Measuring the Uncertainty of Environmental Good Preferences with Bayesian Deep Learning

Ricardo Flores
rflores@wpi.edu
Worcester Polytechnic Institute
Worcester, MA, USA

ML Tlachac
mltlachac@wpi.edu
Bryant University
Smithfield, RI, USA

Elke A. Rundensteiner
rundenst@wpi.edu
Worcester Polytechnic Institute
Worcester, MA, USA

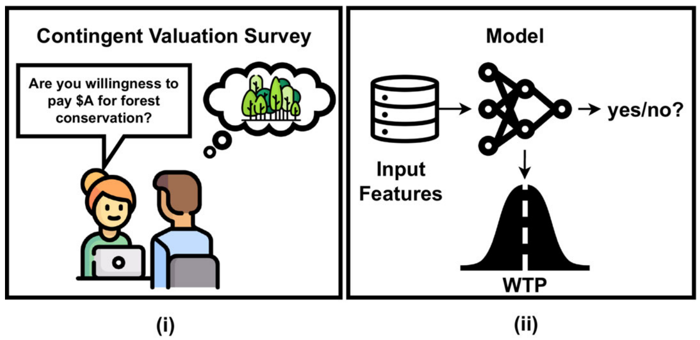

Figure 1: (i) Contingent valuation survey to capture the preference for social goods. For example, "What is it the willingness to pay (WTP) for forest conservation?" (ii) Features from the social good products and/or the population are input for the classification model that will then estimate the distribution of the WTP for a particular social good, such as, forest conservation.

ABSTRACT

Due to climate change and resulting natural disasters, there has been a growing interest in measuring the value of social goods to our society, like environmental conservation. Traditionally, the stated preference, such as contingent valuation, captures an economics-perspective on the value of environmental goods through the willingness to pay (WTP) paradigm. Where the economics theory to estimate the WTP using machine learning is the random utility model. However, the estimation of WTP depends on rather simple preference assumptions based on a linear functional form. These models are therefore unable to capture the complex uncertainty in the human decision-making process. Further, contingent valuation only uses the mean or median estimation of WTP. Yet it has been recognized that other quantiles of the WTP would be valuable to ensure the provision of social goods. In this work, we propose to leverage the Bayesian Deep Learning (BDL) models to capture the uncertainty in stated preference estimation. We focus on the probability of paying for an environmental good and the conditional distribution of WTP. The Bayesian deep learning model connects with the economics theory of the random utility model through the stochastic component on the individual preferences. For testing our proposed model, we work with both synthetic and real world data. The results on synthetic data suggest the BDL can capture the uncertainty consistently with different distribution of WTP. For the real world data, a forest conservation contingent valuation survey, we observed a high variability in the distribution of the WTP, suggesting high uncertainty in the individual preferences for social goods. Our research can be used to inform environmental policy, including the preservation of natural resources and other social good.

CCS CONCEPTS

• **Computing methodologies** → *Neural networks*; **Bayesian network models**; *Supervised learning by classification*; • **Applied computing** → *Economics*; Environmental sciences.

KEYWORDS

Contingent valuation, Environmental conservation, Random utility model, Sustainability

ACM Reference Format:
Ricardo Flores, ML Tlachac, and Elke A. Rundensteiner. 2022. Measuring the Uncertainty of Environmental Good Preferences with Bayesian Deep Learning. In *GoodIT '22: ACM International Conference on Information Technology for Social Good, Sept. 07–09, 2022, Limassol, Cyprus.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3524458.3547250

GoodIT '22, Sept. 07–09, 2022, Limassol, Cyprus

ACM ISBN 978-1-4503-9284-6/22/09…$15.00
https://doi.org/10.1145/3524458.3547250

1 INTRODUCTION

Background. Climate change has become a major focus for international institutions [39], due in part to extreme weather events that cause damage to natural resources [12, 37]. Thus, obtaining an economics value of these social goods is relevant to policy-makers for quantifying the damage. Contingent Valuation (CV), in economics, is a stated preference method, which allows to estimate the value of this non-market good in different environmental areas [3, 11, 32].

Contingent valuation focuses on modeling willingness to pay (WTP) of an environmental good by creating a hypothetical market in which people can declare their preferences, as depicted in Fig. 1 (i). The Random Utility Model (RUM) is one of the economics models utilized to obtain the WTP. Hanemann [22] demonstrated the relationship between the responses to a referendum contingent valuation study and the random utility model suggested by Lancaster [28]. However, the estimation of WTP depends on the assumption that economists make about the preferences of individuals using RUM. Namely, a functional form for capturing the utility function has to be chosen, in most cases assumed linear. Also, in most cases just the mean or median of the WTP is reported. The results of CV studies are therefore not available to capture the uncertainty of individual preferences.

In contrast, Deep Learning (DL) models, which have become increasingly popular and effective for modeling complex problems in recent years are able to capture complex non-linear relationships [19]. They promise to potentially improve the binary classification problem. However, DL maps input and output with a deterministic function, without consider the natural uncertainty models and datasets [17]. Accordingly, some authors suggest [24, 31, 43] leveraging Bayesian analysis for social problems. Bayesian Deep Learning (BDL) is a framework that incorporates the uncertainty in the network structures of the deep learning model architecture by the assumption of randomness in their parameters [20].

Motivating Example. Forest conservation is an example of an important social good. To obtain an economic value for the conservation of a particular forest, we can create an artificial market. This then allows us to capture the preferences of people as well as their willingness to pay for its conservation. Contingent valuation is a methodology that aims to create this hypothetical market by means of conducting a survey, as depicted in Fig. 1 (i). In this survey, the interviewer starts describing in detail the benefits of forest conservation. She then ends with the question "Are you willing to pay A, where A is an amount of money for forest conservation?". The interviewee has to answer "yes" or "no" in a closed-ended format survey. This style of survey is the commonly accepted format used in contingent valuation surveys [22].

After the data collection, the characteristics from interviewee (like age, gender, income) and the forest (like type forest, retention and storage of water in aquifers) are stored to use as input for classification model. Once the model is trained on the binary classification task, an estimation can be obtained of the preference of people through the distribution of willingness to pay for forest conservation. This modeling process is depicted in Fig. 1 (ii).

Problem Definition & Technical Challenges. Given a set of features, our goal is to train a classifier and then infer the distribution of willingness to pay for environmental goods. The data consist of interviewee and social good characteristics, bid (amount of money A), and label with yes/no answers to "Are you willing to pay A for a social good?"

Measuring the distribution of willingness to pay for an environmental good is challenging in the context of the contingent valuation method. This is because it is a non-observable variable that depends on the individual preferences with a stochastic component [41]. Typically, contingent valuation methods make a simple assumption about the preferences using RUM model. Additionally, they simply report the mean or median of the willingness to pay distribution. That is, they are not given any extra information about the distribution of willingness to pay. We note that Carson and Hanemann [9] argued that lower quantiles of the willingness to pay distribution should be used by policy-makers, instead of mean or median, because it implies higher political and social feasibility environmental good provision [42].

Our Approach and Contributions. In this research, we explore the distribution of willingness to pay for environmental goods in the context of the contingent valuation method. Previous contingent valuation studies [21] rely on the simple assumption of individual preferences. Hence, they simply report the mean or median of the willingness to pay distribution for an environmental good, which does however not reveal the uncertainty on the WTP. In this work, we thus propose a methodology that leverages the Bayesian deep learning models for contingent evaluation. Specifically, we connect the economics framework of the random utility model with the Bayesian deep learning, keeping the randomness in modeling individual preferences, but while not assuming any specific functional form for the utility function in the RUM model. Once the Bayesian deep learning model is trained, we can easily obtain the distribution of willingness to pay for both synthetic data and a forest conservation contingent valuation survey. To the best of our knowledge, we are the first to apply Bayesian deep learning to the contingent valuation method for assessing the socially perceived value of environmental goods. Our contributions include:

(1) An approach to use Bayesian deep learning models with the contingent valuation method.
(2) A measurement of the uncertainty of the probability of paying for an environmental good.
(3) An estimation of the empirical distribution of willingness to pay (WTP) for environmental goods.
(4) Evaluation of our approach with synthetic, and real datasets in forest conservation.

2 RELATED WORK

The contingent valuation method has been applied to different areas of economics with the main results summarized in books covering theoretical and empirical issues [3, 11, 32]. Applications of contingent valuation (CV) to specific fields include healthcare [27], education [15], traffic noise [5, 8], and ecosystem services [30].

Next, we discuss related work in the area of using machine learning in the context of CV. Zhao et al. [44] compare different machine learning models, such as Naive Bayes (NB), Support Vector Machine (SVM), and Tree-based Models, with logistic regression traditionally used in the contingent valuation to model the stated preferences for mobility-on-demand transit. Bravo et al. [7] use a feed forward neural network to compute the WTP to reduce road noise annoyance. However, they do not leverage the random utility model, instead, they use the WTP reported by people as target variable, which is not recommend according to NOAA panel for the contingent valuation method [2].

A disadvantage of deep learning models is that they use a deterministic function to map the input to the output. Therefore, uncertainty is often not modeled during the data representation learning process [17]. In contrast, Lavin et al. [29] implement a quantile regression for learning the payments for environmental services. This paper models an approximation of the quantile of the willingness to pay (WTP), although it used the traditional assumption of the contingent valuation method, as a linear utility function. Unlike prior work, our research leverages Bayesian deep learning without assuming any particular distribution on the utility function in order to apply CV to forest conservation.

3 ECONOMICS: RANDOM UTILITY MODEL

As Haab and McConnell [21] described, a contingent value model assumes that the satisfaction that a consumer perceives can be represented by a utility function (u_j) which has a deterministic component and a random component. That is,

$$u_j = v_j(p, I, q_j) + \varepsilon_j, \tag{1}$$

where p denotes a vector of current prices, I is the income, and q_j represents the environmental quality at time j. Also, $j = 0$ denotes the initial situation, while $j = 1$ the new situation. A respondent will be willing to pay the amount $\$A$ only if the utility of paying for this project is higher than the utility of the status quo in which he or she does not pay for the project [3] ($u_1 > u_0$). We denote a positive answer of individual i as $y_i = 1$ and a negative answer as $y_i = 0$. Therefore, the probability (Pr) of an answer being positive (yes) is modeled by:

$$Pr(y_i = 1) = Pr(\Delta v > \eta) = F_\eta(\Delta v), \tag{2}$$

with $\Delta v = v_1(p, I - A, q_1) - v_0(p, I, q_0)$, $\eta = \varepsilon_0 - \varepsilon_1$ and F_η as the distribution function of η. Hanemann [22] noted we need an assumption over the η, In the conventional RUM approach, η follows either a logistic or normal distribution. There are several options for selecting Δv, but most researchers working with the contingent valuation method use a linear utility function. Therefore, assuming $\Delta v = \alpha - \beta A$, $\beta > 0$, the median of willingness to pay (WTP) is:

$$\begin{aligned} 0.5 &= F_\eta(\Delta v), \\ 0.5 &= F_\eta(\alpha - \beta A), \\ \alpha - \beta A &= F_\eta^{-1}(0.5), \\ A &= \frac{\alpha - F_\eta^{-1}(0.5)}{\beta}, \end{aligned} \tag{3}$$

Then, if $F_\eta(\cdot)$ is assumed logistic or normal, we have:

$$WTP = A = \frac{\alpha}{\beta}. \tag{4}$$

In economics, WTP represents the compensation variation, which is the amount of money that someone is indifferent about paying or not paying for an environmental good [21]. The assumption of a linear utility function restricts the possibility of finding realistic preferences. Also, most people in contingent valuation just report the mean or median of WTP instead of its distribution, therefore, it is not possible to measure the uncertainty in this scenario. The estimation of WTP will depend on the assumption about the utility form. For this reason, we propose a more generic approach; we do not impose any assumption on the utility function, but continue to model the stochastic component in the RUM. In other words, we use the random utility model without any restrictive assumption about the deterministic component.

4 BAYESIAN DEEP LEARNING

Our goal is to estimate the probability of paying for an environmental good, $Pr(y_i = 1) = F_\eta(\Delta v)$. A Bayesian deep learning model can be considered to correspond to a neural network with a prior distribution on its weights [16, 18, 33, 35] as depicted in Fig. 2 (right). Following Blundell et al. [6], to estimate the posterior distribution of the weights $Pr(w|D)$, where D represents a set of training samples (x_i, y_i), Variational learning is used to find the parameters θ of a distribution on the weights $q(w|\theta)$ that minimizes the Kullback-Leibler (KL) divergence with the distribution $Pr(w|D)$. The cost function, also called ELBO (Evidence Lower Bound) for Kingma and Welling [26] is:

$$F(D, \theta) = KL[q(w|\theta) \| Pr(w)] - E_{q(w|\theta)}[\log Pr(D|w)]. \tag{5}$$

Then, a deep neural network is trained by assuming an initial distribution of w, usually Gaussian, as in Fig. 2 (right). Our task is to classify if someone agrees or disagrees to pay for an environmental good. Therefore, we set the Bernoulli distribution, which is a discrete probability distribution, with values 1 and 0. Once the model is trained, we can capture the uncertainty of the preference through sampling from the distribution of $Pr(y_i = 1)$. This would provide us with an approximation of the distribution of willingness to pay (WTP) using the quantile of $F_\eta(\Delta v)$ conditional to A,

$$WTP_A = F_\eta^{-1}(\tau), \tau \in (0, 1). \tag{6}$$

Therefore, if we want to get the median of WTP, we set $\tau = 0.5$. Repeating this process for all sampled values of $F_\eta^{-1}(.)$, we can get an approximation of the distribution of WTP as

$$f(WTP) = F_\eta^{-1}(.5)_{sampling}. \tag{7}$$

For the implementation, we use a fully connected layer with three hidden layers, setting the distribution of the weights as Gaussian with a mean of 0 and a variance of 1. Automatic differentiation variational inference (ADVI) is used in mini-batch to improve algorithm speed. Experiments were run using the Python probabilistic programming framework (PyMC3) [38]. For both synthetic and real data, we use 70% of data for training and the remaining 30% for testing. During testing, we evaluate the classification problem using *accuracy*, which is the ratio of correctly classified instances. We also report on the mean WTP and the Bayesian credible interval with a 95% credibility.

5 SYNTHETIC DATASETS

We use two synthetic datasets to explore the effectiveness and limitations of the proposed Bayesian deep learning method. The first experiment follows the traditional assumption of the contingent valuation method. Thus, we expect a perfect Gaussian distribution of the WTP. In the second experiment, we simulate a right-skewed distribution of WTP. For this, we expect the Bayesian deep learning will succeed to capture this distortion in the WTP.

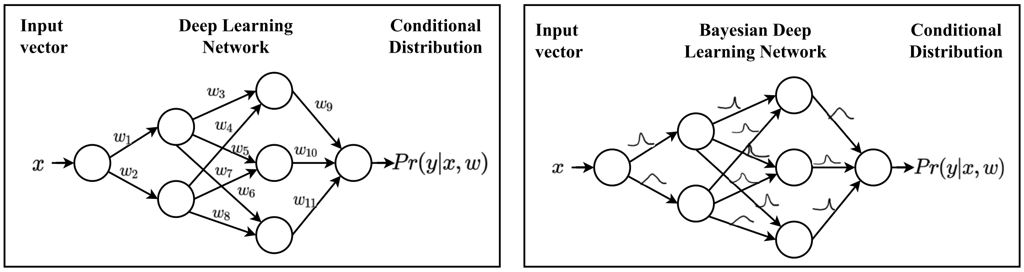

Figure 2: Representation of DL Network (left) and BDL Network (right), where x is an input vector. The weights w are assumed fixed for DL and follow a prior distribution for BDL. Both networks output the conditional distribution of the label y.

5.1 Experiment 1: Gaussian Distribution for Willingness to Pay

In this experiment, we create synthetic datasets following the traditional assumption of the contingent valuation method. Namely, we set a linear utility function as $\Delta v = \alpha - \beta A$, which parameters α and β values are 5 and 0.5 respectively and A a vector of different bids (amount of money someone agrees/disagrese to pay), then Δv is defined as:

$$\Delta v = 5 - 0.5A + \eta. \tag{8}$$

Above, we include the stochastic component η, which is a Gaussian with mean 0 and variance 1 as $\eta \sim N(0, 1)$. Then, the bid A follows a uniform distribution as $A \sim Unif(1, 20)$, where 1 is the minimum value and 20 the maximum value. Because Δv is an unobserved variable, we create the binary label y_i with a sample size of 1000, as follows:

$$y_i = \begin{cases} 1, & \text{if } \Delta v \geq 0, \\ 0, & \text{if } \Delta v < 0. \end{cases} \tag{9}$$

We expect both traditional machine learning models, like logistic regression, and Bayesian deep learning to exhibit a good performance in terms of accuracy in this particular data scenario. The true value WTP is 10 as computed in Equation 4 or 6. However, we note that logistic regression cannot output a representation of the distribution of willingness to pay. This is an unfortunately shortcoming, because this representation could have been leveraged for policy-makers to measure the uncertainty of WTP.

The results of the probability of paying for an environmental good and the distribution of willingness to pay are in Fig.3. As we expected, both plots depict the typical Gaussian property. That is, there is a negative relationship between the bid A and the probability of paying, put differently, the probability of paying decreases if the bid A increase. The distribution of willingness to pay shows a perfect Gaussian distribution, centered around WTP = 10.

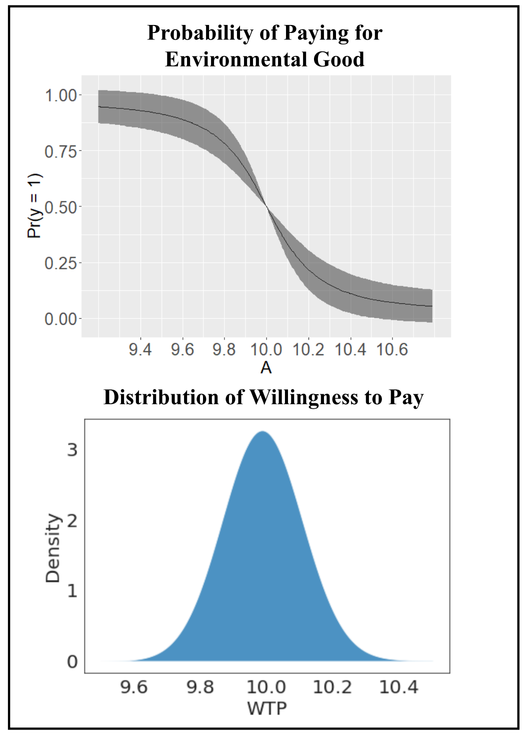

Figure 3: Summary of results of the probability of paying for environmental good and distribution of WTP for the data in experiment 1, where we simulated a Gaussian distribution of WTP using the Bayesian deep learning model.

Table 1: Results of applying Bayesian deep learning and logistic regression to data set from experiment1. We include the mean accuracy, mean WTP, as well as the Bayesian credible interval with a 95% credibility.

Model	Accuracy	WTP (mean)	[Q025, Q975]
BDL	0.93	9.99	[9.98, 10.08]
Logistic	0.91	9.56	**NA**

The results of the Bayesian deep learning (BDL) model and the logistic regression are provided in Table 1. Because of the Gaussian assumption, BDL and the logistic model have similar accuracy, WTP, and Bayesian credible interval. This last one tells us that there is a 95% probability that the WTP would lie withing 9.98 and 10.08 given our datasets. These results suggest there is a low uncertainty in the individual preferences of the surveyed citizens.

5.2 Experiment 2: Right-Skewed Distribution for Willingness to Pay

In this experiment we study a right-skewed (positive skewness) distribution of willingness to pay. In this, we expect the Bayesian deep learning model will be able to capture the distortion in the distribution of WTP, reflecting an hypothetical case with a high uncertainty in the preferences of environmental goods. To do so, we simulate three different utility functions, each reflecting different levels of willingness to pay for an environmental good. They are:

$$\begin{aligned} \Delta v_1 &= 0.5 - 0.05A_1 + \eta, \\ \Delta v_2 &= 0.5 - 0.03A_2 + \eta, \\ \Delta v_3 &= 0.5 - 0.02A_3 + \eta, \\ \Delta v &= (\Delta v_1 \circ \Delta v_2 \circ \Delta v_3). \end{aligned} \tag{10}$$

As for the Gaussian distribution used in experiment 1, we include the stochastic component η, with mean 0 and a variance 1. However, now we also add some variations in the bid A with a uniform distribution of $A_1 \sim Unif(1, 20)$, $A_2 \sim Unif(20, 40)$, and $A_3 \sim Unif(40, 60)$. The β component of the utility function, $\Delta v = \alpha - \beta A$, takes on different values, namely, $0.05, 0.03$, and 0.02. These generate differences on the willingness to pay values, resulting in a high uncertainty supported in the individual preferences. The parameter α is set to 0.5, which represent a constant component between these three groups. We simulate Δv_1 with a sample size of 600 and Δv_2 and Δv_3 with sample of 200 each, because some groups of people will pay less than others. Further, those with the high willingness to pay tend to be the relatively smallest group. Then, Δv is a concatenation of Δv_1, Δv_2, and Δv_2. Finally, we model the binary label y_i again as detailed in Equation 9.

Table 2: Results of applying Bayesian deep learning and logistic regression to experiment 2, a right-skewed distribution. We include the mean accuracy, mean WTP, as well as the Bayesian credible interval with a 95% credibility.

Model	Accuracy	WTP (mean)	$[Q025,\ Q975]$
BDL	0.58	17.38	[5.59, 21.86]
Logistic	0.54	14.23	NA

The results of the synthetic datasets with right-skewed distribution for WTP are displayed in Fig.4. As we expected, the probability of paying for environmental good and the distribution of WTP have high variances. Notably, the Bayesian deep learning captures the right-skewed distribution of the willingness to pay. These results are common in contingent valuations surveys, where most people' willingness to pay corresponds only to a rather small amount of money for environmental conservation goods [42], in contrast, a rather small percentage of people will pay a high value. The accuracy from Table 2 is low for BDL and logistic model, and the Bayesian credible interval is between 5.59 and 21.86. These numbers demonstrate the high uncertainty in this experiment. The mean WTP is 17.38 for BDL and 14.23 for Logistic regression. The higher WTP for the BDL may be because this model can incorporate the right-skewed distribution of WTP, while the logistic function cannot.

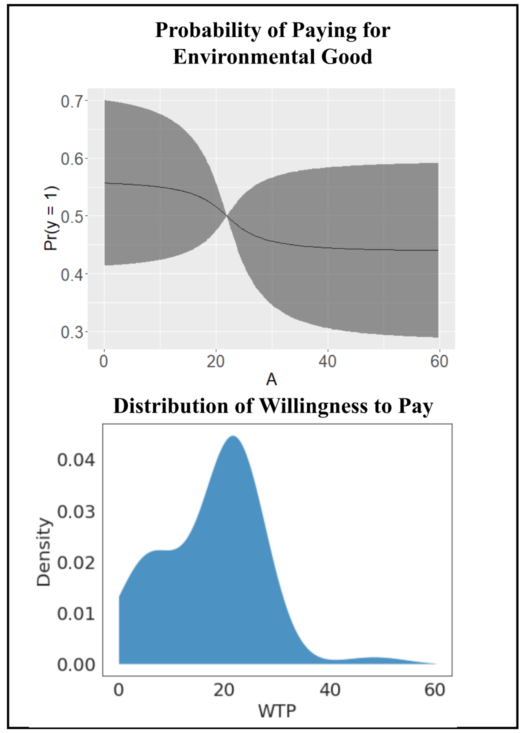

Figure 4: Summary results of the probability of paying for environmental good and the distribution of WTP for the right-skewed (positive skewness) distribution experiment using the Bayesian deep learning model.

6 REAL DATA: FOREST CONSERVATION

The real data set we use in our experiments corresponds to a contingent valuation survey conducted in Bolivia [29]. The goal of this study was to measure the conservation of the forest in the upper and middle basin of the Pirai River, Santa Cruz, Bolivia [29]. The sample size of the survey was 501 observations with an error of 4.4 percent. The target population was concentrated in the middle and lower income levels due to difficulties associated with interviewing people in the high-level income bracket. The main question for obtaining the willingness to pay asked was "Given this information, are you willing to pay monthly $\$A$ Bolivians to support this project and in this way to preserve the forest in the upper and middle basin of the Pirai River and assure the provision of the environmental services including water provision, avoiding floods and droughts, and maintaining favorable weather?"

The descriptive statistics are given in Table 3. From the 501 observations, 236 were negative answers ($y = 0$). The majority of negative answers claimed economic reasons for not paying. Some example are: "I cannot afford additional costs" (23.7 percent); "I'd rather spend that money on other goods" or "The suggested cost is too high for my budget" (30.2 percent). Lastly, 11 participants expressed that they did not believe that the program would have any benefit. The mean bid (A) was 19.40 in Bolivian currency.

In Fig. 5 and Table 4, we summarize the results of applying the Bayesian deep learning model to real data for forest conservation

Table 3: Descriptive statistics of the contingent valuation survey for forest conservation.

Variable	Mean	p50	Min	Max	SD
A	19.40	17	1	55	15.23
y	0.51	1	0	1	0.50
Household size	5.03	5	1	18	2.30
NSE	1.73	2	1	5	0.79
AGE	42.75	45	25	70	13.94
EDUC	1.65	1	0	4	1.16
Water bill	136.53	110	20	800	99.09
Electricity bill	184.44	160	25	810	111.78

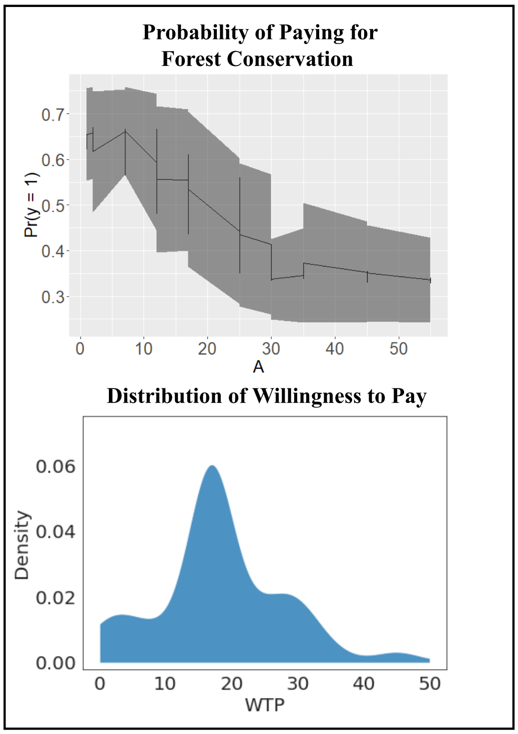

Figure 5: Summary results of the probability of paying for forest conservation and the distribution of WTP using the Bayesian deep learning model.

in Bolivia. Notably, in Fig. 5, the probability of paying for forest conservation, as well as, the distribution of the willingness shows a high variance, instead of a Gaussian distribution. This suggests a high uncertainty in the preferences for forest conservation. From Table 4, we observe the accuracy for different models, where BDL and Logistic have a low performance around 62%. The mean values of WTP are 18.04 and 17.00 Bolivian currency for the Bayesian deep learning model and logistic regression models, respectively.

The difference between the two might be explained because of the right-skewed distribution of WTP estimated by the BDL model, reflected in Fig. 5. The logistic regression method does not allow to estimate the distribution of WTP. Furthermore, the 95% Bayesian

Table 4: Results of applying Bayesian DL and logistic regression (LR) to the real data. The mean WTP in Bolivian currency is 18.04 for BDL and 17 for LR. The Bayesian credible interval, $[Q0.025,\ Q0.975]$, reveals a 95% probability the WTP lies within the interval.

Model	Accuracy	WTP (mean)	$[Q0.025,\ Q0.975]$
BDL	0.633	18.04	[0.91, 45.08]
LR	0.613	17.00	NA

credible interval for the WTP in Table 4 suggests that there is a 95% probability that the WTP would lie within 0.91 and 45.08 Bolivian currency, given the evidence provided by the contingent valuation survey. This is another argument in favor of modeling the high uncertainty of individual preferences for environmental goods.

7 DISCUSSION OF IMPACT, LIMITATIONS, AND FUTURE WORK

Considerations of using deep learning for contingent valuation. Leveraging Bayesian deep learning to obtain the distribution of the willingness to pay (WTP) for environmental goods promises to have a positive impact on policy-makers because they can make better decisions related to environmental conservation. Yet, unfortunately, to date, deep learning has not been implemented widely in environmental economics. A possible explanation is the computational barrier; some knowledge of computer science is required and the GPU hardware required for model training is expensive. Additionally, lack of deep learning model interpretability has been used as an argument against the use of deep learning. On the flip side, the most famous deep learning frameworks, such as Pytorch [25] and TensorFlow [34], now allow for the training of deep learning model with only a couple of line of code. Further, there is research focused on making deep learning models interpretable [10] and explainable [4, 14]. While outside the scope of this research, future work could apply such interpretability to contingent valuation (CV) models.

Accuracy considerations. In this research, we propose a methodology to learn the distribution of WTP for environmental goods in the context of CV. Traditional machine learning methods, like logistic regression, typically used in CV do not reveal the distribution of WTP. However, both synthetic and real datasets do not show a big improvement in the accuracy for the classification task. The reason could be the complex preferences needing to be modeled for multi-class problems, unlike binary classification problems. Future work could thus include experimenting with other distributions that may yield higher accuracies. For instance, a binomial distribution could be used instead of a Bernoulli distribution within the Bayesian deep learning models.

Extension to other stated preference methods. CV is a stated preference method to get the WTP for a non-market good. However, this method just provides the mean or median of WTP, while our approach now allows us to also reveal the actual distribution of the WTP. Nonetheless, the CV method has some disadvantages related to the validity of the results and biases and errors because of the survey format [41]. We can however extend the Bayesian deep learning model to other stated preferences methods, such as contingent behavior [36], contingent ranking [40], and Choice experiments [1]. Notably, choice experiments have become a popular stated preference method in health economics [13] and environmental economics [23], because of the inclusion of more scenarios in the survey design that better capture individual preferences.

8 CONCLUSION

In this research, we present a framework to use Bayesian Deep Learning models with the contingent valuation method to measure the uncertainty of the probability of paying for an environmental good. From this framework, we estimate the empirical distribution of the willingness to pay (WTP) for environmental goods. Our evaluation study using several synthetic datasets demonstrates that our methodology captures the Gaussian and right-skewed distributions for the simulated willingness to pay. In addition, we also leverage real data, namely, a contingent valuation survey about forest conservation in Bolivia, that exhibits a high uncertainty in the preferences for forest conservation with a 95% Bayesian credible interval for the WTP of 0.91 and 45.08. The results thus suggest that the amount of money the population is willing to pay for forest conservation varies greatly. Therefore, our research can be used to provide more information to environmental policy-makers for the conservation of environmental goods.

ACKNOWLEDGMENTS

We thank the Fulbright Foreign Student Program, the Chilean National Agency for Research and Development (ANID)/Scholarship Program/DOCTORADO BECAS CHILE/2015-56150007, and the Data Science department at WPI for funding. We thank Thomas Hartvigsen and the DAISY research lab at WPI for support.

REFERENCES

[1] Francisco Alpizar, Fredrik Carlsson, Peter Martinsson, et al. 2001. Using choice experiments for non-market valuation. *Working papers in economics/Göteborg University, Dept. of Economics; no. 52* (2001).

[2] Kenneth Arrow, Robert Solow, Paul R Portney, et al. 1993. Report of the NOAA panel on contingent valuation. *Federal register* 58, 10 (1993), 4601–14.

[3] Kenneth J Arrow. 2001. *Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries.* Oxford University Press on Demand.

[4] Xiao Bai, Xiang Wang, Xianglong Liu, Qiang Liu, Jingkuan Song, Nicu Sebe, and Been Kim. 2021. Explainable deep learning for efficient and robust pattern recognition: A survey of recent developments. *Pattern Recognition* 120 (2021).

[5] Thomas Bue Bjørner. 2004. Combining socio-acoustic and contingent valuation surveys to value noise reduction. *Transportation Research Part D: Transport and Environment* 9, 5 (2004), 341–356.

[6] Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. 2015. Weight uncertainty in neural network. In *Int. conference on machine learning*. PMLR, 1613–1622.

[7] Luis Bravo-Moncayo, José Lucio Naranjo, Ignacio Pavón García, and Roberto Mosquera. 2017. Neural based contingent valuation of road traffic noise. *Transportation Research Part D: Transport and Environment* 50 (2017), 26–39.

[8] Luis Bravo-Moncayo, Ignacio Pavón-García, José Lucio-Naranjo, and Roberto Mosquera. 2017. Contingent valuation of road traffic noise: A case study in the urban area of Quito, Ecuador. *Case studies on transport policy* 5, 4 (2017), 722–730.

[9] Richard T Carson and W Michael Hanemann. 2005. Contingent valuation. *Handbook of environmental economics* 2 (2005), 821–936.

[10] Supriyo Chakraborty et al. 2017. Interpretability of deep learning models: A survey of results. In *IEEE smartworld, ubiquitous intelligence & computing, advanced & trusted computed, scalable computing & communications, cloud & big data computing, Internet of people and smart city innovation.*

[11] Patricia A Champ, Kevin J Boyle, Thomas C Brown, and L George Peterson. 2003. *A primer on nonmarket valuation.* Vol. 3. Springer.

[12] Declan Conway, Edward Allison, Richard Felstead, and Marisa Goulden. 2005. Rainfall variability in East Africa: implications for natural resources management and livelihoods. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 363, 1826 (2005).

[13] Esther W de Bekker-Grob, Mandy Ryan, and Karen Gerard. 2012. Discrete choice experiments in health economics: a review of the literature. *Health economics* 21, 2 (2012), 145–172.

[14] Filip Karlo Došilović, Mario Brčić, and Nikica Hlupić. 2018. Explainable artificial intelligence: A survey. In *41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, 210–215.

[15] Colette M Escobar, W Steven Barnett, and John E Keith. 1988. A contingent valuation approach to measuring the benefits of preschool education. *Educational Evaluation and Policy Analysis* 10, 1 (1988), 13–22.

[16] Yarin Gal et al. 2016. Uncertainty in deep learning. (2016).

[17] Yarin Gal and Zoubin Ghahramani. 2016. A theoretically grounded application of dropout in recurrent neural networks. *Advances in neural information processing systems* 29 (2016).

[18] Andrew Gelman, John B Carlin, Hal S Stern, and Donald B Rubin. 1995. *Bayesian data analysis.* Chapman and Hall/CRC.

[19] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep learning.* MIT press.

[20] Alex Graves. 2011. Practical variational inference for neural networks. *Advances in neural information processing systems* 24 (2011).

[21] Timothy C Haab and Kenneth E McConnell. 2002. *Valuing environmental and natural resources: the econometrics of non-market valuation.*

[22] W Michael Hanemann. 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American journal of agricultural economics* 66, 3 (1984), 332–341.

[23] Nick Hanley, Robert E Wright, and Vic Adamowicz. 1998. Using choice experiments to value the environment. *Environmental and resource economics* 11, 3 (1998), 413–428.

[24] Simon Jackman. 2009. *Bayesian analysis for the social sciences.* J. Wiley & Sons.

[25] Nikhil Ketkar and Jojo Moolayil. 2021. Introduction to pytorch. In *Deep learning with python.* Springer, 27–91.

[26] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).

[27] Thomas Klose. 1999. The contingent valuation method in health care. *Health policy* 47, 2 (1999), 97–123.

[28] Kelvin J Lancaster. 1966. A new approach to consumer theory. *Journal of political economy* 74, 2 (1966), 132–157.

[29] Felipe Vásquez Lavín, Ricardo Flores, and Verónica Ibarnegaray. 2017. A Bayesian quantile binary regression approach to estimate payments for environmental services. *Environment and Development Economics* 22, 2 (2017), 156–176.

[30] John Loomis, Paula Kent, Liz Strange, Kurt Fausch, and Alan Covich. 2000. Measuring the total economic value of restoring ecosystem services in an impaired river basin: results from a contingent valuation survey. *Ecological economics* 33, 1 (2000), 103–117.

[31] Scott M Lynch. 2007. *Introduction to applied Bayesian statistics and estimation for social scientists.* Springer Science & Business Media.

[32] Robert Cameron Mitchell and Richard T Carson. 2013. *Using surveys to value public goods: the contingent valuation method.* Rff Press.

[33] Radford M Neal. 2012. *Bayesian learning for neural networks*. Vol. 118. Springer Science & Business Media.

[34] Bo Pang, Erik Nijkamp, and Ying Nian Wu. 2020. Deep learning with tensorflow: A review. *J of Educational and Behavioral Stats* 45, 2 (2020), 227–48.

[35] Ankit Patel, Minh Nguyen, and Richard Baraniuk. 2016. A probabilistic framework for deep learning. *Advances in neural info. processing systems* 29 (2016).

[36] John W Payne. 1982. Contingent decision behavior. *Psychological bulletin* 92, 2 (1982), 382.

[37] Roberto Roson, Alvaro Calzadilla, and Francesco Pauli. 2006. Climate change and extreme events: an assessment of economic implications. (2006).

[38] John Salvatier, Thomas V Wiecki, and Christopher Fonnesbeck. 2016. Probabilistic programming in Python using PyMC3. *PeerJ Comp Sci* 2 (2016).

[39] Quirin Schiermeier. 2019. Eat less meat: UN climate-change report calls for change to human diet. *Nature* 572, 7769 (2019), 291–293.

[40] Ulla Slothuus, Mette L Larsen, and Peter Junker. 2002. The contingent ranking method—a feasible and valid method when eliciting preferences for health care? *Social science & medicine* 54, 10 (2002), 1601–1609.

[41] Lingappan Venkatachalam. 2004. The contingent valuation method: a review. *Environmental impact assessment review* 24, 1 (2004), 89–124.

[42] Dale Whittington and Stefano Pagiola. 2012. Using contingent valuation in the design of payments for environmental services mechanisms: A review and assessment. *The World Bank Research Observer* 27, 2 (2012), 261–287.

[43] Arnold Zellner. 1988. Bayesian analysis in econometrics. *Journal of Econometrics* 37, 1 (1988), 27–50.

[44] Xilei Zhao, Xiang Yan, Alan Yu, and Pascal Van Hentenryck. 2018. Modeling Stated preference for mobility-on-demand transit: a comparison of Machine Learning and logit models. *arXiv preprint arXiv:1811.01315* (2018).