall, this suggests that respondents were considering cleaning products and clothing less than toiletries and furniture when responding to the valuation questions.

People who describe their health as well below or below average relative to others of their age and gender *(health_below)* are less willing to buy the safer products than those who perceive their health as average (reference category). The opposite holds for those who evaluate their health as well above or above average *(health_above)* compared to respondents who perceive it as average. Experience with being born with low birth weight *(LBW)* or having a child born with low birth weight *(LBWkid)* has no statistical effect on WTP to reduce the risk of very low birth weight. People who had Covid-19 *(had_COVID)* are also less willing to pay for the safer products than people who did not have it. There is also no significant association between coverage of medical costs from respondents' budget (*own_pocket*) and preference for the safer products.

Table 5.6. WTP model with sociodemographic and health characteristics

	(A)		(B)		(C)		(D)	
ΔRISK	0.066	***	0.067	***	0.067	***	0.069	***
	(0.013)		(0.013)		(0.013)		(0.013)	
female	-0.240	***	-0.180	**	-0.183	***	-0.178	***
	(0.052)		(0.055)		(0.053)		(0.053)	
AGE_25-34	-0.248	**	-0.247	**	-0.235	**	-0.222	**
	(0.083)		(0.083)		(0.084)		(0.084)	
AGE_35-39	-0.113		-0.110		-0.054		-0.045	
	(0.085)		(0.086)		(0.088)		(0.088)	
AGE_40-44	0.002		0.009		0.087		0.094	
	(0.114)		(0.114)		(0.116)		(0.116)	
AGE_45+	-0.158		-0.132		-0.071		-0.040	
	(0.175)		(0.177)		(0.177)		(0.178)	
ln(household income)	0.338	***	0.308	***	0.316	***	0.301	***
	(0.041)		(0.042)		(0.041)		(0.042)	
info about income missing	6.097	***	5.588	***	5.738	***	5.485	***
	(0.755)		(0.781)		(0.763)		(0.767)	
university	0.234	**	0.209	**	0.174	*	0.173	*
	(0.078)		(0.079)		(0.079)		(0.079)	
upper secondary	(0.157)	.	0.147	.	0.122		0.124	
	(0.080)		(0.081)		(0.081)		(0.081)	
married	-0.121	*	-0.118	*	-0.097	.	-0.100	.
	(0.054)		(0.054)		(0.056)		(0.056)	
when_2yrs	0.016		0.027					
	(0.060)		(0.060)					
when_3yrs	0.048		0.062					
	(0.071)		(0.071)					
when_4yrs	-0.071		-0.058					
	(0.100)		(0.100)					
when_5yrs	-0.014		-0.004					
	(0.090)		(0.091)					
kids1	0.019		-0.022		0.017		-0.017	
	(0.057)		(0.057)		(0.060)		(0.060)	
kids2	0.047		-0.021		0.083		0.030	
	(0.065)		(0.066)		(0.070)		(0.071)	
urban			0.283	***	0.305	***	0.281	***
			(0.068)		(0.068)		(0.068)	
suburban			0.007		0.009		0.014	
			(0.074)		(0.074)		(0.074)	
own pocket					0.086		0.086	
					(0.072)		(0.072)	
morekid1					0.138	*	0.137	*
					(0.062)		(0.062)	
morekids2					0.248	***	0.248	***
					(0.065)		(0.065)	

Table continues…	(A)		(B)		(C)		(D)	
morekids3					0.393 (0.106)	***	0.376 (0.107)	***
morekids4					0.036 (0.155)		-0.081 (0.156)	
had COVID					-0.120 (0.051)	*	-0.122 (0.051)	*
health below average					-0.280 (0.079)	***	-0.292 (0.079)	***
health above average					0.104 (0.050)	*	0.061 (0.051)	
LBW					0.018 (0.082)		-0.017 (0.082)	
LBWkid					-0.073 (0.096)		-0.124 (0.096)	
less1year					0.071 (0.079)		0.025 (0.079)	
one-two-years					-0.031 (0.087)		-0.022 (0.087)	
three-four-years					0.116 (0.072)		0.107 (0.072)	
five-six-years					0.192 (0.071)	**	0.184 (0.071)	**
fre_clothing			0.138 (0.082)	.			0.141 (0.082)	.
fre_toiletries			0.244 (0.071)	***			0.242 (0.071)	***
fre_furniture			0.244 (0.090)	**			0.240 (0.090)	**
fre_cleaningp			0.142 (0.069)	*			0.135 (0.069)	.
employed			0.078 (0.067)					
part-time employed			0.061 (0.084)					
exposed							0.079 (0.048)	.
log(bid)	-0.950 (0.017)	***	-0.961 (0.017)	***	-0.960 (0.017)	***	-0.966 (0.017)	***
Distribution	Log-logistic		Log-logistic		Log-logistic		Log-logistic	
Post-strata weights	Yes		Yes		Yes		Yes	
LL	-8 727.39		-8 670.65		-8 675.14		-8 640.07	
AIC	17 510.78		17 413.30		17 430.28		17 370.14	
BIC	17 702.44		17 659.72		17 704.08		17 678.16	
Observations	6 940		6 940		6 940		6 940	

Notes : Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1, standard errors in parentheses. Speeders and respondents who failed the probability quiz excluded. Intercept and country dummies are included in the model but not displayed in this table for clarity.

5.6. Additional robustness checks

The VSC estimates are also robust to various modelling choices (error distribution) and to various methods of screening out various respondents. However, the method used to compute the mean and median VSC has a non-negligible impact on the magnitude of the VSC.

First, the mean and median VSC were estimated assuming five different distributions of the errors, including log-logistic, log-normal, Weibull, and log-normal and Weibull with spike. Table C.2 in Annex C reports the estimation results for the different distributions. The mean VSC is USD_{2022} PPP 1 186 000 and the median VSC is USD_{2022} PPP 666 000 for Weibull with spike (see column (5)) that is also the preferred model in other studies conducted in the SWACHE project. Both these statistical moments are very close to the baseline estimate, deviating only by 1%, and 7%, respectively. Column (1) to (3) reports the results for the log-logistic, log-normal and Weibull distributions without spike. Column (4) shows the results for the lognormal distribution with spike that present the largest deviation from the baseline estimate in terms of mean VSC.

Second, the different ways of screening respondents – excluding speeders, who failed the probability quiz, both of them, and none of them – does not affect the statistical significance of the risk coefficients, nor the WTP and VSC estimates, see Table C.3 to Table C.7 in Annex C. The screening strategy influences slightly the size of significance of some of the country dummies. However, all these variables stay statistically significant at least at the 5% level. The WTP estimates are highest when only speeders are excluded and lowest when only those who failed the probability quiz are excluded. The WTP estimates based on the complete dataset (including speeders and failed quiz) are similar to estimates based on the subsample where speeders and probability quiz failures are not included. Looking only at the preferred log-logistic model, the estimates of individual mean WTP (truncated at the maximum bid with adjustment) ranges from USD_{2022} PPP 1 155 000 to 1 224 000 while the median VSC ranges from USD_{2022} PPP 571 000 to 674 000, depending on the screening out approach.

Unsurprisingly, the method used to compute the mean and median WTP has a significant impact on the magnitude of the WTP statistics. Table C.12 shows the resulting mean and median WTP when integrating the probability of an individual to agree with the safer products over the interval from zero to the maximum bid, with and without adjustment to derive WTP statistics at the respondent level. Median WTPs are systematically lower than mean WTPs while using the adjustment leads to larger mean WTPs. Table C.12 also shows the resulting mean and median WTP when using respondent level mean or median WTP to aggregate the results at the sample level. This has a very small impacts on the obtained magnitudes. The results are declined for the Weibull and log-logistic distribution and for the use of post-stratification weights.

Finally, it is worth comparing the baseline VSC estimate derived from the DBDC to the VSC estimate derived from the single-bounded dichotomous choice (SBDC). Table C.11 report the results of the logit model estimation using only the response from the first dichotomous choice. Panel A of Table C.11 shows the coefficient and the statistical significance while Panel B shows the mean and median VSC by country. The mean VSC of very low birth weight for the pooled sample based on the SBDC is equal to USD2022 PPP 613 000 that is twice as small as the mean VSC of very low birth weight obtained based on the DBDC equal to USD2022 PPP 1 194 000. The median VSC based on the SBDC equals USD2022 PPP 531 000 is closed to the median VSC based on the DBDC that is equal to USD2022 PPP 614 000. Country-specific VSC means range from USD2022 PPP 234 000 (United Kingdom) to USD2022 PPP 929 000 (Italy). VSC based on the SBDC will be compared to the results of previous studies using a similar approach in Section 6.2.

6 Recommended policy values

6.1. Central value across countries and by country

The model reported in column (1) of Table 5.2 was identified as the preferred model to provide the recommended value of a statistical case of very low birth weight. The baseline specification corresponds to a maximum likelihood estimation of the joint probabilities assuming a log-logistic distribution. The baseline model excluded speeders and respondents who failed the probability quiz and included the post-stratification weights entered the model through the interaction terms with country dummies.

The recommended value of the VSC of very low birth weight using the pooled data is based on the mean of individual WTP truncated at the maximum bid with adjustment and equals USD_{2022} PPP 1 194 000. The median VSC of very low birth weight is USD_{2022} PPP 614 000. Mean VSC of very low birth weight by country ranges from USD_{2022} PPP 805 000 (United Kingdom) to USD_{2022} PPP 1 744 000 (Italy).

Table 6.1 Values per statistical case of very low birth weight

	Mean VSC (in USD_{2022})	Mean VSC (in local currency)
Pooled across countries	1 194 000	USD 1 194 000
Canada	872 000	CAD 1 117 000
Czech Republic	1 120 000	CZK 13 535 000
Italy	1 744 000	EUR 1 220 000
Mexico	1 528 000	MXN 14 448 000
Netherlands	1 013 000	EUR 832 000
Switzerland	1 107 000	CHF 1 464 000
Türkiye	1 274 000	TRY 2 262 000
United Kingdom	805 000	GBP 602 000
United States	1 389 000	USD 1 389 000

Note: The conversions are done using Purchasing Power Parities for actual individual consumption of 2019 since this figure was also used to convert bid levels across countries. Data are provided by the OECD. https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm. For clarity and simplicity, VSC are expressed in USD_{2022} because the field implementation of the surveys was completed in 2022.

6.2. Comparison with previous studies

In Table B.2 the estimates obtained in this study are compared to the existing estimates from the two previous studies commissioned by ECHA (Ščasný and Zvěřinová, 2014[36]) and Health Canada (Ščasný and Zvěřinová, 2016[37]). To allow for a proper comparison, all values from the previous studies are adjusted to GDP per capita growth and inflation and expressed in USD_{2022} PPP using equation (3). Moreover, since the two previous studies rely on the SBDC elicitation format, estimates of the present studies derived using the SBDC are also presented.

The VSC of very low birth weight obtained in the present study are considerably larger than the estimates obtained in the two previous studies, even when considering only SBDC estimates. This difference might be due to several things. First, the payment vehicle used in the private scenarios differs. It might be the case that a chemical policy that affects the attributes of products was deemed more acceptable to respondents than a novel complex of vitamins and minerals used in the previous studies. Second, the SWACHE surveys were carried out six and eight years after the previous studies and transferring values over time only based on GDP per capita and inflation might generate benefit transfer errors.

6.3. Strengths and weaknesses of results

This study provides useful and internationally validated estimates of the VSC of very low birth weight for several OECD countries using an original, state-of-the-art stated preference survey. The survey was administered electronically to samples selected to be demographically representative of each country's population. Using various validity and robustness checks, the survey performs well and as intended. The coefficients have signs that are consistent with expectations and scope sensitivity is statistically significant. The cost for the reduced risk option has a negative effect on the probability to choose the reduced risk option that is statistically different from 0. The statistically significant determinants of WTP include the size of the risk reduction, gender, income, education and behaviour linked to the purchase of safer products.

In this study, the WTP is elicited in the context of chemicals management options in which respondents are asked if they would be willing to pay an extra amount of money to purchase safer products in order to reduce the risk of very low birth weight in their future child. It is known from the stated preference literature that the context influence the resulting WTP estimates. Using a private good scenario has several advantages such as avoiding potential double counting due to a paternalistic altruism that can appear in public good scenarios. However, relying on the private good scenario restricts the population to couples who wish and plan to have a child or another child. Therefore, the study does not provide any information about preferences of people who do not wish to have a baby or who are already beyond reproductive age.

While respondents were presented health benefits of using the safer products only in terms of reduced risk of very low birth weight, some of them stated that they were also thinking about other effects, potentially increasing stated WTP. Consequently, the baseline estimates of WTP and VSC of very low birth weight may implicitly include values for these other effects, notably a reduced probability of miscarriage and infant mortality as well as a reduction in other health impacts. These issues are analysed thoroughly in Section 5.4 of the paper and the implications for the use of WTP and VSC estimates in cost benefit analyses are discussed in Section 6.4.

Although the samples are close to the target quotas for gender for each of the countries surveyed, several samples missed the quotas for other key demographics – most notably the samples from Canada, the United Kingdom and the United States where respondents with low education are underrepresented and the samples from the Czech Republic, Netherlands and Switzerland where respondents aged 18-24 are underrepresented. However, using post-stratification weights as additional regressors allows to control for these deviations from the population structure.

While the study vastly expands the WTP estimates for infertility available for policy analysis, they were obtained in nine countries. Non-surveyed countries without country-specific values will need to conduct benefit transfer using best practices. In the absence of benefit transfer guidance specific to the health effects covered by the SWACHE project, it is recommended as a starting point that non-surveyed countries use the value estimated for a surveyed country from Table 6.1 that shares similar characteristics such as income, population by age, and public health care systems.

It is noted that the SWACHE survey was conducted during the COVID-19 pandemic that resulted in serious adverse respiratory effect in a significant number of contaminated people, including premature death. However, the implementation of surveys was not conducted during the peaks of the pandemics following the recommendations of Mourato and Schreedar (2021[50]) to avoid biasing the results. Moreover, the results show that people who had covid are not willing to pay more for reducing the risk of very low birth weight.

6.4. Using these recommended values in policy analysis

The obtained VSC estimates for very low birth weight can used in cost-benefit analyses addressing proposed regulations of chemicals or other pollutants that are linked to very low birth weight. Presented here is the recommended use.

Assume a policy is appraised over T years in country c. Compared to the status quo, this policy is estimated to lead to a reduction of SC_{ct} statistical cases of very low birth weight in country c in year t. The discounted benefits of the policy in terms of avoided very low birth weight should thus be computed as follows:

$$Discounted\ benefits_c = \sum_{t=0}^{T} \frac{VSC_{ct} \times SC_{ct}}{(1+k_c)^t} \tag{2}$$

where k_c is the discount rate used in country c[12], VSC_{ct} is the recommended value of a statistical case of very low birth weight in country c in year t. VSC_{ct} is based on the recommended values $VSCI_{c,2022}$ reported in USD PPP in Table 6.1 and should reflect increase in prices and in GDP per capita over time such that:

$$VSC_{ct} = VSC_{c,2022} \times PPP_{c,2019} \times \left(1 + \%\Delta P_{c,2022-t}\right) \times \left(1 + \%\Delta Y_{c,2022-t}\right)^{\beta} \tag{3}$$

where $PPP_{c,2019}$ stands Purchasing Power Parity for actual individual consumption in national currency per USD for the year 2019 that was used to convert the bid levels in the survey, $\%\Delta P_{c,2022-t}$ is the increase in consumer price index from 2022 to year t, $\%\Delta Y_{c,2022-t}$ is GDP per capita growth from 2022 to year t and β is the income elasticity.

An example for a fictional policy that reduces the number of statistical cases of very low birth weight by 1 000 every year in Czech Republic for 2022-2025 is provided in

Table 6.2 for illustrative purpose. Based on a VSC of USD 1 120 000 in 2022, the discounted benefits of the policy over the 4 years equals CZK2022 54 520 million.

Finally, the discounted costs of the policy should be subtracted from these discounted benefits to compute the net present value of the policy.

Table 6.2. Measuring the benefits of policy intervention in Czech Republic: an illustrative example using the value of a statistical case of very low birth weight

		2022	2023	2024	2025
GDP per capita, volume in USD, at constant PPP (USD_{2015})		36 933	37 131	38 186	40 155[a]
GDP per capita growth since 2022 ($\%\Delta Y_{c,2022-t}$)			0.53%	3.39%	8.72%[a]
Consumer Price Index (2015)		134	139[b]	144[b]	149[b]
Consumer Price Index growth since 2022 ($\%\Delta P_{c,2022-t}$)			5%[b]	10%[b]	15%[b]
PPP for actual individual consumption ($PPP_{c,2019}$)		12.09			
Value of a statistical case of very low birth weight (VSC)	(USD_{2022} PPP thousand)	1 120			
	(CZK_{2022} thousand)	13 535			
	(CZK thousand)	13 535	14 235	15 039	15 961
Annual statistical cases of very low birth weight avoided (SC_{ct})		1 000	1 000	1 000	1 000

[12] Note that the discount rate may vary over time, but is generally stable over shorter horizons.

Discounted annual benefits (CZK2022 million)	13 535	13 557	13 640	13 788
Discount rate	5%[c]	5%[c]	5%[c]	5%[c]
Discounted benefits (CZK2022 million)	54 520			

Note: This illustrative example assumes a fictional policy that would reduce the number of statistical cases of very low birth weight by 1 000 every year in Czech Republic from 2022 to 2025. GDP per capita projections for 2022-2024 are provided by the OECD Economic Outlook (2022[51]). [a]GDP per capita for 2025 is computed by the authors based on the linear fit of 2022-2024 values over time and is not an OECD forecast. Consumer Price Index data for 2022 comes from the OECD Dataset: Consumer price indices (CPIs) as of January 2022. [b]a 5% increase per year is assumed for Consumer Price Index for 2023-2025 and is not an OECD forecast. PPP for actual individual consumption data is for year 2019 as used to convert bid levels across countries and comes from the OECD Dataset: PPPs and exchange rates as of January 2022. [c]The discount rate is based on the EU recommendation for CEE countries,. The income elasticity equals 0.3 as estimated in this paper.

Incidence and prevalence of very low birth weight vary across countries and regions and will most likely change over time. In order to reflect these changes, this study elicited preferences for reducing the probability of very low birth weight for three different status quo prevalence levels (i.e. the baseline risk levels). These baseline risk levels are 6, 10, and 12 in 1 000 and this range of prevalence covers the prevalence levels that are typical for most countries worldwide. Despite the prior expectation that the baseline risk would influence the stated WTP, it was found that it does not affect WTP and hence VSC values. Therefore, the baseline estimates of VSC of very low birth weight can be applied in a benefit transfer analysis regardless of what the baseline risk level is, at least in the range 6 to 12 in 1 000.

The way other health effects are monetised in benefit-cost analyses influence the estimates of VSC of very low birth weight to use. If any other effects, positive or negative, aside from the decrease in the probability of very low birth weight are removed, the mean (median) VSC of very low birth weight is USD2022 PPP 1 008 000 (499 000) (see Panel A of Table 6.3). If responses that were identified as incentive incompatible are also excluded, the mean (median) VSC of very low birth weight become USD2022 PPP 992 000 (478 000) (see Panel B of Table 6.3). These conservative estimates of VSC of very low birth weight should be used in a cost-benefit analysis that also includes VSC of miscarriage and infant mortality because these adverse health risks were often in the minds of respondents to the present survey. However, if a cost-benefit analysis does not include these health risk or even any other health risk on new-borns or infants, then it is recommended to use the baseline estimates presented in Table 6.1, including potentially a value for other effects, aside from the decrease in the probability of very low birth weight.

Table 6.3. Country-specific VSC of very low birth weight, conservative estimates, means

Panel (A) – Responses without other effects

	USD PPP	Country currency	Difference from central estimate
Canada	737 000	CAD 948 000	-15%
Czech Republic	869 000	CZK 11 069 000	-22%
Italy	1 490 000	EUR 960 000	-15%
Mexico	1 198 000	MXN 11 865 000	-22%
Netherlands	876 000	EUR 664 000	-14%
Switzerland	993 000	CHF 1 102 000	-10%
Türkiye	1 024 000	TRY 2 669 000	-20%
United Kingdom	696 000	GBP 465 000	-14%
United States	1 272 000	USD 1 272 000	-8%

Panel (B) – Responses without other effects and only incentive compatible

	USD PPP	Country currency	Difference from central estimate
Canada	723 000	CAD 930 000	-17%
Czech Republic	870 000	CZK 11 082 000	-22%
Italy	1 463 000	EUR 943 000	-16%
Mexico	1 166 000	MXN 11 549 000	-24%
Netherlands	844 000	EUR 640 000	-17%
Switzerland	965 000	CHF 1 070 000	-13%
Türkiye	949 000	TRY 2 473 000	-26%
United Kingdom	683 000	GBP 456 000	-15%
United States	1 327 000	USD 1 327 000	-4%

7 Conclusion

This study provides estimates of VSC of very low birth weight for nine OECD countries: Canada, the Czech Republic, Italy, Mexico, the Netherlands, Switzerland, Türkiye, the United Kingdom, and the United States. To date, there have been only three valuation studies, including this one, that elicited preferences for reducing the risk of very low birth weight. The present study share similarities with the previous two studies (Ščasný and Zvěřinová, 2014[36]; Ščasný and Zvěřinová, 2016[37]). However, it relies on a different contingent scenario that is a private good scenario linked to a policy regulating chemicals in products sold in the market and on a different elicitation format that is a DBDC with follow-up open-ended question to identify true zero responses.

The central estimate of VSC of very low birth weight across countries is equal to USD_{2022} PPP 1 194 000. Country-specific estimates range from USD_{2022} PPP 805 000 (United Kingdom) to USD_{2022} PPP 1 744 000 (Italy). These estimates are robust to various methodological choices such as the assumption on the distribution of errors, to various methods of respondents screening, and to whether the post-stratification weights are used or not. The WTP function is skewed resulting in a much smaller median VSC of very low birth weight equal to USD_{2022} PPP 614 000 and ranging from USD_{2022} PPP 338 000 (United Kingdom) to USD_{2022} PPP 1 050 000 (Italy).

Three different levels of baseline risk were applied in the valuation design and these three levels (6, 10, and 12 in 1 000) cover prevalence that is typical for most countries in the world. The analysis of survey responses reveals that the baseline risk does not have a statistically significant effect on WTP. Consequently, the estimates of VSC of very low birth weight can be used in a benefit cost analysis regardless of the baseline risk level in the country under consideration.

Other effects, positive or negative, aside from the reduction in the risk of very low birth weight, were considered by a significant share of respondents to the present survey. When other effects are considered the stated WTP value also include the other effects and can result in an overvaluation of the health endpoint of interest. Almost a half of the respondents in the survey considered other effects, and two third of them reported that these other effects increased the WTP they stated. These respondents stated mean (median) VSC equal to USD_{2022} PPP 1 450 000 (797 000) that are much higher than mean (median) VSC equal to USD_{2022} PPP 1 008 000 (499 000) stated by respondents who did not consider other effects. If incentive incompatible responses are excluded as well, the mean (median) VSC shrank to USD_{2022} PPP 992 000 (478 000). These VSC estimates may be applied as conservative values in benefit cost analysis.

Finally, the present study finds that WTP and hence VSC values differ amongst the nine surveyed countries. Moreover, the highest country-specific VSC values are observed in the three countries with the lowest average household income (Italy, Mexico, and Türkiye). In contrast, respondents from the countries with higher incomes are willing to pay less, for example Canada and the United Kingdom. Income elasticity within the sample is 0.3, which is comparable with the values found in similar health risk-reducing stated preference studies and in the other SWACHE survey analyses. This elasticity can be used to transfer WTP values and VSC over time and across countries. However, since cross-country difference in income does not seem to be the main driver of cross-country difference in VSC, an adequate benefit transfer methodology should be developed in future work to transfer VSC into non surveyed countries.

References

Alberini, A. (2017), "Measuring the economic value of the effects of chemicals on ecological systems and human health", *OECD Environment Working Papers*, No. 116, OECD Publishing, Paris, https://doi.org/10.1787/9dc90f8d-en. [39]

Alberini, A. and M. Ščasný (2011), "Context and the VSL: Evidence from a stated preference study in Italy and the Czech Republic", *Environmental and Resource Economics*, Vol. 49/4, pp. 511–538. [41]

Alberini, A. and M. Ščasný (2021), "On the validity of the estimates of the VSL from contingent valuation: Evidence from the Czech Republic", *Journal of Risk and Uncertainty*, Vol. 62/1, pp. 55-87. [43]

Alberini, A. and M. Ščasný (2018), "The benefits of avoiding cancer (or dying from cancer): Evidence from a four-country study", *Journal of health economics*, Vol. 57, pp. 249-262. [42]

Almond, D. and B. Mazumber (2005), "The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data.", *American Economic Review*, Vol. 95/2, pp. 258-262. [16]

Ancker, J. et al. (2006), "Design features of graphs in health risk communication: A systematic review.", *Journal of the American Medical Information Association*, Vol. 13/6, pp. 608–618. [40]

Barker, D. (1995), "Fetal origins of coronary heart disease", *British Journal*, Vol. 311/6998, pp. 171-174. [15]

Behrman, J. and M. Rosenzweig (2004), "Returns to birthweight.", *Review of Economics and statistics*, Vol. 86/2, pp. 586-601. [28]

Bell, M. et al. (2010), "Prenatal exposure to fine particulate matter and birth weight: variations by particulate constituents and sources", *Epidemiology (Cambridge, Mass.)*, Vol. 21/6, p. 884. [1]

Bharadwaj, P., J. Eberhard and C. Neilson (2018), "Health at birth, parental investments, and academic outcomes", *Journal of Labor Economics*, Vol. 36/2, pp. 349-394. [24]

Birks, L. and M. Casas (2016), "Occupational exposure to endocrine-disrupting chemicals and birth weight and length of gestation: a European meta-analysis", *Environmental health perspectives*, Vol. 124/11, pp. 1785-1793. [3]

Black, S., P. Devereux and K. Salvanes (2007), "From the cradle to the labor market? The effect of birth weight on adult outcomes", *The Quarterly Journal of Economics*, Vol. 122/1, pp. 409-439. [33]

Boyle, K., M. Welsh and R. Bishop (1988), "Validation of empirical measures of welfare change: Comment", *Land Economics*, Vol. 64/1, pp. 94–98. [47]

Carson, R. and W. Hanemann (2005), *Contingent valuation*, Elsevier B.V. [44]

Clarke, D., S. Oreffice and C. Quintana-Domeque (2017), "On the Value of Birth Weight", *Human Capital and Economic Opportunity Working Group Working Papers,*, Vol. 18. [34]

Clarke, D., S. Oreffice and C. Quintana-Domeque (2019), "The demand for season of birth", *Journal of Applied Econometrics*, Vol. 34/5, pp. 707-723. [35]

Conley, D. (2000), "Is biology destiny? Birth weight and life chances.", *American Sociological Review*, pp. 458-467. [27]

Cook, C. and J. Fletcher (2015), "Understanding heterogeneity in the effects of birth weight on adult cognition and wages.", *Journal of health economics*, Vol. 41, pp. 107-116. [29]

Currie, J. and R. Hyson (1999), "Is the impact of health shocks cushioned by socioeconomic status? The case of low birthweight", *American Economic Review*, Vol. 89/2, pp. 245-250. [30]

Currie, J. and E. Moretti (2007), "Biology as destiny? Short-and long-run determinants of intergenerational transmission of birth weight", *Journal of Labor economics*, Vol. 25/2, pp. 231-264. [26]

Czajkowski, M. et al. (2022), "On the Inference About Willingness to Pay Distribution Using Contingent Valuation Data", https://doi.org/10.2139/ssrn.4073496. [45]

de Cock, M. et al. (2016), "Prenatal exposure to endocrine disrupting chemicals and birth weight—a prospective cohort study", *Journal of Environmental Science and Health, Part A*, Vol. 51/2, pp. 178-185. [4]

Dockins, C. et al. (2023), "Valuing a reduction in the risk of chronic kidney disease: a large scale multi-country stated preference approach", *Forthcoming OECD Environment Working Paper*. [49]

Edmond, K. and R. Bahl (2007), "Optimal feeding of low-birth-weight infants: technical review", *World Health Organization*, pp. 1-121. [11]

Figlio, D. et al. (2014), "The effects of poor neonatal health on children's cognitive development.", *American Economic Review*, Vol. 104/12, pp. 3921-55. [23]

Fletcher, J. (2011), "The medium term schooling and health effects of low birth weight: Evidence from siblings.", *Economics of Education Review*, Vol. 30/3, pp. 517-527. [19]

Govarts, E. et al. (2016), "Combined effects of prenatal exposures to environmental chemicals on birth weight", *International journal of environmental research and public health*, Vol. 13/5, p. 495. [5]

Hack, M., N. Klein and H. Taylor (1995), "Long-Term Developmental Outcomes of Low Birth Infants", *Future Child*, Vol. 5, pp. 176-196. [14]

Hanemann, W. (1984), "Welfare evaluations in contingent valuation experiments with discrete responses", *American Journal of Agricultural Economics*, Vol. 66/3, pp. 332–341. [48]

Hanemann, W., J. Loomis and B. Kanninen (1991), “Statistical efficiency of double-bounded dichotomous choice contingent valuation”, *American Journal of Agricultural Economics*, Vol. 73/4, pp. 1255–1263. [46]

Hu, J. et al. (2021), “Prenatal exposure to endocrine disrupting chemical mixtures and infant birth weight: A Bayesian analysis using kernel machine regression”, *Environmental research*, Vol. 195/110749. [10]

Hunt, A. (2011), “Policy Interventions to Address Health Impacts Associated with Air Pollution, Unsafe Water Supply and Sanitation, and Hazardous Chemicals”, *OECD Environment Working Papers*, Vol. No. 35/OECD Publishing, Paris., https://doi.org/10.1787/5kg9qx. [13]

Johnson, R. and R. Schoeni (2011), “Early-life origins of adult disease: national longitudinal population-based study of the United States”, *American journal of public health*, Vol. 101/12, pp. 2317-2324. [17]

Johnson, R. and R. Schoeni (2011), “The influence of early-life events on human capital, health status, and labor market outcomes over the life course”, *The BE journal of economic analysis & policy*, Vol. 11/3. [31]

Kormos, C. et al. (2013), “Low birth weight and intelligence in adolescence and early adulthood: a meta-analysis”, *Journal of Public Health*, Vol. 36/2, pp. 213-224. [20]

Kriström, B. (1997), “Spike Models in Contingent Valuation”, *American Journal of Agricultural Economics*, Vol. 79/3, pp. 1013-1023. [53]

Lenters, V. et al. (2016), “Prenatal phthalate, perfluoroalkyl acid, and organochlorine exposures and term birth weight in three birth cohorts: multi-pollutant models based on elastic net regression”, *Environmental health perspectives*, Vol. 124/3, pp. 365-372. [6]

Lin, M. and J. Liu (2009), “Do lower birth weight babies have lower grades? Twin fixed effect and instrumental variable method evidence from Taiwan.”, *Social Science & Medicine*, Vol. 68/10, pp. 1780-1787. [21]

Manski, C. and S. Lerman (1977), “The estimation of choice probabilities from choice based samples”, *Econometrica*, Vol. 45/8, pp. 1977-1988. [52]

Miao, M. et al. (2011), “In utero exposure to bisphenol-A and its effect on birth weight of offspring”, *Reproductive toxicology*, Vol. 32/1, pp. 64-68. [2]

Mourato, S. and G. Shreedhar (2021), “Conducting economic valuation surveys during extreme events”, *OECD Environment Working Papers*, OECD Publishing, Paris, https://doi.org/10.1787/85477cd9-en. [50]

Navrud, S. (2018), “Assessing the economic valuation of the benefits of regulating chemicals: Lessons learned from five case studies”, *OECD Environment Working Papers*, No. 136, OECD Publishing, Paris, https://doi.org/10.1787/9a061350-en. [38]

OECD (2022), *OECD Economic Outlook, Volume 2022 Issue 2*, OECD Publishing, Paris, https://doi.org/10.1787/f6da2159-en. [51]

Rosenzweig, M. and J. Zhang (2013), “Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of Chinese twins”, *Journal of Development Economics*, Vol. 104, pp. 245-260. [32]

Royer, H. (2009), "Separated at girth: US twin estimates of the effects of birth weight", *American Economic Journal: Applied Economics*, Vol. 1/1, pp. 49-85. [25]

Ščasný, M. and I. Zvěřinová (2016), *A Valuation of Environment Related Health Impacts on Fertility and Birth Outcomes in Canada*, Health Canada. [37]

Ščasný, M. and I. Zvěřinová (2014), *Stated-preference study to examine the economic value of benefits of avoiding selected adverse human health outcomes due to exposure to chemicals in the European Union, Part II: Fertility and Developmental Toxicity*, European Chemicals Agency (ECHA). [36]

Sun, X. et al. (2016), "The associations between birth weight and exposure to fine particulate matter (PM2. 5) and its chemical constituents during pregnancy: A meta-analysis", *Environmental pollution*, Vol. 211, pp. 38-47. [7]

Torche, F. and G. Echevarría (2011), "The effect of birthweight on childhood cognitive development in a middle-income country", *International journal of epidemiology*, Vol. 40/4, pp. 1008-1018. [22]

Van der Bergh, G., M. Lindeboom and F. Portrait (2006), "Economic Conditions Early in Life and Individual Mortality.", *American Economic Review*, Vol. 96/1, pp. 290-302. [18]

Wikström, S. et al. (2020), "Maternal serum levels of perfluoroalkyl substances in early pregnancy and offspring birth weight", *Pediatric Research*, Vol. 87/6, pp. 1093-1099. [9]

Woods, M. et al. (2017), "Gestational exposure to endocrine disrupting chemicals in relation to infant birth weight: a Bayesian analysis of the HOME Study", *Environmental Health*, Vol. 16/1, pp. 1-12. [8]

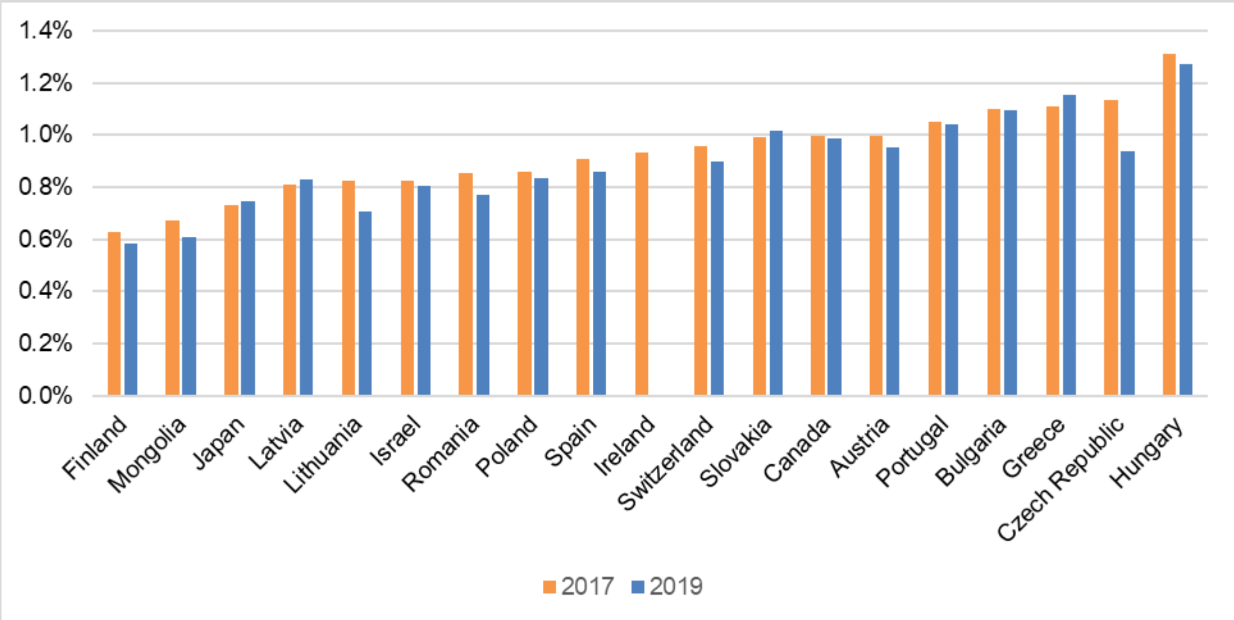
World Health Organization (2004), "International statistical classification of diseases and related health problems, tenth revision, 2nd ed.". [12]

Annex A. Additional figures

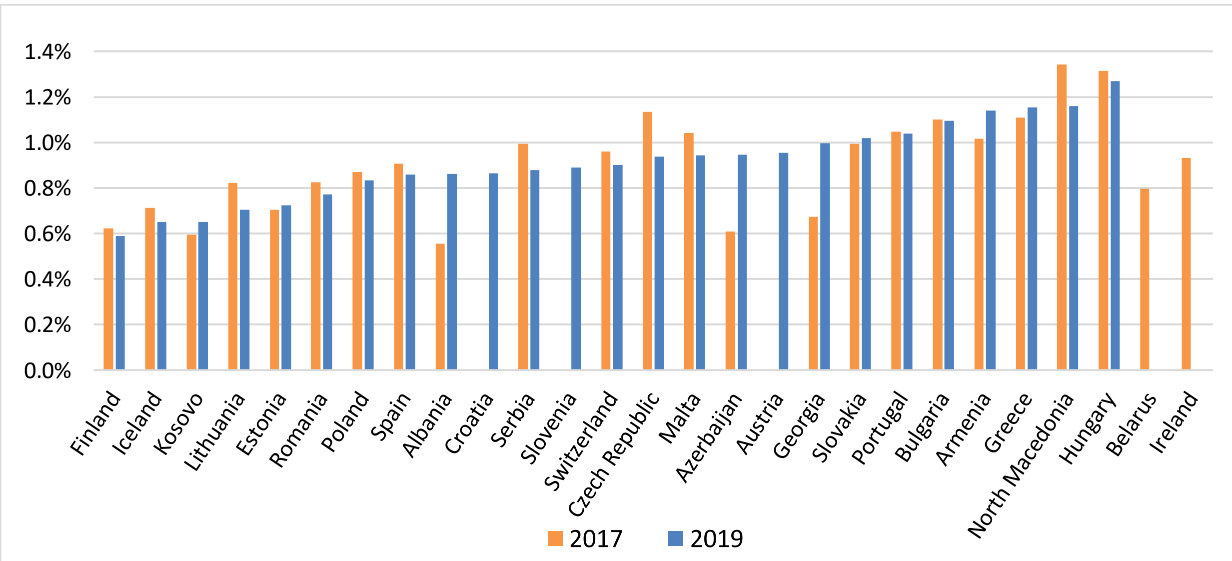
Prevalence of very low birth weight by country

Figure A.1. Very low birth weight prevalence in 2017 and 2019 (percentages of total live births)

Panel A: UN data



Panel B: Eurostat data



Source: Authors' own calculation based on available UN data and Eurostat data, Live births by birth weight and duration of gestation

Description of the health problems associated with very low birth weight and its consequences as shown in the questionnaire

Figure A.2. Neurosensory Problems

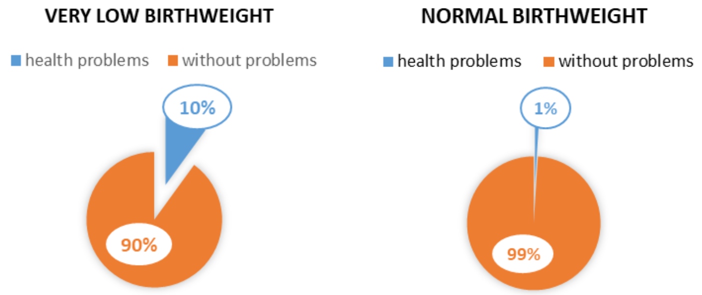
- abnormal development of the brain that control movement, balance, and posture
- fluid collecting in the brain, blindness or deafness, or epilepsy

Treatment

- is not curable - only improvement
- medicines, orthopaedic surgery, pain management, physical therapy

Quality of life impact

- disorder of movement and motor function, respiratory problems, impaired functional status
- need of assistance



Source: Hack, Klein and Taylor (1995[14]).

Figure A.3. Behavioural and Social Competency Problems

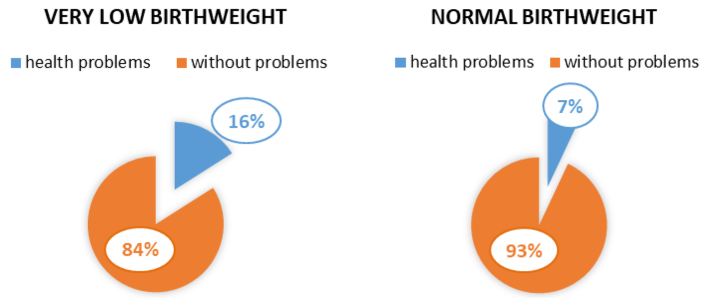
- hyperactivity and attentional weaknesses
- disruptive behaviour and impulsivity

Treatment

- is not curable - only improvement
- medication, diet, psychotherapy, education or training

Quality of life impact

- social problems, difficulty organizing activities, diminished school performance and vocational achievement
- special educational needs



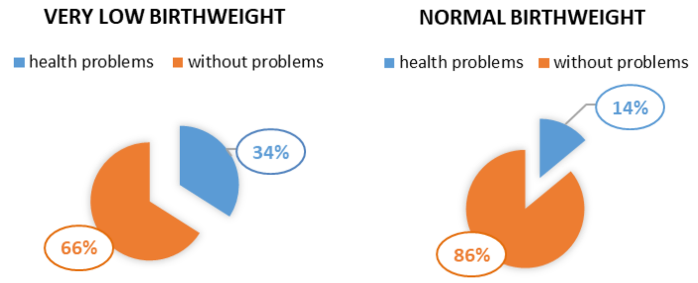
Source: Hack, Klein and Taylor (1995[14]).

Figure A.4. Learning Disabilities

- poorer language abilities, memory, motor coordination and problem solving abilities
- learning problems

  Treatment
- special education assistance and help

  Quality of life impact
- impairments in life skills (communication, self-care, home living, social skills)
- school problems (grade repetition or placement in special education programs)

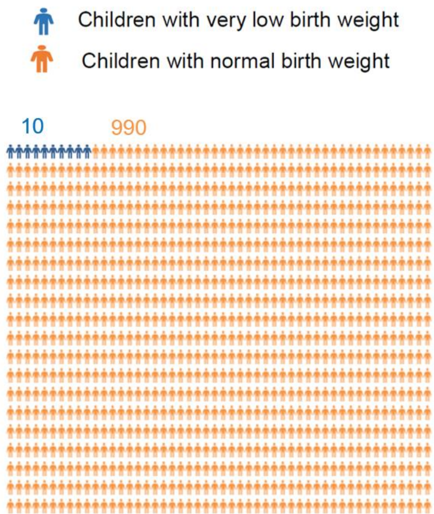


Source: Hack, Klein and Taylor (1995[14]).

Visuals for risk used in the questionnaire

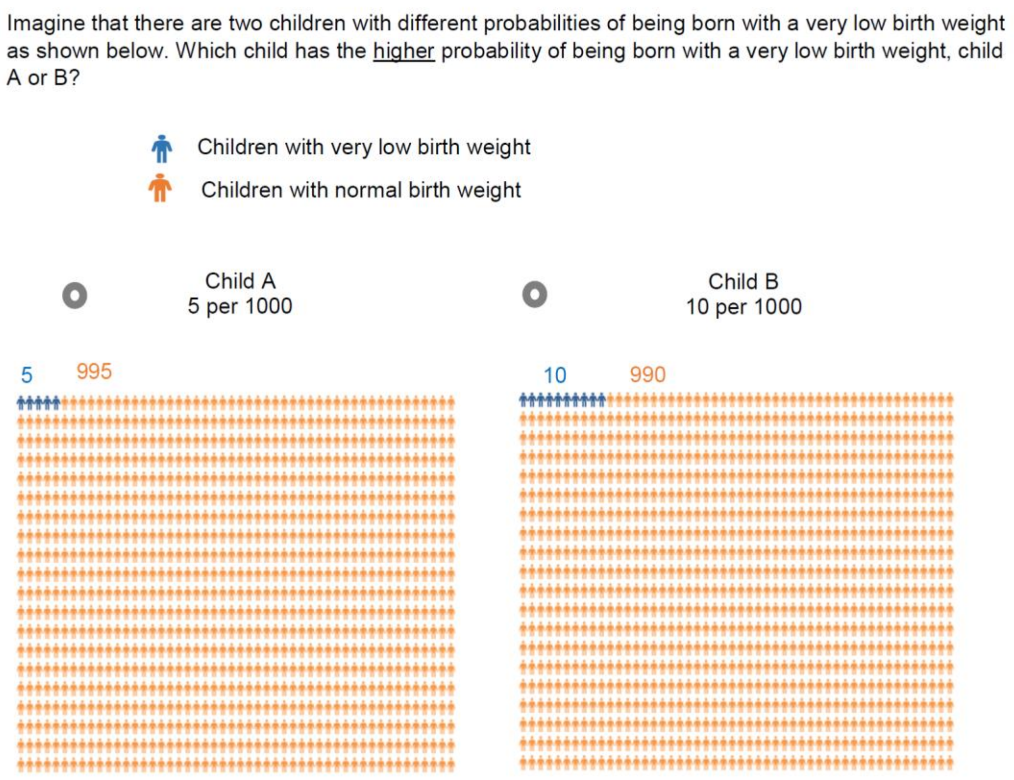
Figure A.5. Example of a graph shown in the questionnaire to explain probability of a child to be born with a very low birth weight

The graph shows the probability of a child to be born with a very low birth weight. There are 10 blue children with very low birth weight and 990 orange children with normal birth weight.



Note: Authors' own elaboration.

Figure A.6. Example of one probability quiz



Note: Authors' own elaboration.

Annex B. WTP from previous studies

Table B.1. Value of a Statistical Case of very low birth weight and low birth weight, in EUR PPP

Health outcome	Very low birth weight			Low birth weight		
Target population	Who plan to have a baby		General population	Who plan to have a baby		General population
Type of the good / Context	private good (vitamins)	public good (regulation)	public good (regulation)	private good (vitamins)	public good (regulation)	public good (regulation)
Canada (no Zika)*	201 858	875 380	423 547	136 032	702 401	220 834
Czech Republic	120 558	405 517	546 737	*NA*	*NA*	*NA*
Italy	245 157	532 549	669 255	*NA*	*NA*	*NA*
Netherlands	*NA*	620 842	*NA*	*NA*	*NA*	*NA*
United Kingdom	80 090	420 130	316 092	*NA*	*NA*	*NA*
ECHA-4**	120 165	386 114	477 838	*NA*	*NA*	*NA*
EU28 ***	126 200	405 500	548 300	*NA*	*NA*	*NA*

Note: VSC of very low birth weight and low birth weight estimates for ECHA and Health Canada studies taken from Ščasný & Zvěřinová (2014; 2016)
* Based on WTP for respondents who did not think about Zika virus.
** WTP means estimated for the four European countries included in the ECHA study, namely Czech Republic, Italy, Netherlands, and the United Kingdom, population weighted.
*** Based on income-adjusted benefit transfer, using the income elasticity of 0.7, population weighted, rounded at hundreds.
All nominal values are expressed in EURO PPP, using purchasing power standard for individual consumption for the year 2012 by Eurostat that is for CZK 17.0603 for the Czech Republic, EUR 1.02356 for Italy, EUR 1.11216 for the Netherlands, and GBP 0.945661 per Euro. For Canada CAD were converted using Purchase Power Standard at 1.645 CAD per 1 EUR.

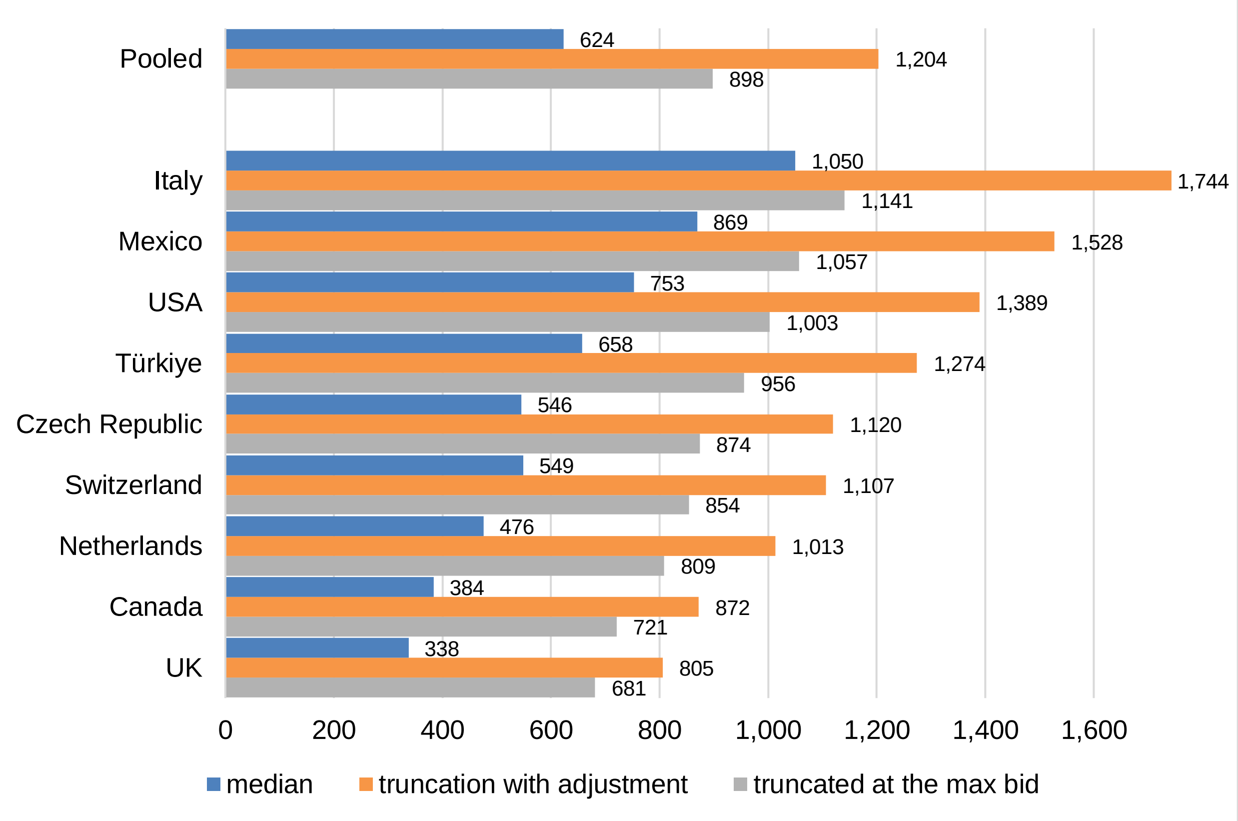
Table B.2. Value of a statistical case of very low birth weight, private good scenario, respondents who plan to have a baby, comparison with other studies

Study	Payment vehicle	Country	VSC of very low birth weight elicited by the study	VSC in USD_{2022} adjusted for GDP per capita growth and inflation
Present study, DBDC	Safer products	Canada	USD_{2022} 872 000	872 000
Present study, SBDC	Safer products	Canada	USD_{2022} 395 000	395 000
Ščasný and Zvěřinová (2016[37])	Complex of vitamins	Canada*	EUR_{2016} 202 000	325 000
Present study, DBDC	Safer products	Czech Republic	USD_{2022} 1 120 000	1 120 000
Present study, SBDC	Safer products	Czech Republic	USD_{2022} 651 000	651 000
Ščasný and Zvěřinová (2014[36])	Complex of vitamins	Czech Republic	EUR_{2014} 121 000	231 000
Present study, DBDC	Safer products	Italy	USD_{2022} 1 744 000	1 744 000
Present study, SBDC	Safer products	Italy	USD_{2022} 929 000	929 000
Ščasný and Zvěřinová (2014[36])	Complex of vitamins	Italy	EUR_{2014} 245 000	379 000
Present study, DBDC	Safer products	United Kingdom	USD_{2022} 805 000	805 000
Present study, SBDC	Safer products	United Kingdom	USD_{2022} 234 000	234 000
Ščasný and Zvěřinová (2014[36])	Complex of vitamins	United Kingdom	EUR_{2014} 80 000	137 000

Note: * Based on WTP estimated only for respondents who did not think about Zika virus when answering the valuation question. All nominal values for Ščasný and Zvěřinová (2014[36]) and Ščasný and Zvěřinová (2016[37]) are expressed in EUR PPP, using purchasing power parities for individual consumption for the year 2012 by Eurostat that is for CZK 17.0603 for the Czech Republic, EUR 1.02356 for Italy and GBP 0.945661 per Euro. For Canada CAD were converted using purchasing power parities at 1.645 CAD per 1 EUR. VSC in USD_{2022} are adjusted for GDP per capita growth and inflation using equation (3) and an income elasticity of 0.7 derived in the previous studies.

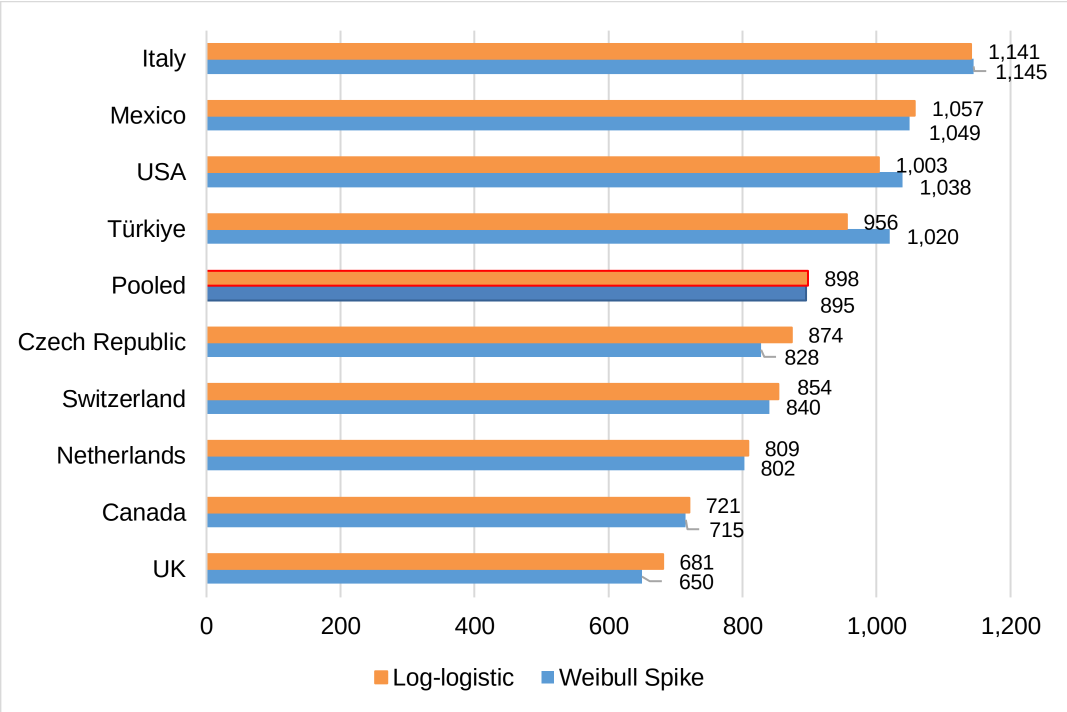
Annex C. Additional results

Figure C.1. Values of a statistical case of very low birth weight for countries and pooled for all countries estimated from models with log-logistic, assuming various individual WTP estimates, (USD PPP)



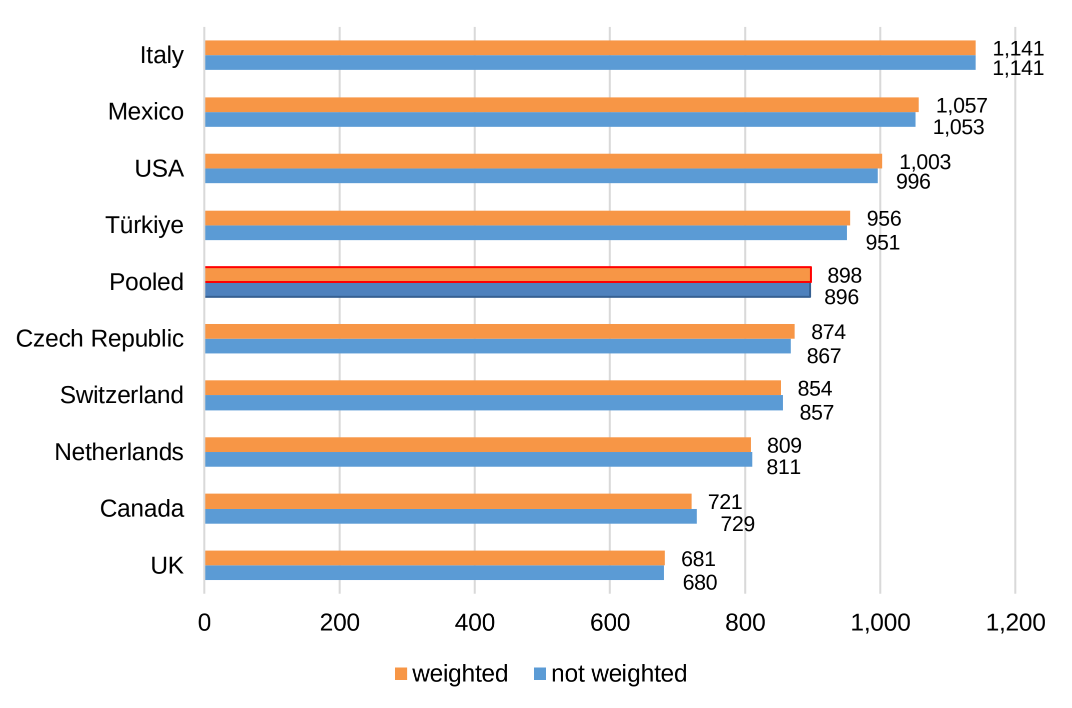
Note: Mean of VSC estimates are based on individual WTP estimates assuming log-logistic distribution and weighted, with speeders and quiz failures excluded.

Figure C.2. Values of a statistical case of very low birth weight for countries and pooled for all countries estimated from models with log-logistic and Weibull distribution with “spike” (USD PPP)



Note: Values are based on individual mean WTP truncated at the maximum bid, speeders and quiz failures excluded, weighted

Figure C.3. Values of a statistical case of very low birth weight for countries and pooled for all countries with and without weighting estimated from log-logistic model (USD PPP)



Note: Values based on individual mean WTP truncated at the maximum bid, speeders quiz failures excluded

Table C.1. Comparison between target population and samples for gender, age, and education

	CZ	IT	MX	NL	CH	TR	UK	USA	CA
GENDER									
Male_quota	49%	49%	49%	50%	50%	49%	49%	49%	50%
Male_sample	49%	49%	47%	43%	47%	47%	37%	43%	41%
Male_diff wrt quota	0%	0%	-4%	-14%	-6%	-4%	-24%	-12%	-18%
AGE									
18-24_quota	9%	6%	36%	6%	5%	18%	13%	18%	8%
25-34_quota	59%	49%	48%	63%	56%	58%	57%	57%	60%
35-39_quota	22%	28%	11%	22%	27%	17%	21%	17%	23%
40-44_quota	8%	13%	4%	7%	9%	5%	7%	5%	7%
45-65_quota	2%	4%	0%	2%	3%	1%	3%	2%	2%
18-24_sample	20%	13%	46%	22%	15%	33%	22%	22%	16%
25-34_sample	48%	42%	38%	53%	48%	43%	48%	49%	47%
35-39_sample	22%	28%	11%	20%	25%	17%	22%	20%	26%
40-44_sample	8%	13%	5%	4%	9%	5%	7%	6%	9%
45-65_sample	2%	4%	0%	1%	3%	2%	2%	2%	2%
18-24_diff wrt quota	122%	117%	28%	267%	200%	83%	69%	22%	100%
25-34_diff wrt quota	-19%	-14%	-21%	-16%	-14%	-26%	-16%	-14%	-22%
35-39_diff wrt quota	0%	0%	0%	-9%	-7%	0%	5%	18%	13%
40-44_diff wrt quota	0%	0%	25%	-43%	0%	0%	0%	20%	29%
45-65_diff wrt quota	0%	0%	0%	-50%	0%	100%	-33%	0%	0%
EDUCATION									
Low_quota	6%	38%	60%	20%	11%	58%	20%	9%	8%
Medium_quota	70%	43%	22%	39%	45%	20%	33%	42%	32%
High_quota	24%	20%	18%	40%	44%	22%	47%	48%	59%
Low_sample	4%	8%	20%	11%	8%	13%	2%	1%	1%
Medium_sample	72%	71%	59%	36%	39%	63%	35%	41%	22%
High_sample	25%	21%	21%	53%	53%	24%	62%	59%	78%
Low_diff wrt quota	-33%	-79%	-67%	-45%	-27%	-78%	-90%	-89%	-88%
Medium_diff wrt quota	3%	65%	168%	-8%	-13%	215%	6%	-2%	-31%
High_diff wrt quota	4%	5%	17%	33%	20%	9%	32%	23%	32%

Note: "diff wrt quota" describes a relative difference with respect to quota, in percentage.

Table C.2. Robustness check, WTP models assuming with different distributions, speeders and quiz failures excluded

	Log-logistic		Log-normal		Weibull		Spike log-normal		Spike Weibull	
Intercept	6.958	***	4.266	***	5.804	***	3.232	***	5.395	***
	(0.131)		(-0.086)		(0.115)		(0.070)		(0.106)	
Czech Republic	-0.285	***	-0.171	**	-0.331	***	-0.184	***	-0.322	***
	(0.084)		(0.057)		(0.070)		(0.057)		(0.070)	
Italy	0.326	***	0.206	***	0.175	*	0.177	**	0.171	*
	(0.087)		(0.059)		(0.074)		(0.059)		(0.074)	
Mexico	0.139	.	0.116	.	0.026	.	0.125	*	0.032	.
	(0.087)		(0.060)		(0.075)		(0.060)		(0.075)	
Netherlands	-0.426	***	-0.238	***	-0.340	***	-0.240	***	-0.336	***
	(0.087)		(0.059)		(0.072)		(0.059)		(0.072)	
Switzerland	-0.256	**	-0.148	*	-0.245	***	-0.200	***	-0.250	***
	(0.086)		(0.059)		(0.071)		(0.058)		(0.071)	
Türkiye	0.048	.	-0.116	.	-0.073	.	-0.049	.	-0.054	.
	(0.093)		(0.064)		(0.082)		(0.064)		(0.082)	
United Kingdom	-0.717	***	-0.438	***	-0.582	***	-0.439	***	-0.574	***
	(0.085)		(0.057)		(0.068)		(0.057)		(0.068)	
Canada	-0.553	***	-0.351	***	-0.459	***	-0.345	***	-0.452	***
	(0.086)		(0.058)		(0.070)		(0.058)		(0.070)	
$\Delta RISK$	0.055	***	0.038	***	0.045	***	0.037	***	0.045	***
	(0.012)		(0.008)		(0.009)		(0.008)		(0.009)	
log(bid)	-0.911	***	-0.565	***	-0.706	***	-0.424	***	-0.653	***
	(0.015)		(0.010)		(0.012)		(0.007)		(0.011)	
Distribution	Log-logistic		Log-normal		Weibull		Spike log-normal		Spike Weibull	
Log-likelihood	-8 798.9		-8 777.06		-8 796.93		-9 407.48		-9 163.84	
AIC	17 621.8		17 576.13		17 615.86		18 836.96		18 349.67	
BIC	17 703.9		17 651.42		17 691.15		18 912.26		18 424.97	
Observations	6 940		6 940		6 940		6 940		6 940	
VSC of very low birth weight, USD PPP										
Mean VSC	1 194 223		1 211 924		1 144 026		1 414 745		1 185 965	
Mean VSC*	894 810		899 792		883 762		953 201		893 536	
Median VSC	613 816		621 743		664 818		690 634		666 064	

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses.
Notes: mean derived from individual mean WTP truncated at the maximum bid with adjustment; * mean derived from individual mean WTP truncated at the maximum bid.

Table C.3. Estimates of average and median WTP based on different screening using DBDC, log-logistic distribution

Excluded	speeders quiz failed		speeders		quiz failed		none	
	Coeff		Coeff		Coeff		Coeff	
Intercept	7.118	***	6.95752	***	6.68956	***	6.58146	***
	(0.15)		(0.131)		(0.133)		(0.114)	
Czech Republic	-0.310	**	-0.28492	***	-0.16397	.	-0.14449	.
	(0.095)		(0.084)		(0.088)		(0.077)	
Italy	0.316	**	0.32564	***	0.39019	***	0.36678	***
	(0.099)		(0.087)		(0.09)		(0.079)	
Mexico	0.143	.	0.13872	.	0.26659	**	0.22791	**
	(0.1)		(0.087)		(0.091)		(0.078)	
Netherlands	-0.423	***	-0.42645	***	-0.25687	**	-0.25622	***
	(0.099)		(0.087)		(0.09)		(0.077)	
Switzerland	-0.269	**	-0.25608	**	-0.13911	.	-0.1602	*
	(0.098)		(0.086)		(0.088)		(0.077)	
Türkiye	-0.141	.	0.0478	.	-0.4812	***	-0.32241	***
	(0.11)		(0.093)		(0.096)		(0.081)	
United Kingdom	-0.750	***	-0.71714	***	-0.56287	***	-0.55155	***
	(0.096)		(0.085)		(0.087)		(0.076)	
Canada	-0.613	***	-0.55332	***	-0.45143	***	-0.41806	***
	(0.097)		(0.086)		(0.088)		(0.077)	
$\Delta RISK$	0.063	***	0.05538	***	0.05467	***	0.0436	***
	(0.013)		(0.012)		(0.012)		(0.01)	
log(bid)	-0.939	***	-0.91098	***	-0.89951	***	-0.87362	***
	(0.017)		(0.015)		(0.015)		(0.013)	
Distribution	Log-logistic		Log-logistic		Log-logistic		Log-logistic	
Log-likelihood	8 775.2805		11 172.758		10 540.73		13 893.856	
AIC	17 572.561		22 367.516		21 103.459		27 809.712	
BIC	17 647.857		2 2445.561		21 180.574		27 889.872	
Observations	6 940		8 910		8 188		10 800	
Mean VSC, in USD PPP								
individual mean WTP truncated at the maximum bid	896 453		933 950		869 922		903 698	
individual mean WTP truncated at the maximum bid with adjustment	1 201 945		1 284 213		1 155 448		1 223 771	
individual median WTP	622 760		673 745		571 134		609 182	

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses.

Table C.4. Estimates of average and median WTP based on different screening using DBDC, log-normal distribution

Excluded	speeders quiz failed	speeders	quiz failed	none
	Coeff	Coeff	Coeff	Coeff
Intercept	4.26616 ***	4.16641 ***	4.03488 ***	3.96986 ***
	(-0.086)	(0.075)	(0.077)	(0.066)
Czech Republic	-0.17108 **	-0.15968 **	-0.09268 .	-0.08298 .
	(0.057)	(0.051)	(0.053)	(0.047)
Italy	0.20579 ***	0.20581 ***	0.24548 ***	0.22629 ***
	(0.059)	(0.052)	(0.054)	(0.047)
Mexico	0.11595 .	0.11220 *	0.18221 ***	0.15973 ***
	(0.060)	(0.052)	(0.055)	(0.047)
Netherlands	-0.23844 ***	-0.24226 ***	-0.14573 **	-0.14659 **
	(0.059)	(0.052)	(0.054)	(0.046)
Switzerland	-0.14836 *	-0.14405 **	-0.07547 .	-0.09143 *
	(0.059)	(0.052)	(0.053)	(0.046)
Türkiye	-0.11578 .	-0.00854 .	-0.31630 ***	-0.22817 ***
	(0.064)	(0.054)	(0.056)	(0.048)
United Kingdom	-0.43784 ***	-0.42165 ***	-0.33299 ***	-0.32913 ***
	(0.057)	(0.051)	(0.053	(0.046
Canada	-0.35116 ***	-0.31858 ***	-0.26085 ***	-0.24347 ***
	(0.058)	(0.051)	(0.053)	(0.046)
$\Delta RISK$	0.03787 ***	0.03351 ***	0.03308 ***	0.02633 ***
	(0.008)	(0.007)	(0.007)	(0.006)
log(bid)	-0.56463 ***	-0.54669 ***	-0.54342 ***	-0.52744 ***
	(0.010)	(0.008)	(0.008)	(0.007)
Distribution	Log-normal	Log-normal	Log-normal	Log-normal
Log-likelihood	-8 777.06	-11 178.39	-10 540.57	-13 896.78
AIC	17 576.13	22 378.78	21 103.13	27 815.56
BIC	17 651.42	22 456.82	21 180.25	27 895.72
Observations	6 940	8 910	8 188	10 800
Mean VSC, in USD PPP				
individual mean WTP truncated at the maximum bid	899 792	937 486	873 561	907 219
individual mean WTP truncated at the maximum bid with adjustment	1 211 924	1 295 709	1 165 329	1 234 749
individual median WTP	621 743	671 938	572 515	610 084

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses.

Table C.5. Estimates of average and median WTP based on different screening using DBDC, log-normal distribution, "spike" configuration

Excluded	speeders quiz failed		speeders		quiz failed		none	
	Coeff		Coeff		Coeff		Coeff	
Intercept	3.23152	***	3.22621	***	3.22621	***	3.22621	***
	(0.070)		(0.062)		(0.062)		(0.056)	
Czech Republic	-0.18406	***	-0.17518	***	-0.17518	***	-0.17518	***
	(0.057)		(0.051)		(0.051)		(0.047)	
Italy	0.17749	**	0.17424	***	0.17424	***	0.17424	***
	(0.059)		(0.052)		(0.052)		(0.047)	
Mexico	0.12474	*	0.11396	*	0.11396	*	0.11396	*
	(0.060)		(0.052)		(0.052)		(0.047)	
Netherlands	-0.24003	***	-0.24808	***	-0.24808	***	-0.24808	***
	(0.059)		(0.052)		(0.052)		(0.046)	
Switzerland	-0.20031	***	-0.19121	***	-0.19121	***	-0.19121	***
	(0.058)		(0.052)		(0.052)		(0.046)	
Türkiye	-0.04928	.	0.04411	.	0.04411	.	0.04411	.
	(0.064)		(0.054)		(0.054)		(0.048)	
United Kingdom	-0.43917	***	-0.43905	***	-0.43905	***	-0.43905	***
	(0.057)		(0.051)		(0.051)		(0.046))	
Canada	-0.34471	***	-0.32596	***	-0.32596	***	-0.32596	***
	(0.058)		(0.051)		(0.051)		(0.046)	
$\Delta RISK$	0.03748	***	0.03376	***	0.03376	***	0.03376	***
	(0.008)		(0.007)		(0.007)		(0.006)	
log(bid)	-0.42391	***	-0.41869	***	-0.41869	***	-0.41869	***
	(0.007)		(0.006)		(0.006)		(0.005)	
Distribution	Spike log-normal		Spike log-normal		Spike log-normal		Spike log-normal	
Log-likelihood	-9 407.48		-11 911.60		-11 911.60		-14 703.12	
AIC	18 836.96		23 845.20		23 845.20		29 428.24	
BIC	18 912.26		23 923.24		23 923.24		29 508.40	
Observations	6 940		8 910		8 910		10 800	
Mean VSC								
individual mean WTP truncated at the maximum bid	953 201		985 921		921 590		950 083	
individual mean WTP truncated at the maximum bid with adjustment	1 414 745		1 495 954		1 334 728		1 396 006	
individual median WTP	690 634		755 845		614 327		656 553	

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table C.6. Estimates of average and median WTP based on different screening using DBDC, Weibull distribution

Excluded	speeders quiz failed		speeders		quiz failed		none	
	Coeff		Coeff		Coeff		Coeff	
Intercept	5.80449	***	5.71338	***	5.42535	***	5.37843	***
	(0.115)		(0.102)		(0.102)		(0.088)	
Czech Republic	-0.33053	***	-0.28463	***	-0.20235	**	-0.16070	**
	(0.070)		(0.062)		(0.064)		(0.056)	
Italy	0.17547	*	0.20053	**	0.24263	***	0.24132	***
	(0.074)		(0.066)		(0.067)		(0.059)	
Mexico	0.02593	.	0.03603	.	0.13579	*	0.11822	*
	(0.075)		(0.066)		(0.068)		(0.058)	
Netherlands	-0.34013	***	-0.32982	***	-0.21866	***	-0.21115	***
	(0.072)		(0.063)		(0.064)		(0.055)	
Switzerland	-0.24474	***	-0.22181	***	-0.13074	*	-0.13374	*
	(0.071)		(0.063)		(0.063)		(0.055)	
Türkiye	-0.07281	.	0.06996	.	-0.24503	***	-0.14092	*
	(0.082)		(0.072)		(0.069)		(0.059)	
United Kingdom	-0.58222	***	-0.53370	***	-0.44008	***	-0.41533	***
	(0.068)		(0.061)		(0.061)		(0.054)	
Canada	-0.45886	***	-0.40438	***	-0.33932	***	-0.30599	***
	(0.070)		(0.062)		(0.063)		(0.055)	
$\Delta RISK$	0.04524	***	0.03923	***	0.03956	***	0.03085	***
	(0.009)		(0.008)		(0.009)		(0.008)	
log(bid)	-0.70565	***	-0.69033	***	-0.67063	***	-0.65756	***
	(0.012)		(0.011)		(0.011)		(0.009)	
Distribution	Weibull		Weibull		Weibull		Weibull	
Log-likelihood	-8 796.93		-11 198.89		-10 568.62		-13 922.24	
AIC	17 615.86		22 419.78		21 159.23		27 866.49	
BIC	17 691.15		22 497.82		21 236.35		27 946.65	
Observations	6 940		8 910		8 188		10 800	
Mean VSC								
individual mean WTP truncated at the maximum bid	883 762		923 638		856 025		891 881	
individual mean WTP truncated at the maximum bid with adjustment	1 144 026		1 228 750		1 093 621		1 162 541	
individual median WTP	664 818		716 730		607 261		645 993	

Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses.

Table C.7. Estimates of average and median WTP based on different screening using DBDC, Weibull distribution with spike

Excluded	speeders quiz failed		speeders		quiz failed		none	
	Coeff		Coeff		Coeff		Coeff	
Intercept	5.39514	***	5.34923	***	5.12923	***	5.12829	***
	(0.106)		(0.093)		(0.093)		(0.081)	
Czech Republic	-0.32202	***	-0.27947	***	-0.19846	**	-0.15891	**
	(0.070)		(0.062)		(0.064)		(0.056)	
Italy	0.17134	*	0.19619	**	0.23860	***	0.23773	***
	(0.074)		(0.066)		(0.067)		(0.059)	
Mexico	0.03178	.	0.03960	.	0.13805	*	0.11957	*
	(0.075)		(0.066)		(0.068)		(0.058)	
Netherlands	-0.33565	***	-0.32746	***	-0.21775	***	-0.21125	***
	(0.072)		(0.063)		(0.064)		(0.055)	
Switzerland	-0.25027	***	-0.22714	***	-0.13762	*	-0.13960	*
	(0.071)		(0.063)		(0.063)		(0.055)	
Türkiye	-0.05367	.	0.08513	.	-0.23014	***	-0.12904	*
	(0.082)		(0.072)		(0.069)		(0.059)	
United Kingdom	-0.57381	***	-0.53133	***	-0.43577	***	-0.41465	***
	(0.068)		(0.061)		(0.061)		(0.054)	
Canada	-0.45167	***	-0.40062	***	-0.33543	***	-0.30434	***
	(0.070)		(0.062)		(0.063)		(0.055)	
$\Delta RISK$	0.04459	***	0.03892	***	0.03906	***	0.03066	***
	(0.009)		(0.008)		(0.009)		(0.008)	
log(bid)	-0.65288	***	-0.64340	***	-0.63214	***	-0.62516	***
	(0.011)		(0.010)		(0.010)		(0.008)	
Distribution	Spike Weibull		Spike Weibull		Spike Weibull		Spike Weibull	
Log-likelihood	-9 163.84		-11 636.84		-10 974.48		-14 416.34	
AIC	18 349.67		23 295.67		21 970.95		28 854.69	
BIC	18 424.97		23 373.72		22 048.07		28 934.85	
Observations	6 940		8 910		8 188		10 800	
Mean VSC								
individual mean WTP truncated at the maximum bid	893 536		932 147		863 480		897 935	
individual mean WTP truncated at the maximum bid with adjustment	1 185 965		1 269 776		1 122 982		1 188 919	
individual median WTP	666 064		721 474		605 092		645 275	

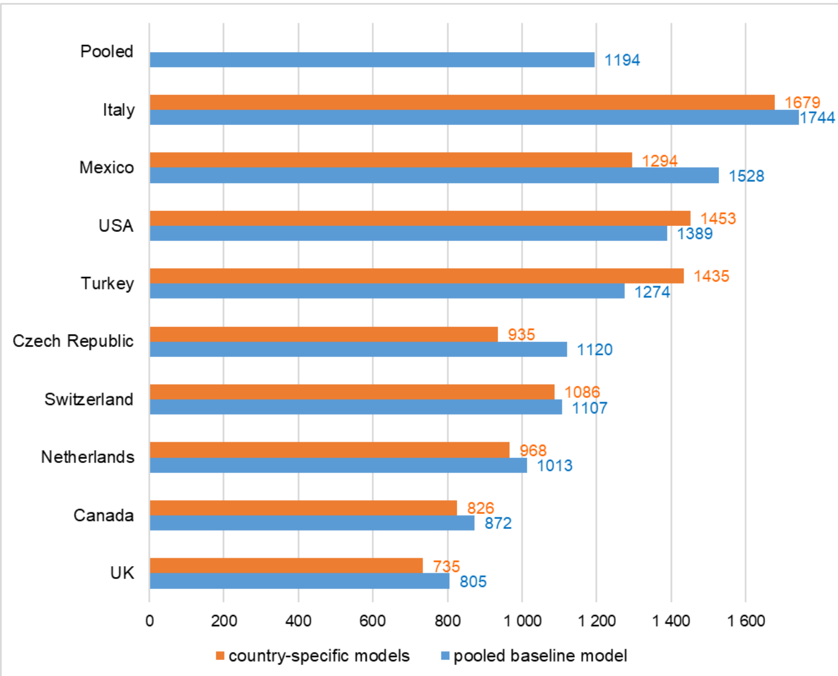
Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses.

Table C.8. Estimations based on country samples taken separately

	Czech Republic	Canada	Italy	Mexico	Netherland	Switzerland	Türkiye	UK	US
Intercept	8.481***	6.061***	8.056***	8.894***	6.579***	7.321***	4.042***	6.684***	5.680***
	(0.411)	(0.364)	(0.43)	(0.468)	(0.386)	(0.41)	(0.364)	(0.372)	(0.357)
$\Delta RISK$	0.039	0.104**	0.017	0.012	0.097*	0.041	0.053	0.106**	0.076.
	(0.037)	(0.038)	(0.039)	(0.041)	(0.04)	(0.039)	(0.045)	(0.037)	(0.039)
log(bid)	-1.158***	-0.900***	-0.997***	-1.132***	-0.942***	-0.992***	-0.518***	-1.009***	-0.751***
	(0.053)	(0.047)	(0.052)	(0.058)	(0.05)	(0.051)	(0.044)	(0.049)	(0.044)
Log-likelihood	-1102.6	-1042.0	-971.7	-907.3	-956.7	-1007.0	-671.1	-1103.8	-950.7
AIC	2211.3	2090.0	1949.4	1820.6	1919.5	2020.0	1348.2	2213.6	1907.3
BIC	2225.4	2104.0	1963.5	1834.4	1933.2	2033.9	1361.5	2227.7	1921.2
Observations	838	795	829	754	732	765	627	833	767
VSC, 1000 USD PPP									
Mean VSC	935	826	1679	1294	968	1086	1435	735	1453
Mean VCS*	750	652	1083	878	759	851	745	626	918
Median VSC	523	380	1044	812	468	553	1074	330	831

Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; standard errors are given in parentheses. Log-logistic distribution, weighted, speeders and who failed the probability quiz excluded. Mean VSC are derived from individual mean WTP truncated at the maximum bid with adjustment; * Mean VSC are derived from individual mean WTP truncated at the maximum bid.

Figure C.4. Value of statistical case of very low birth weight by country, pooled data and country-specific models



Note: Mean of VSC of very low birth weight estimates are based on individual WTP truncated at the maximum bids with adjustment, assuming log-logistic distribution, weighted, with speeders and quiz failures excluded. The estimates are derived from column (1) of

Table 5.2 and from Table C.8.

Table C.9. WTP model that controls for the baseline risk level in interaction with risk reduction

	(A)		(A23)		(B)		(B23)	
Intercept	6.720	***	6.559	***				
	(0.133)		(0.221)					
$\Delta RISK \times SQ6$	0.067	*	0.083		943.894	***	930.676	***
	(0.033)		(0.064)		(41.267)		(39.248)	
$\Delta RISK \times SQ10$	0.080	***	0.078		616.416	***	928.615	***
	(0.017)		(0.064)		(17.728)		(38.768)	
$\Delta RISK \times SQ12$	0.057	***	0.091		511.485	***	954.295	***
	(0.014)		(0.064)		(13.460)		(39.331)	
log(bid)	-0.920	***	-0.903	***				
	(0.016)		(0.022)					
Distribution	Log-logistic		Log-logistic		Normal		Normal	
Data on ∆RISK used	(2,3,5,6,7)		(2,3)		(2,3,5,6,7)		(2,3)	
Log-likelihood	-8 880.04		-4 853.73		-9 739.86		-5 268.46	
Observations	6 940		3 752		6 940		3 752	
Excluded	speeders, failed		speeders, failed		speeders, failed		speeders, failed	
Pair-wise test of differences between the coefficients								
$\Delta RISK \times SQ6 = \Delta RISK \times SQ10$	χ^2 = 0.2870 (p = 0.592)		χ^2 = 0.0274 (p = 0.868)		χ^2 = 0.2870 (p = 0.592)		χ^2 = 0.0274 (p = 0.868)	
$\Delta RISK \times SQ6 = \Delta RISK \times SQ12$	χ^2 = 0.1357 (p = 0.713)		χ^2 = 0.0703 (p = 0.791)		χ^2 = 0.1357 (p = 0.713)		χ^2 = 0.0703 (p = 0.791)	
$\Delta RISK \times SQ10 = \Delta RISK \times SQ12$	χ^2 = 4.1313 (p = 0.042)*		χ^2 = 0.1878 (p = 0.665)		χ^2 = 4.1313 (p = 0.042)*		χ^2 = 0.1878 (p = 0.665)	

Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses. Models (A) and (B) use responses on all choice questions, while models (A23) and (B23) use responses on choice questions with the risk reduction by 2 or 3 in 1,000 – that are the same risk reduction levels used with all three baseline risk levels. The terms SQ6, SQ10, and SQ12 are dummies equal to one if the baseline risk level was 6, 10 and 12, respectively, in 1000, and zero otherwise.

Table C.10. WTP model that controls for other effects, interaction terms with the change in the risk reduction

	(1)		(2)	
Intercept	6.808	***	6.806	***
	(0.131)		(0.131)	
$\Delta RISK$	0.022	·	0.025	.
	(0.013)		(0.013)	
$\Delta RISK \times$ increase			0.138	***
			(0.012)	
$\Delta RISK \times$ decrease			-0.011	
			(0.021)	
$\Delta RISK \times$ environment	0.034	*		
	(0.017)			
$\Delta RISK \times$ own health	0.034	*		
	(0.015)			
$\Delta RISK \times$ health family	0.056	***		
	(0.016)			
$\Delta RISK \times$ economy	-0.012			
	(0.019)			
$\Delta RISK \times$ miscarriage	0.006			
	(0.016)			
$\Delta RISK \times$ infant death	0.067	***		
	(0.016)			
log(bid)	-0.930	***	-0.931	***
	(0.017)		(0.017)	
Distribution	Log-logistic		Log-logistic	
Log-likelihood	-8 817.33		-8 809.46	
AIC	17 652.65		17 628.92	
BIC	17 714.26		17 663.15	
Observations	6 940		6 940	
Excluded	speeders, failed		speeders, failed	

Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in parentheses.

Table C.11. Single-bounded dichotomous choice model, bids in USD PPP

Panel A – logit model

	(1)		(2)	
$\Delta RISK$	1.801E-01	***	5.875E-02	***
	(9.393E-03)		(1.466E-02)	
bid1	-2.516E-04	***	-5.171E-04	***
	(3.620E-05)		(4.360E-05)	
Canada			0.478	***
			(0.102)	
Czech Republic			0.898	***
			(0.101)	
Switzerland			0.600	***
			(0.104)	
Italy			1.374	***
			(0.108)	
Mexico			1.357	***
			(0.111)	
Netherlands			0.660	***
			(0.104)	
Türkiye			1.079	***
			(0.111)	
United Kingdom			0.204	*
			(0.099)	
United States			1.002	***
			(0.106)	
Log-likelihood	-4 520.88		-4 346.72	
Observations	6 940		6 940	
screening	speeders quiz failed		speeders quiz failed	
VSC of very low birth weight mean,	715 865		612 848	

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors are given in brackets. The VSC mean is estimated from the estimated coefficients for model (1). The mean VSC is obtained as the mean of individual WTP.

Panel B – VSC estimates by country in USD PPP

	Mean	Standard deviation	Median
Canada	394 772	123 567	421 893
Czech Rep	650 781	237 448	692 558
Italy	928 753	363 962	999 644
Mexico	910 079	358 182	988 136
Netherlands	504 060	173 925	539 029
Switzerland	459 425	155 018	500 133
Türkiye	765 403	285 020	808 954
United Kingdom	234 451	53 540	244 973
United States	706 903	267 522	759 693
ALL	612 848	333 077	530 823

Table C.12. Robustness check, WTP models assuming with different distributions, speeders and quiz failures excluded

USD PPP	Weibull Spike						Log-logistic					
	mean			median			mean			median		
	individual truncated at max bid	individual truncated adjusted	median WTP	individual truncated at max bid	individual truncated adjusted	median WTP	individual truncated at max bid	individual truncated adjusted	median WTP	individual truncated at max bid	individual truncated adjusted	median WTP
not weighted												
Canada	723	845	424	783	909	453	729	881	388	791	951	415
Czech Republic	822	1 004	524	878	1 065	553	867	1 112	542	927	1 183	572
Italy	1 145	1 768	1 109	1 237	1 891	1 177	1 141	1 743	1 049	1 232	1 867	1 115
Mexico	1 044	1 497	889	1 138	1 614	951	1 053	1 521	866	1 147	1 643	927
Netherlands	805	979	509	868	1 048	542	811	1 016	477	875	1 091	508
Switzerland	843	1 062	569	931	1 161	617	857	1 110	551	947	1 218	598
Türkiye	1 014	1 390	797	1 076	1 463	834	951	1 268	655	1 008	1 337	686
United Kingdom	649	731	354	697	780	376	680	804	337	732	862	358
United States	1 031	1 453	852	1 115	1 557	905	996	1380	749	1 076	1 480	796
Pooled	894	1 186	666	783	1 065	574	896	1202	623	791	1 081	566
weighted												
Canada	715	836	420	783	909	453	721	872	384	791	951	415
Czech Republic	828	1 011	528	878	1 065	553	874	1 120	546	927	1 183	572
Italy	1 145	1 768	1 109	1 237	1 891	1 177	1 141	1 744	1 050	1 232	1 867	1 115
Mexico	1 049	1 503	892	1 138	1 614	951	1 057	1 528	869	1 147	1 643	927
Netherlands	802	976	508	868	1 048	542	809	1 013	476	875	1 091	508
Switzerland	840	1 059	567	931	1 161	617	854	1 107	549	947	1 218	598
Türkiye	1 020	1 397	801	1 076	1 463	834	956	1 274	658	1 008	1 337	686
United Kingdom	650	732	355	697	780	376	681	805	338	732	862	358
United States	1 038	1 463	857	1 115	1 557	905	1 003	1 389	753	1 076	1 480	796
Pooled	895	1 188	667	783	1 065	574	895	1 194	614	840	1 101	561

Annex D. Core principles of survey analysis

Detect potentially problematic Detect potentially problematic responses

1. Generate a dummy variable for people failing the probability test
2. Speeder management: Generate one dummy variable for *survey* speeders and one dummy for *valuation* speeder. A respondent taking less than 48% of the median time is a speeder (ISS definition). Median values should be country specific to account for difference in languages that impact reading time.
3. Generate two dummies variable for distracted respondents: respondents who took an abnormally long time to respond:
   a. 48% longer than the median *survey* time,
   b. 48% longer than the median *valuation* time.
4. Optional. Generate a dummy variable for straightliners: when survey respondents give identical (or nearly identical) answers to items in a battery of questions using the same response scale. Note that there should not be any of them in the data sent by the internet panel provider.
5. Optional. Generate a dummy variable for respondents having incoherent answers:
   a. E.g., mismatch between the number of children, number of people in the household, or year of youngest child
6. Generate a dummy variable for unrealistic max WTP in open-ended question
7. Generate a dummy variable for probability test failers
8. Generate a dummy variable for protesters. This varies between endpoints. For example, in the asthma survey, people who disagree with the description of asthma provided in the survey or who are very doubtful that the information provided by the survey is correct or who thought they could just lower consumption of cleaning products can be considered as protesters.
9. Generate a dummy variable for respondents stating high co-benefits
10. Generate a dummy variable for consequentiality (real life debrief)
11. Optional. Read written responses to open ended questions to detect potentially problematic responses
12. Optional. Compute number of problematic responses to debriefing:
    a. that could overestimate WTP
    b. that could underestimate WTP
    c. that could go in either direction or a non-directional

Screen out problematic responses

- Baseline:

  - Exclude survey and valuation speeder (reinforced compared to Ipsos)
  - Exclude straightliners (already done by Ipsos)
  - Exclude respondents who fail the probability test (not applicable for IQ loss)
  - Keep pilot respondents if the survey design is the same even if parameters (such as bid levels) changed except if the changes are significant
  - Keep co-benefiters
  - Keep protesters to have a conservative estimate
  - Keep distracted respondents
  - Variations to perform as robustness checks:
- Optional robustness: stricter screening
  - Exclude survey and valuation speeder (same as option A)
  - Exclude straightliners (same as option A)
  - Exclude respondents who fail the probability test (same as option A)
  - Keep pilot respondents if the survey design is the same even if parameters (such as bid levels) changed (same as option A)
  - Keep co-benefiters (same as option A)
  - Exclude protesters because no does not mean true zero
  - Exclude distracted respondents
  - Exclude pilot respondents if pilot parameters differ too much (case of very low birth weight)
- Optional: exclude respondents that took more than 12h to complete the survey

Provide information on the sample of respondents

1. Compute summary statistics to describe the screened sample
   - Put main descriptive in body of text
   - And other e.g., country level in the appendix
2. Check that achieved quotas (age, education, location, gender) and income distribution in the screened sample are consistent with available population statistics (target quotas) at the country level (from OECD.Stat and Eurostat).
3. For each country separately, compute post-stratification weights to reweight later the observations through an iterative proportional fitting procedure (raking algorithm) using the following strata:
   - Gender × Age: (1) males aged 18-24; (2) males aged 25-34; (3) males aged 35-39; (4) males aged 40-44; (5) males aged 45-65; (6) females aged 18-24; (7) females aged 25-34; (8) females aged 35-39; (9) females aged 40-44.
   - Educational level: (1) low, (2) medium, and (3) high
   - Geographic region: country-specific NUTS 2 regions

It is important to consider the efficiency of the weights, such that ideally the overall weighting efficiency remains above a certain value to avoid any significant impact on the effective sample sizes obtained and, consequently, on the statistical power of the analyses conducted. Weighting efficiency can be further improved by collapsing weighting cells and capping weights at each of the steps to reduce the impact on the variance of the final weights. At the end of each iteration of the algorithm, any weights larger than 3.0 or lower than 1/3 should be automatically set to equal this cap.

Analyse responses to the valuation questions after baseline screening

1. Compute the DBDC response matrix for both the pooled dataset and each country of the dataset
2. Scope analysis:
    - Verify that the share of yes response decreases with the cost to be paid
    - Verify that the share of yes response increases with the risk reduction offered
3. Analyse written (open-ended) questions:
    - Use examples to illustrate the thinking of respondents if they were asked why they made their choice
    - Optional. Check consistency between OE and DBDC responses
4. As a preliminary step, regress SBDC (response to first dichotomous choice) on income, bid amount, baseline risk (if relevant) and risk reduction using a logit model
5. Optional. Try to find determinants of no-no and yes-yes responses using responses to debriefing questions

Compute harmonised variables

1. Compute continuous income level in USD PPP[13] based on unequivalised income range selected by the respondents:
    - Average of each interval
    - 0.5 lowest interval and 1.5 highest interval
2. Predict missing income using the following strategy
    - Generate the following dummies
        - Missing income dummy equal to 1 if the respondent did not provide income information
        - Couple dummy equals 1 if the respondent is married or have a partner
        - Employed dummy equals 1 if the respondent is in one of the following situations:
            - employed full-time
            - self employed
            - military
            - Own business manager
        - Part time dummy equals 1 if the respondent is employed part time
        - Retired dummy equals 1 if the respondent is retired

[13] This is OECD standard. PPS is the technical term used by Eurostat for the common currency in which national accounts aggregates are expressed when adjusted for price level differences using PPPs. Thus, PPPs can be interpreted as the exchange rate of the PPS against the euro.

  - Replace employed and part time dummies by 0 if they are missing
  - Replace retired dummy by 1 if it is missing and the person is aged 60 or more or by 0 if it is missing and the person is younger than 60 years old.
- For each surveyed country separately, run the OLS regression of log(income) on age dummies, high education dummy, female dummy, couple dummy, number of persons in the household, employed dummy, part time dummy and retired dummy. For surveys targeting couples planning to have children, do not include couple dummy nor retired dummy that are naturally omitted since perfectly colinear.
- Predict income based on the regressions
- Replace missing income with predicted value in the main dataset

3. Compute a variable for age:
   - Option A (preferred). One dummy variable for each category ➔ better identification
   - Option B. Continuous age as for income ➔ preserves statistical power
     - 18-26 ➔22
     - 27-34 ➔ 30.5
     - 35-39 ➔37
     - 40-44 ➔ 42
     - 45-65 ➔ 55
     - 65+ ➔ 70
4. Compute a variable for education using Ipsos's low, medium and high category (directly available)
5. For all countries except the United States, compute bid level in USD PPP equivalent using OECD data on PPP for actual individual consumption. Because of rounding after currency conversion, respondents in non-US countries had bid levels that are slightly different than the bid levels seen by US respondents. Reconverting actual bid levels to USD PPP equivalent allows to obtain a more precise bid amount.

Apply a standard specification

1. Baseline:
   - All surveys: intercept, female, age, kids02, category dummies, log(income), missing income dummy, low, medium, high education dummies, baseline risk (if relevant), risk reduction
   - Add country dummies interacted by the post stratification weights to account for the difference between target and achieved sample quotas. This is similar to—albeit less complex than—the correction method for choice-based samples proposed by Manski and Lerman (1977[52]). Do not add country dummies to these interactions to avoid multi collinearity.
   - Add the number of children for fertility loss and very low birth weight
2. Robustness checks:
   - Health augmentation: own health perception, know someone having the condition, lifestyle, covid
   - Run the estimation without the missing income dummy.

Estimate average and median WTP based on DBDC

1. Estimator: DBDC or SBDC:
   - Baseline: interval-data maximum likelihood estimator using DBDC
   - Robustness check: Estimate WTP based on SB choice with logit model to compare to DB estimate
2. Distribution of the error:
   - Baseline (preferred to allow comparison across endpoints): Weibull. The Weibull distribution has desirable characteristics. Specifically, this specification offers a flexible survival function which mimics other distributional forms quite well, and thanks to its shorter right tail it typically performs better than the log-normal distribution (Carson and Hanemann, 2005[44]).
   - Robustness checks:
     - Non-parametric: Turnbull (e.g., Kaplan-Meier)
     - Basic parametric: normal, log normal, logistic, log logistic
     - Identify estimator with the lowest Akaike information criterion ($AIC = 2k - 2\ln \hat{L}$)
3. Spike configuration:
   - Baseline: use spike configuration (Kriström, 1997[53]; Carson and Hanemann, 2005[44]) if the spike variable is higher or equal to 5%. In other words, use spike when the average probability that people are indifferent to the valued item is higher or equal to 5%. Spike configuration can still be used if spike is lower than 5% but close to it. Spike is less likely to be relevant when people that have a priori no preference for the good are screened out by design. This is the case of the infertility and very low birth weight where only people planning to have a child over the next years were able to respond to the survey.
   - Robustness check: Compare estimates using spike and without using spike.
4. Compute WTP and VSC on pooled dataset based on a simple model with constant, country dummies interacted with weights and risk reduction as the only covariates using the following formulas:
   - Baseline: $\widehat{VSC} = \frac{1}{n}\sum_i \widehat{VSC}_i$ where $\widehat{VSC}_i = \widehat{WTP}_i/RR_i$ and $\widehat{WTP}_i$ is the individual mean WTP (truncated at the maximum bid with adjustment)
   - Robustness check (optional): Compute average WTP at sample mean: $\overline{WTP} = \widehat{b_0} + \widehat{b_1}\overline{RR} \rightarrow \widehat{VSC} = \overline{WTP}/\overline{RR}$
5. Compute WTP and VSC for each country based on the ***pooled*** regression estimated above. Do not use separate country-level regressions to generate country-level WTP and VSC as indicated in the previous version. Using the pooled model allows to capture the “cultural” differences between the countries (by also taking into account the fact that the sample is not perfectly representing the population in the country), by multiplying the country dummies with the weights, and using this as a coefficient to predict the values in each country. The pooled approach also increases dramatically the statistical power.
6. Perform the estimation using the standard specification defined above to test determinants of WTP:
   - Assess scope sensitivity:

  - Inference of the risk reduction coefficient
  - Optional. Estimate WTP for different risk reduction separately
- Estimate income elasticity by simulating an increase in income by 1% for all respondents.
  - Increase income of all respondents by 1% before computing individual WTP. This relies on the same estimates derived from original data.
  - Compute the new mean of the individual mean WTP (truncated at the maximum bid with adjustment)
  - The elasticity is equal to this % change between this new mean and the baseline mean WTP.
- Other effects using the regressors of the specification: age, gender, etc.

Derive central value and range of VSC for pooled dataset and each country

1. Estimate central value (mean VSC) using the baseline approach. The central value should be clearly identified for regulators to choose.
2. Clearly present country-specific values as recommended values because they can be directly use in cost benefit analyses.
3. Provide pooled (all countries) mean VSC for information.
4. Provide pooled and country specific median WTP and VSC in the appendix
5. Provide an example of how the VSC can be used in CBA.
6. Compare WTP and VSC with magnitude of available WTP, QALY and Cost-of-Illness estimates from the literature for similar endpoints.

Prepare and share your code

1. Baseline: Prepare your code in R because it is free and more flexible (see dbchoice and dbspike packages). In contrast, only interval data ML estimators based on normal distribution are directly available for Stata (intreg, doubleb). In the long run, it is planned to make the code of the working paper publicly available.
2. Comment your code sufficiently so that a third person can run your code from scratch.
3. Share your code in shared folders.

Valuing a reduction in the risk of very low birth weight

There is ample evidence that exposure to various chemicals can increase the probability of children to be born with low or very low birth weight. Infants born with very low birth weight have a higher risk of suffering from neurosensory problems, issues related to behavioural and social competencies, and learning disabilities than infants born with normal birth weight. Authorities face challenges in regulating chemical substances through actions such as bans and prohibitions, because of the difficulty in explicitly considering the economic benefits and costs of such regulations. Moreover, existing Values of a Statistical Case (VSC) of very low birth weight are rare and cannot be directly applied to the cost benefit analysis of chemical management options for a wide range of countries.

This paper is part of the series of large scale WTP studies resulting from the Surveys to elicit Willingness to pay to Avoid Chemicals related negative Health Effects (SWACHE) project that intends to improve the basis for doing cost benefit analyses of chemicals management options and environmental policies in general. The present paper details a stated preference survey estimating WTP to reduce the risk of very low birth weight, filling an important gap in the valuation literature and addressing a need for applied benefits analysis for chemicals regulation. The SWACHE very low birth weight survey was fielded in nine countries: Canada, the Czech Republic, Italy, Mexico, the Netherlands, Switzerland, Türkiye, the United Kingdom, and the United States.

In each country, a sample of 1 200 adults, representative of the general population and of childbearing age who are in a relationship and plan for a(nother) child within the next five years, was collected and empirically analysed.

The mean Value of a Statistical Case (VSC) of very low birth weight estimated in this study equals USD_{2022} Purchasing Power Parity (PPP) 1 194 000 and the median VSC equals USD_{2022} PPP 614 000. Country-specific mean VSC of very low birth weight vary between USD_{2022} PPP 805 000 for the United Kingdom and USD_{2022} PPP 1 744 000 for Italy.

Recommended citation: Ščasný, M.,I. Zvěřinová and D. Dussaux (2023), *"Valuing a reduction in the risk of very low birth weight: A large scale multi-country stated preference approach"*, OECD Environment Working Papers, No. 217, OECD Publishing, Paris, https://doi.org/10.1787/dfd159a1-en.

The OECD SWACHE project has received the financial assistance of the European Union. The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

For more information:

https://oe.cd/SWACHE

@OECD_ENV

OECD Environment

OECD
BETTER POLICIES FOR BETTER LIVES