

ISSN: 2543-6821 (online)

Journal homepage: <http://ceej.wne.uw.edu.pl>

Nan Zhang

Comparative Analysis of Hedonic Wage and Discrete Choice Models in Valuing Job Safety

To cite this article

Zhang, N. (2025). Comparative Analysis of Hedonic Wage and Discrete Choice Models in Valuing Job Safety. Central European Economic Journal, 12(59), 1-16.

DOI: 10.2478/ceej-2025-0001

 To link to this article: <https://doi.org/10.2478/ceej-2025-0001>



Nan Zhang 

Trento University, Department of Economics and Management, Via Vigilio Inama, 5,
38122, Trento TN, Italy
corresponding author: nan.zhang@unitn.it

Comparative Analysis of Hedonic Wage and Discrete Choice Models in Valuing Job Safety

Abstract

This is the first study to compare the hedonic wage and discrete choice models in the context of real job market choices. The hedonic wage method is the most widely adopted approach in estimating the value of marginal changes in mortality risks (or value of a statistical life, VSL). Alternatively, the discrete choice model offers a framework for valuing non-market goods, often used in choice experiments involving hypothetical job selection scenarios. This study employs both the hedonic wage and discrete choice models, using real market data from Taiwan's Panel Survey of Family Dynamics (PSFD 2016-2018), to examine workers' industry choices and infer their willingness to pay for job safety. By comparing the estimation results from the discrete choice model and the hedonic wage model, this study finds that the marginal willingness to pay (MWTP) for risk reduction estimated from discrete choice models is unreasonably large and therefore prefers the hedonic wage model in estimating the value of job safety. The findings of this study contribute to the literature comparing the hedonic model and the discrete choice model in evaluating environmental goods, which are mainly conducted using housing market data.

Keywords

Nonmarket valuation | Hedonic wage model | Random utility model | Value of safety | VSL

JEL Codes

J17, J28, Q51

1. Introduction

The hedonic wage model and discrete choice model are the two most widely used nonmarket valuation methods. The hedonic wage model is a revealed preference method which is widely used in estimating the job safety (referred as “value of a statistical life, VSL” in literature). When the job market is in equilibrium, the price of jobs (wages) is a non-linear function of job attributes and worker attributes. By regressing wages on a specific job/worker attribute, one can infer the marginal implicit price of this job/worker attribute, which also equals the marginal willingness to pay (MWTP) for the chosen quantity of that attribute.

Discrete choice models provide an alternative way of valuing nonmarket goods and can be applied to both revealed data and hypothetical data. The Random Utility Model (RUM) is the most widely adopted discrete choice model. The RUM model maximises the probability that a specified bundle of attributes is chosen by an individual to maximise the utility, which

is defined over attributes (such as the housing attributes in residential sorting literature, or job attributes in survey-based choice experiment studies focusing on the job market). To the author's knowledge, the discrete choice model has not been applied to real labour market data to study the willingness to pay for job attributes, such as job risk reduction.

This study applies the discrete choice model to real labour market data in Taiwan to estimate the MWTP for job safety and compares the results from the hedonic wage model using the same data. This study finds that the hedonic wage model outperforms the discrete choice model in valuing job safety, suggesting that the hedonic wage model provides a more reliable estimate of marginal willingness to pay (MWTP) for job risk reduction. This study also explores the possible explanations for the unsatisfying performance of the discrete choice model.

This study contributes to three branches of valuation literature: the choice experiment (CE) studies that focus on the labour market, the residential sorting

literature, and the comparison between the hedonic price model and the discrete choice model.

First, this study applies the discrete choice model to real labour market data to examine workers' industry choices. In choice experiment (CE) studies, subjects are presented with hypothetical industry options, each characterised by different levels of attributes, including wages and other non-monetary job features, and asked to select the industry that maximises their utility. By analysing workers' choices, researchers can assess the relative importance of various attributes and infer their willingness to pay (WTP) for these non-monetary features. However, CE studies rely on stated preferences and suffer from hypothetical bias.

The second line of literature this study connects to is the residential sorting literature. These studies adopt discrete choice models to investigate consumers' residential choices (according to the living cost, earning opportunity, moving cost, and environmental attributes of different residential locations) and to infer their WTP for environmental goods such as air quality and climate (Cragg and Kahn, 1997; Sinha and Cropper, 2013; Chay and Greenstone, 2005; Bayer et al., 2009; Freeman et al., 2019). Currently, this sorting literature has never been applied to the labour market to study WTP for job attributes, such as job risk.

The third line of literature that this study contributes to is the comparison between the performance of the hedonic and discrete choice models in the housing market. For now, there is no unanimous agreement on which model performs better. For instance, Cropper et al. (1993) compare the logit model and hedonic price model in estimating the preference for housing attributes in a single market through simulation. They find that the two approaches perform equally well in estimating the marginal value, but the logit model outperforms the hedonic price model in valuing non-marginal changes of attributes. Palmquist and Israngkura (1999) investigate the willingness to pay for air quality improvements with both the hedonic price model and the random utility model. They conclude that the hedonic wage model is more successful than the random utility model when multiple markets are available. Sinha et al. (2021) compare the two approaches in estimating MWTP for climate attributes with housing market data. They find that the discrete choice model results in an MWTP for warm winters twice as large as the hedonic price model, though the MWTP for cooler summers is about the same for both approaches.

The structure of this paper is arranged as follows: Section 2 compares the theoretical framework of the two valuation approaches. Section 3 describes the data used in this study. Section 4 estimates the MWTP for job safety with the three discrete choice specifications and the hedonic wage model. Section 5 discusses the differences between the two models and concludes.

2. Theoretical Framework

2.1. Random Utility Model

The discrete choice model is widely used in valuing environmental goods and recreation demand. By exploring how individuals make choices among discrete alternatives, such as different brands of commodities, transportation modes, fishing sites, occupation/industries, etc., the discrete choice model can be used to infer consumers' WTP for an attribute and to predict the changes in demand for a certain alternative associated with the changes in attributes. The random utility model (RUM) is the most widely adopted discrete choice model, which is based on the work of economists including Manski and McFadden (Manski and Lerman, 1977; McFadden, 1974; McFadden, 1980; McFadden & Train, 2000).

Suppose a worker i 's utility from working in industry j consists of four parts: the consumption of the ordinary commodity Q , job risk R , other industry-specific job attributes (such as working hours and the unemployment rates) X , and Y , workers' personal attributes (such as education, age, and gender). Including these personal attributes can help investigate the observed heterogeneity in preferences. A worker chooses an industry to maximise his/her utility under the budget constraint:

$$U = V(Q, R, X, Y) + \varepsilon, s.t. Q = W + I \quad (1)$$

where ε is the unobserved term that affects the utility, which has the IID extreme value distribution. V is the representative utility. W is the wage, and I is the non-wage income.

Following the random utility studies, this study also assumes a linear representative utility function. Worker i 's utility from choosing alternative j is:

$$U = \alpha(W + I) + \beta R + \theta X + \delta Y + \varepsilon \quad (2)$$

Since job risk is bad, we expect β to be negative. j is set to zero in this study for simplicity. Usually, one income/cost variable is included to measure the marginal utility of money. Thus, a utility maximisation consumer will choose the amount of risk according to the first-order condition:

$$dV = \alpha dW + \beta dR + \theta dX + \delta dY = 0 \quad (3)$$

The role of these personal attributes can help to investigate the heterogeneity in preference. To be specific, we include workers' gender, age, and education in the utility function. These worker attributes are case-specific. The other job attributes included here are working hours and the unemployment rates for each industry. These variables are alternative-specific. Under the specification, the utility of worker i working in industry j can be written as follows:

$$U_{ij} = V_{ij} + \varepsilon_{ij} = \alpha W_{ij} + \beta R_j + \theta X_j + \delta Y_i + \varepsilon_{ij} \quad (4)$$

where ε_{ij} is distributed IID extreme value. Thus, the probability that worker i will choose industry j is:

$$P_{ij}(\forall l \neq j) = \frac{e^{(\alpha W_{ij} + \beta R_j + \theta X_j + \delta Y_i)}}{\sum_l e^{(\alpha W_{il} + \beta R_l + \theta X_l + \delta Y_i)}} \quad (5)$$

which can be estimated with the conditional logit model, as in most random utility models.

The fixed coefficient β in the conditional logit model indicates that the preferences over an attribute are the same across consumers, which is a major weakness of the conditional logit model. Another shortcoming of the conditional logit model is the IIA (independence from irrelevant alternatives) property, which assumes that the probability ratio between two alternatives is irrelevant to a third alternative. The IIA property is a strong restriction which is sometimes inappropriate in real life.

The mixed logit model is a very flexible model that can solve the limitations of the standard logit model by allowing for unobserved random taste variation and unrestricted substitution pattern. In addition,

the panel data mixed logit model can also address the correlation of the unobserved tastes over time. McFadden and Train (2000) show that the mixed logit model can be used to approximate any random utility model.

In the mixed logit model, the coefficient (β) is a distribution rather than a fixed value, and the choice probability can be written as:

$$P_{ij} = \int \frac{e^{(\alpha W_{ij} + \beta R_j + \theta X_j + \delta Y_i)}}{\sum_l e^{(\alpha W_{il} + \beta R_l + \theta X_l + \delta Y_i)}} f(\beta) d\beta \quad (6)$$

Further, considering workers are making repeated choices over time, let $j=(j_1, \dots, j_T)$ denote the vector of alternatives a worker faces, then the probability of the worker choosing alternative j can be written as:

$$P_{ij} = \int \prod_{t=1}^T \frac{e^{(\alpha W_{ijt} + \beta R_{jt} + \theta X_{jt} + \delta Y_i)}}{\sum_l e^{(\alpha W_{ilt} + \beta R_{lt} + \theta X_{lt} + \delta Y_i)}} f(\beta) d\beta \quad (7)$$

which can be estimated through the panel data mixed logit model. The coefficients $\alpha, \beta, \theta, \delta$ are parameters to be estimated.

As shown in the first order condition (equation 3), α measures the marginal utility of money, β, θ, δ measures the marginal utility associated with changes in job/worker attributes, dividing the β, θ, δ by the marginal utility of money (α) provides a monetary estimate of the MWTTP of these job/worker attributes. Specifically, the marginal willingness to pay for risk reduction is given by:

$$MWTTP = -\frac{\partial U / \partial R}{\partial U / \partial W} = -\frac{\beta}{\alpha} \quad (8)$$

The change in consumer surplus is equal to:

$$\Delta CS = \frac{1}{\alpha} (V^1 - V^0) \quad (9)$$

where V^1 and V^0 are the utility after and before the change (in job risk, for instance), respectively.

2.2. Hedonic Wage Model

The hedonic wage model is built based on the hedonic price theory by Thaler and Rosen (1976). A job can be seen as a composite good with its attributes.

From a worker's perspective, a job can be considered as a differentiated good consisting of various job characteristics, such as flexibility in working hours, the number of paid vacation days, working conditions, prestige, training and enhancement of skills, the risk of accidental injury or exposure to toxic substances, etc. Different jobs have different bundles of characteristics and different wage rates. Wage differentials can be interpreted as the implicit prices of job characteristics. From the perspective of employers, they are the suppliers of the job attributes. At the same time, employers can be viewed as choosing from among a set of workers of different characteristics. Thus, the job market also functions as the implicit market of job attributes and worker attributes.

The hedonic wage equation is an equilibrium relationship that reflects the interaction of supply and demand for worker attributes. It is the double-envelope of both the firm's iso-profit curves and workers' indifferent curves. In the hedonic wage model, the wages in a job market reflect the equilibrium relationship between workers and firms. However, the discrete choice model considers the wage as a job attribute that is exogenous. The hedonic wage equation is a function of the worker attributes Y , job attributes X , and job risk R and is exogenous to each individual in this market.

$$W = W(R, X, Y) \quad (10)$$

The marginal implicit price of each characteristic can be calculated by taking the derivative of (10) with respect to each characteristic. For example, the implicit price of risk, P_R , is the partial derivative of W with respect to R :

$$P_R \equiv \frac{\partial W}{\partial R} \quad (11)$$

It is often assumed that the shape of (10) is exponential:

$$W = e^{\pi_0 + \pi_1 R + \pi_2 X + \pi_3 Y} \quad (12)$$

Econometrically, the hedonic wage equation is therefore estimated using a log-linear form:

$$\ln W = \pi_0 + \pi_1 R + \pi_2 X + \pi_3 Y + \varepsilon_1 \quad (13)$$

where $\pi_0, \pi_1, \pi_2, \pi_3$ are parameters to be estimated, $\varepsilon_1 \sim N(0, \sigma^2)$.

After the estimation, the implicit price of risk (P_R) can be calculated as follows:

$$P_R = \frac{\partial \hat{W}}{\partial R} = \hat{\pi}_1 e^{\hat{\pi}_0 + \hat{\pi}_1 R + \hat{\pi}_2 X + \hat{\pi}_3 Y + 0.5 \hat{\sigma}^2} \quad (14)$$

where $\hat{\sigma}^2$ is the unbiased estimator of σ^2 . The nonlinearity of the hedonic wage equation implies that the marginal price of risk is not a constant but a function of the quantity of risk. If $\pi_1 > 0$, there is a positive wage premium for risk. In addition, the partial derivative of P_R is also positive, suggesting that the price is increasing in risk.

The implicit price of risk (P_R) can only measure the value of a marginal change in risks. To measure the value of non-marginal changes in risks, the demand for safety needs to be estimated, which requires data from multiple labour markets to meet the identification criteria. Please see Zhang et al. (2023) for details about the identification and estimation of the second-stage hedonic wage model.

3. Data

Three levels of data are combined in this study. The first is the individual-level data from the PSFD (Panel Survey of Family Dynamics) conducted in Taiwan. The PSFD survey started in 1999 and collected information on workers' personal information (wage, education, gender, work experience, etc.). The survey is of high quality and has a panel structure, conducted every year before 2012 and every two years after 2012.

The currently unemployed respondents are excluded in the study, following the literature. The agricultural sector is excluded because the wage equation in the agriculture industry is different from other industries. The mining sector and the public sector are also excluded because there are not enough observations in these two sectors. The lack of observations in a given industry will make it hard to get an accurately estimated wage equation for this industry, which is required for the discrete choice model. For the same reason, only data from the years 2016 and 2018 are used because other waves of survey do not have enough observations for each of the seven industries. Thus, we end up with seven

industries: Manufacturing, Construction, Catering, Transportation, Finance, Science, and Service. The number of observations in the years 2016 and 2018 are 2293 and 1903, respectively.

Then comes the industry-level data, including job risk, working hours, and unemployment rates. The industry mortality risk (*risk_alt*) is our core independent variable. I use the moving average of the industry fatality rates for the past three years rather than the current year's fatality rate to reduce the fluctuations due to stochastic shocks, following the VSL literature (Zhang et al., 2022, Hintermann et al., 2010, Kniesner et al., 2012; Guo and Hammitt, 2009). For instance, the fatality rate in the mining industry in 2015 was 0; however, in the years before (2014) and after (2016), it was extremely high (0.766 and 0.532 per thousand). Thus, using the average risk values over the past several years rather than the current year value may better reflect the risks of this industry.

There are two sources of variations in the job risk of an individual, one being the change in risk over time if the worker stays in the same industry, and the other comes from the worker switching to a different sector.

There are two considerations for choosing the industry-specific job attributes except job risk. Following the choice experiment (CE) studies on labour market choices, wage is considered one of the most important factors in workers' labour market choices by most studies (Guo et al., 2010; Mangham and Hanson, 2008). Besides wages, working intensity is also an important factor in deciding which job to take (Kornstad and Thoresen, 2007). In this study, I use the average monthly total working hours (*workhour_alt*) to proxy the work intensity in each industry. The industry unemployment rate (*umem_alt*) is used to proxy the competition. A higher unemployment rate denotes higher competition.

The industry mortality risk, average working hours by industry, and unemployment rates by industry are all from the Bureau of Labour Insurance, Ministry of Labour, Taiwan (https://statfy.mol.gov.tw/statistic_DB.aspx). These three industry-specific attributes are recategorised according to the Statistical Classification of Industry System of the Republic of China (Rev.6, 1996), as in the hedonic wage model. The mean and standard deviation of the industry-specific attributes, including *wage_alt*, are organised by industry and shown in appendix, Table A3.

This study also incorporates town-level climate data, recognising climate as an important factor in

both residential location choices and its influence on labour supply and wage income in a given city. The analysis focuses on two key climate variables: the average temperatures in January and July (1981-2005), representing the coldest and warmest months in Taiwan. The climate data, provided by the Taiwan Climate Change Projection and Information Platform (TCCIP)¹, is of high quality. Our survey dataset covers 218 out of 352 townships in Taiwan, with the town-level climate data matched to individual worker data based on the workplace addresses reported in the survey.

The variable definition and the summary statistics of the variables used in this study are shown in Table 1.

4. Empirical results

4.1. The Discrete Choice Model

We do not observe the wage a worker can earn in every industry but the wage in the industry he/she has already chosen. The first step, which is the same as in the sorting literature, is to predict wages in alternative industries that are not chosen by a worker.

For each of the seven industries, a wage equation was estimated in both years to capture the wage structure in this specific industry and then used to predict the wage income one specific worker would get if he/she chose this industry. The control variables mainly include factors that affect one's productivity, such as gender, age, education, etc. The results of the first stage wage regressions for 2016 and 2018 are shown in Tables A1 and A2 in the Appendix. For later analysis in the discrete choice model, *wage_alt* will be the reported real wage for industries chosen by workers, and the wages in alternative industries that workers do not choose will be the predicted wages based on these industry wage regressions.

This study estimates three different discrete choice models: the conditional logit model, which assumes all coefficients are fixed (that workers have homogenous tastes over industry attributes); the mixed logit model, which assumes the coefficient of job attributes to be random, thus allowing for the heterogeneity in tastes variety; and the panel-data mixed logit model, which allows the coefficients to be correlated over time.

¹ <http://tccip.ncdr.nat.gov.tw/NCDR/main/usage.aspx>

Table 1. Variable Descriptive Statistics

Category	Variable	Definition	Mean	Min	Max
Individual (Obs. 4035)	wage	Yearly wage in 2014 value (unit: 1000 TWD)	537.2	28.5	5587.7
	eduear	Education years	14.1	0.0	22.0
	feduear	Father's education years	9.2	0.0	22.0
	female	1 if female, 0 otherwise	0.4	0.0	1.0
	age	age	35.1	25.0	79.0
	marriage	1 if married, 0 otherwise	0.5	0.0	1.0
	scale	Number of employees in the firm	212.2	2.0	500.0
	wexp	Working experience years	13.6	1.0	71.0
	healthlevel	1 for very good health, 5 for very bad health	3.6	1.0	5.0
Industry (seven industries)	risk	industry mortality rates, unit: 1/1,000,000	23.6	0.0	124.3
	workhour_alt	industry average monthly total work hours, unit: hour	169.8	159.8	176.0
	unem_alt	industry unemployment rates, unit: %	3.2	2.6	4.6
Town (218 towns)	TJAN	township-level average January temperature (Celsius) from 1981 to 2005	16.4	13.1	20.8
	TJUL	township-level average July temperature (Celsius) from 1981 to 2005	28.8	26.4	30.3
Regional dummy	NORTH	1 if respondent lives in Northern Taiwan, 0 otherwise	0.5	0.0	1.0
	CENTER	1 if respondent lives in central Taiwan, 0 otherwise	0.2	0.0	1.0
	SOUTH	1 if respondent lives in southern Taiwan, 0 otherwise	0.2	0.0	1.0
	EAST	1 if respondent lives in eastern Taiwan, 0 otherwise	0.0	0.0	1.0

Source: PSFD 2016, 2018; Bureau of Labour Insurance, Ministry of Labour. Taiwan Climate Change Projection and Information Platform (TCCIP) project

Specification 1: Conditional Logit Model

First, the random utility model in equation (4) was estimated using the conditional logit model with pooled observations from 2016 and 2018. The results are shown in Columns (1)-(2), Table 2.

As can be seen from Columns (1) and (2) in Table 2, workers prefer industries that have higher wage income, longer working hours, lower risks, and lower unemployment rates. The preferences are heterogeneous across workers. Compared with Manufacturing, high-education workers are less likely to choose Construction and Catering but are more likely to work in the Finance and Service industries. Older workers are less likely to choose a job in the Catering and Science industries. Female

workers are less likely to get a job in the Construction and Transportation industries and are more likely to work in Catering, Finance, Science, and Service, as expected. Including the worker attributes can help solve the observed heterogeneity in preferences. However, there are cases in which unobserved worker attributes could also affect workers' industry choices and can be addressed with the mixed logit model.

4.3. Specification 2: Mixed Logit Model

Because of the advantages of the mixed logit model in the flexible substitution pattern and allowing for unobserved heterogeneity, this study estimates a mixed logit model with pooled observations of 2016

Table 2. Regression Results-Discrete Choice Model

	CL		ML		PML	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
wage_alt	0.0017***	0.0013***	0.0022***	0.0018***	0.0018***	0.0016***
risk_alt	-0.1091***	-0.0986***	-0.1244***	-0.1162***	-0.1632***	-0.1422***
workhour_alt	0.0251**	0.0096	0.1633***	0.1039***	0.1529***	0.1085***
unem_alt	-0.7097***	-0.5777***	0.3434	0.1138	-0.1729	-0.2892
sd(risk_alt)			0.3826***	0.3186***	0.2546***	0.2323***
sd(workhour_alt)			0.4681***	0.4270***	0.4601***	0.4771***
sd(unem_alt)			0.0026	0.0041	2.3677***	2.7743***
Construction						
eduyear		-0.2003***		-1.0500**		-1.0161***
age		-0.0134		0.0941		0.1238**
female		-1.0217***		-15.3042***		-11.4380***
_cons	9.0968***	11.2592***	-42.8469***	-19.2340	-19.0045***	-5.0814
Catering						
eduyear		-0.1084***		-0.1538*		-0.1549**
age		-0.0296***		-0.0946***		-0.0997***
female		0.6325***		3.2816***		3.2010***
_cons	-2.0309***	0.7968	-1.3451***	2.2099	-1.3025***	2.5388*
Transportation						
eduyear		-0.0522		-0.1804		-0.1984***
age		-0.0162		-0.0123		0.0004
female		-0.3839*		-3.2871***		-2.4575***
_cons	-2.9876	1.8339**	-8.5789***	-2.9876	-5.4116***	0.1871
Finance						
eduyear		0.2072***		0.2016**		0.2588***
age		0.0058		-0.0370		-0.0454**
female		0.6392***		3.1610***		2.9912***
_cons	-3.7770***	-7.1861***	-8.5789***	-9.0142***		-8.6711***
Science						
eduyear		0.0321		0.0385		0.0385
age		0.0105		-0.0011		0.0315
female		0.4214***		0.6866***		1.1440***
_cons	-0.6070***	-1.8486***	-1.9667***	-2.4552***	-2.9669***	-5.4924***
Service						
eduyear		0.1624***		0.1287		0.2505***
age		-0.0345***		-0.1047***		-0.1237***
female		1.1372***		3.9960***		4.1265***
_cons	-2.2029***	-4.0223***	-2.1834***	-2.8356	-2.4155***	-4.1438**
N	32,347	32,347	32,347	32,347	32,347	32,347
Pseudo R2	1.41%	1.32%	30.93%	31.52%	27.36%	28.19%
Log-likelihood	-7388.89	-7085.43	-7127.41	-6923.56	-6107.19	-5918.74
AIC	14797.78	14226.87	14280.81	13909.11	12750.25	11899.47

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

and 2018. Following the literature, such as Huang et al. (2018), the coefficients of industry attributes are assumed to follow a random distribution, while the coefficient of income/cost (in our case, wage) is assumed to be fixed. In Model 6, personal attributes such as gender, education, and age are included to explore the observable heterogeneity in tastes. The results are shown in Models 3-4 in Table 2. In Model 3, only industry attributes are considered, and in Model 4, worker attributes are also included.

We can see that workers still prefer industries with higher wages, longer working hours, and lower mortality rates, similar to what we find in the conditional logit model. However, the coefficients of unemployment become insignificant, suggesting that unemployment has no impact on workers' industry choices. In Model 4, the coefficient of *risk_alt* has a mean of -0.1162 and a standard deviation of 0.3186, both the mean value and the standard deviation are significant, suggesting rich variations in the population.

Similarly, the mean of the coefficient of workhour is 0.104, and the mean of the standard deviation of workhour is 0.427. Both are significant. The relatively large standard deviations suggest that workers have various tastes over working hours. However, the mean and the standard deviation of the coefficients of unemployment rates are insignificant. The individual-specific attributes affect the industry choices in a similar way as those found in the conditional logit model.

Specification 3: Panel data Mixed Logit Model

The panel-data mixed logit model has the advantage of capturing the correlation of workers' choices over time and is very useful for modelling consumer behaviour such as habit formation and/or variation-seeking. In our context, it is reasonable to assume that workers prefer to choose the same industry as they did in the last period; otherwise, their work experience and skills accumulated in the last job will be less appreciated. In this specification, we continue our assumption that all workers have the same preference regarding wages but heterogeneous preferences for risks, workhours, and unemployment. The regression results are shown in Models 5 and 6, Table 2.

As we can see, workers show heterogeneous preferences for job attributes: they prefer jobs with higher wages, longer working hours, and lower job risks. The coefficients of workhour are positive, with substation variations among the population, suggesting that average workers are more likely to be engaged in industries that have longer working hours. The coefficients of unemployment rates are negative but not significant, suggesting that overall, unemployment rates do not affect how workers choose industries.

Compared with Manufacture, high-education workers prefer the Finance and Science industries over Construction, Catering, and Transportation. Older workers are more likely to be working in Construction and less likely to work in Catering, Finance, and Service. Female workers prefer Service, Catering, Finance, and Science and are least likely to be engaged in Construction. By comparing the regression results of the three specifications using the AIC and log-likelihood, we find that the panel mixed logit models perform better than the other models.

In the discrete choice models, the marginal willingness to pay (MWTP) for one unit of risk reduction (1/1000,000) can be calculated as follows:

$$MWTP = -\frac{\hat{\beta}}{\hat{\alpha}}$$

where $\hat{\alpha}$ is the estimated coefficient of wage and $\hat{\beta}$ is the estimated coefficient of risk for the conditional logit model. In the mixed logit model and the panel data logit model, the coefficient of *risk_alt* ($\hat{\beta}$) is assumed to have a normal distribution. The mean and standard deviation of MWTP can be calculated as follows:

$$mean(MWTP) = -\frac{mean(\hat{\beta})}{\hat{\alpha}}$$

$$sd(MWTP) = -\frac{sd(\hat{\beta})}{\hat{\alpha}}$$

The mean, upper, and lower limits of MWTP are shown in Table 4.

Table 3. The Hedonic Wage Model Results

	(1)	(2)	(3)	(4)	(5)
	year2016	year2018	pool_1	pool_2	FE
risk_alt	0.0006*	0.0011***	0.0008***	0.0012***	0.0012***
	(0.0003)	(0.0004)	(0.0002)	(0.0003)	(0.0003)
eduyear	0.0660***	0.0765***	0.0710***	0.0711***	0.0706***
	(0.0052)	(0.0054)	(0.0038)	(0.0038)	(0.0038)
feduyear	0.0095***	0.0063**	0.0082***	0.0079***	0.0078***
	(0.0029)	(0.0030)	(0.0021)	(0.0021)	(0.0021)
female	-0.2149***	-0.2370***	-0.2248***	-0.2301***	-0.2299***
	(0.0166)	(0.0190)	(0.0125)	(0.0128)	(0.0128)
age	0.0158***	0.0113***	0.0142***	0.0141***	0.0138***
	(0.0039)	(0.0035)	(0.0026)	(0.0026)	(0.0026)
wexp	-0.0017	0.0011	-0.0005	-0.0003	-0.0003
	(0.0034)	(0.0031)	(0.0023)	(0.0023)	(0.0023)
marry	0.1471***	0.1156***	0.1351***	0.1349***	0.1346***
	(0.0170)	(0.0198)	(0.0130)	(0.0129)	(0.0129)
scale	0.0004***	0.0004***	0.0004***	0.0005***	0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
TJAN	-0.0503***	-0.0519***	-0.0502***	-0.0503***	-0.0507***
	(0.0066)	(0.0078)	(0.0050)	(0.0050)	(0.0050)
TJUL	0.0874***	0.1035***	0.0933***	0.0899***	0.0907***
	(0.0159)	(0.0166)	(0.0115)	(0.0115)	(0.0115)
workhour_alt				-0.0045***	-0.0042***
				(0.0011)	(0.0011)
unem_alt				-0.0238*	-0.0170
				(0.0137)	(0.0137)
year					0.0158**
					(0.0061)
_cons	2.7904***	2.3943***	2.6150***	3.5419***	-28.3320**
	(0.4185)	(0.4370)	(0.3026)	(0.3666)	(12.4421)
N	2293	1742	4035	4035	4035
R ²	0.349	0.361	0.357	0.359	0.360

Note: Standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2. The Hedonic Wage Model

First, a separate hedonic wage model (equation 14) was estimated using data from 2016 and 2018, respectively. Then, the hedonic wage model was re-estimated with pooled data from these two years. The hedonic wage model regression results are shown in Table 3.

From Table 3, we can see that, in both years, workers with higher job risks are compensated with higher wages, indicating that risk is a disamenity. Similarly, July temperature is a disamenity, and workers engaged with higher July temperature are paid more. At the same time, January's temperature is an amenity, and workers are willing to sacrifice some wages for

Table 4. MWTP for risk reduction (safety) by the magnitude of 1/1000,000 (unit: USD)

		MWTP		
		Mean	l1	u1
ML	Model 1	1885	1537	2233
	Model 2	2152	1798	2506
PML	Model 3	3022	2739	3305
	Model 4	2963	2672	3253
HWM	FE	25	13	36

a warmer winter. Both industry unemployment and workhour have a negative and significant coefficient. The workers' individual attributes have expected effects on wages. For example, senior married male workers with a higher education (or the respondent's father has a higher education, higher *feduyear*) receive higher wages, and workers who work in firms with more employees have higher wages, other factors being equal. We can see that the regression results from the fixed effect model and pooled_2 are very close to each other and equally good.

Given the estimation results of the hedonic wage model, we can measure the marginal willingness to pay for one unit of risk reduction (1/1000,000) as follows:

$$MWTP = \hat{\pi}_1 \times \overline{wage}$$

where $\hat{\pi}_1$ is the estimated coefficient of risk.

The estimated MWTP from both discrete choice models and hedonic wage models are calculated and transferred to US dollars at the exchange rate of 30 TWD: 1 USD, as shown in Table 4.

The MWTP estimated from the two mixed logit models are comparable to each other but are significantly larger than the MWTP from the hedonic wage model. In fact, the MWTP estimated from the discrete choice model is unreasonably larger than any related studies in Taiwan or abroad. For example, Zhang et al. (2022) found a VSL of \$14 million based on Taiwan labour market data during 1999 and 2014. Banzhaf (2022) conducted a meta-analysis using the results of five meta-analyses about the US and found a median VSL of \$7 million, with a 90% confidence interval of \$2.4 to \$11.2 million. Based on the stated preference method, the OECD estimate of VSL is about \$4 million (OECD, 2012). Giergiczny, M. (2008)

estimated the VSL for Poland and found a value between \$0.79 and \$2.41 million. VSLs estimated in mid- or low-income countries are generally lower.

Given the linear utility function assumed in the discrete choice model, the nonmarginal value of risk reduction can be calculated by multiplying the changes in risks by the MWTP; thus, the unreasonably large MWTP from the discrete choice model also implies an unreasonably large nonmarginal value of risk reduction. Therefore, we believe the hedonic wage model outperforms the discrete choice model in evaluating safety with labour market data.

One possible explanation for the surprisingly large MWTP estimates by the discrete choice model stems from the well-acknowledged non-existence of moments of a ratio distribution resulting from dividing two normally distributed variables. As we know, the MWTP of risk is measured as the absolute value of the marginal utility of risk (β in equation (8)) divided by the marginal value of money (α in equation (8)). When the distribution of the denominator (α) spans zero, the WTP ratio distribution may be undefined. Carson and Czajkowski (2019) suggest that restricting the coefficient of the cost/income variable (in this case, wage) to be random and lognormally distributed could help solve the problem of non-existent moments in the WTP ratio. However, setting the coefficient of wage to be random could cause non-convergence problems. Unfortunately, the Carson and Czajkowski approach does not work in this study since the mixed logit models fail to converge and thus are not reported here.

Another possible explanation of the extremely large MWTP from the discrete choice model is that job choices are actually not very responsive to wages. In our discrete choice model results, the coefficients of wages are quite small. Though they are statistically significant, they suggest that wages do not have a large effect on workers' industry choices. In our case, the wage varies across individuals rather than industries, which could cause the workers' industry choices to be insensitive to wages.

5. Discussion and Conclusion

5.1. Discussion

The two approaches differ in their underlying assumptions and theoretical foundations. The hedonic wage method assumes a competitive market with

numerous heterogeneous firms and consumers, while the discrete choice model typically assumes a known utility function that is linear in individual and job or housing attributes. The hedonic model results from consumers' utility maximisation and firms' profit maximisation, while the random utility model focuses solely on consumer behaviour (Jonker, 2001). Despite these differences, there is a connection between the two models. Wong (2010) demonstrates a duality between the hedonic price model and the discrete choice model: the gradient of the hedonic price function represents the ratio of a weighted average of individuals' marginal utilities, where the weights are derived from choice probabilities in the discrete choice logit model.

Compared to the hedonic wage model, the discrete choice model has two widely recognised disadvantages, aside from the issue of non-existent moments in the WTP ratio distribution. First, the hedonic price equation reflects market equilibrium that balances consumers' willingness to pay and firms' marginal costs, while the discrete choice model only captures the demand side. This may explain why the discrete choice model often yields higher WTP estimates than the hedonic price model. For example, Palmquist and Israngkura (1999) analyse MWTP for air quality improvements using the same housing market data with both models. They find that the discrete choice model produces larger estimates. Similarly, Sinha et al. (2021) compare the two approaches in evaluating MWTP for climate attributes and find that the discrete choice model generates MWTP estimates for warmer winters that are twice as large as those from the hedonic price model.

A second disadvantage of the discrete choice model is that it relies heavily on assumptions about the utility function form and the distribution of the random term, which are, unfortunately, restrictive and hard to test (Freeman et al., 2014). Moreover, most studies adopting the discrete choice model in environmental goods valuation (Chay and Greenstone, 2005; Bayer et al., 2009; Freeman et al., 2019) have adopted a multinomial logit estimation procedure that is subjected to the IIA property. The IIA property imposes a strict substitution pattern so that introducing a third option does not change the ratio of the choice probabilities between two alternatives – a condition rarely met in real-world scenarios.

5.2. Conclusion

This is the first study to apply the discrete choice model to real labour market data to investigate the value of job safety and compare the results with those from the hedonic wage model. This study finds that the hedonic wage model performs better than the discrete choice model in modelling real labour market decisions and provides a more reliable marginal willingness to pay (MWTP) for job safety. Future research should further explore the comparison between these two approaches using real market data or simulation in both the housing and labour markets regarding both the marginal value and nonmarginal value.

Acknowledgements

I sincerely thank Botao Qin, Carlo Fezzi, Hui Mao, and Yongkun Yin for their insightful comments and suggestions for this paper.

References

- Banzhaf, H. S. (2022). The Value of Statistical Life: A Meta-Analysis of Meta-Analyses. *Journal of Benefit-Cost Analysis*, 13(2), 182–197. <https://doi.org/10.1017/bca.2022.9>
- Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and Hedonic Valuation: The Case of Air Quality. *Journal of Environmental Economics and Management*, 58(1), 1–14. <https://doi.org/10.1016/j.jeem.2008.08.004>
- Carson, R. T., & Czajkowski, M. (2019). A New Baseline Model for Estimating Willingness to Pay from Discrete Choice Models. *Journal of Environmental Economics and Management*, 95, 57–61. <https://doi.org/10.1016/j.jeem.2019.03.003>
- Chay, K. Y., & Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, 113(2), 376–424. <https://www.jstor.org/stable/10.1086/427462>
- Cragg, M., & Kahn, M. (1997). New Estimates of Climate Demand: Evidence from Location Choice. *Journal of Urban Economics*, 42(2), 261–284. <https://doi.org/10.1006/juec.1996.2027>

- Cropper, M. L., Deck, L., Kishor, N., & McConnell, K. E. (1993). Valuing Product Attributes Using Single Market Data: A Comparison of Hedonic and Discrete Choice Approaches. *The Review of Economics and Statistics*, 225–232.
- Freeman, R., Liang, W., Song, R., & Timmins, C. (2019). Willingness to Pay for Clean Air in China. *Journal of Environmental Economics and Management*, 94, 188–216. <https://doi.org/10.1016/j.jeem.2019.01.005>
- Giergiczny, M. (2008). Value of a Statistical Life—The Case of Poland. *Environmental and Resource Economics*, 41, 209–221. <https://doi.org/10.1007/s10640-007-9188-2>
- Guo, J., Feng, T., & Timmermans, H. J. P. (2020). Co-Dependent Workplace, Residence and Commuting Mode Choice: Results of a Multi-Dimensional Mixed Logit Model with Panel Effects. *Cities*, 96, 102448. <https://doi.org/10.1016/j.cities.2019.102448>
- Guo, X., & Hammitt, J. K. (2009). Compensating Wage Differentials with Unemployment: Evidence from China. *Environmental and Resource Economics*, 42, 187–209. <https://doi.org/10.1007/s10640-008-9217-9>
- Hintermann, B., Alberini, A., & Markandya, A. (2010). Estimating the Value of Safety with Labour Market Data: Are the Results Trustworthy? *Applied Economics*, 42(9), 1085–1100. <https://doi.org/10.1080/00036840802260940>
- Jonker, N. (2001). Constructing Quality Adjusted Price Indexes: A Comparison of Hedonic and Discrete Choice Methods. *De Nederlandsche Bank, Econometric Research and Special Studies Department*. <https://dx.doi.org/10.2139/ssrn.358104>
- Kniesner, T. J., Viscusi, W. K., Woock, C., & Ziliak, J. P. (2012). The Value of a Statistical Life: Evidence from Panel Data. *Review of Economics and Statistics*, 94(1), 74–87. https://doi.org/10.1162/REST_a_00229
- Lagarde, M., & Blaauw, D. (2009). A Review of the Application and Contribution of Discrete Choice Experiments to Inform Human Resources Policy Interventions. *Human Resources for Health*, 7, 1–10. <https://doi.org/10.1186/1478-4491-7-62>
- Lim, H.-E., & Jan-Jan, S. (2006). Job Selection Criteria and Job Sector Preference of Economics Student: An Ordered Logit Model Analysis. *International Journal of Business and Society*, 7(1), 53.
- Mangham, L. J., Hanson, K., & McPake, B. (2009). How to Do (or Not to Do)... Designing a Discrete Choice Experiment for Application in a Low-Income Country. *Health Policy and Planning*, 24(2), 151–158. <https://doi.org/10.1093/heapol/czn047>
- Manski, C. F., & Lerman, S. R. (1977). The Estimation of Choice Probabilities from Choice Based Samples. *Econometrica: Journal of the Econometric Society*, 1977–1988. <https://doi.org/10.2307/1914121>
- McFadden, D. (1974). The Measurement of Urban Travel Demand. *Journal of Public Economics*, 3(4), 303–328. [https://doi.org/10.1016/0047-2727\(74\)90003-6](https://doi.org/10.1016/0047-2727(74)90003-6)
- McFadden, D. (1980). Econometric Models for Probabilistic Choice Among Products. *Journal of Business*, S13–S29.
- McFadden, D., & Train, K. (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, 15(5), 447–470. [https://doi.org/10.1002/1099-1255\(200009/10\)15:5%3C447::AID-JAE570%3E3.0.CO;2-1](https://doi.org/10.1002/1099-1255(200009/10)15:5%3C447::AID-JAE570%3E3.0.CO;2-1)
- OECD. (2012). *Mortality Risk Valuation in Environment, Health and Transport Policies*. OECD Publishing.
- Palmquist, R. B., & Israngkura, A. (1999). Valuing Air Quality with Hedonic and Discrete Choice Models. *American Journal of Agricultural Economics*, 81(5), 1128–1133. <https://doi.org/10.2307/1244096>
- Sinha, P., & Cropper, M. L. (2013). The Value of Climate Amenities: Evidence from US Migration Decisions. (No. w18756). National Bureau of Economic Research.
- Sinha, P., Caulkins, M., & Cropper, M. (2021). The Value of Climate Amenities: A Comparison of Hedonic and Discrete Choice Approaches. *Journal of Urban Economics*, 126, 103371. <https://doi.org/10.1016/j.jue.2021.103371>
- Smits, M.-F., Witter, S., Lemiere, C., Hoang-Vu Eozenou, P., Lievens, T., Zaman, R. U., Engelhardt, K., & Hou, X. (2016). Understanding Health Workers' Job Preferences to Improve Rural Retention in Timor-Leste: Findings from a Discrete Choice Experiment. *PloS ONE*, 11(11), e0165940. <https://doi.org/10.1371/journal.pone.0165940>
- Thaler, R., & Rosen, S. (1976). The Value of Saving a Life: Evidence from the Labour Market. In *Household Production and Consumption* (pp. 265–302). National Bureau of Economic Research.
- Wong, M. (2010). The Relationship between Marginal Willingness-to-Pay in the Hedonic and Discrete Choice Models. *MPRA Working Paper*.

Zhang, N., Shaw, D., & Lin, C. Y. (2022). Implicit Prices of Job Risk, Climate, and Air Pollution: Evidence from Taiwan. *Climate Change Economics*, 13(4), 2250007. <https://doi.org/10.1142/S2010007822500075>

Zhang, N., Mendelsohn, R., & Shaw, D. (2023). How to Identify and Estimate the Demand for Job Safety? *MPRA Working Paper*.

Appendix

Table A1. Wage Regression by Industry (2016)

	Manufactory	Construction	Catering	Transportation	Finance	Science	Service
female	-0.272*** (0.0232)	-0.198** (0.0785)	-0.201*** (0.0337)	-0.242*** (0.0798)	-0.0981 (0.0703)	-0.161*** (0.0575)	-0.179*** (0.0575)
eduyear	0.100*** (0.0059)	0.0347** (0.0169)	0.0769*** (0.0101)	0.0495** (0.0216)	0.113*** (0.0193)	0.0915*** (0.0140)	0.0501*** (0.0130)
lnwexp	0.0990*** (0.0212)	0.119* (0.0633)	0.0923** (0.0439)	0.0393 (0.0621)	0.193** (0.0838)	0.150*** (0.0423)	0.0162 (0.0430)
healthlevel	0.00211 (0.0144)	0.0684* (0.0395)	0.0174 (0.0204)	0.0403 (0.0420)	-0.0068 (0.0445)	-0.0373 (0.0448)	0.0357 (0.0279)
age	0.0088*** (0.0028)	0.0008 (0.0054)	0.0058 (0.0046)	0.0163* (0.0089)	0.0132* (0.0078)	0.0001 (0.0050)	0.0150** (0.006)
_cons	4.27*** (0.120)	5.21*** (0.368)	4.60*** (0.188)	4.82*** (0.532)	3.94*** (0.391)	4.66*** (0.326)	4.83*** (0.252)
N	1008	129	495	117	194	167	459
R ²	0.390	0.122	0.206	0.186	0.283	0.301	0.117

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A2. Wage Regression by Industry (2018)

	Manufactory	Construction	Catering	Transportation	Finance	Science	Service
female	-0.243*** (0.0282)	-0.175** (0.0828)	-0.240*** (0.0399)	-0.211** (0.0985)	-0.102 (0.0740)	-0.105* (0.0548)	-0.231*** (0.0535)
eduyear	0.110*** (0.0066)	0.0622*** (0.0155)	0.0788*** (0.0110)	0.0730*** (0.0271)	0.096*** (0.0197)	0.110*** (0.0152)	0.0611*** (0.0115)
lnwexp	0.0804** (0.0335)	0.256*** (0.0704)	0.211*** (0.0567)	0.313*** (0.101)	0.178** (0.0885)	0.189*** (0.0573)	-0.007 (0.0503)
healthlevel	0.0183 (0.0155)	0.00363 (0.0405)	0.0606** (0.0278)	0.159*** (0.0564)	0.0279 (0.0408)	0.0310 (0.0413)	0.0277 (0.0255)

Table A2. Wage Regression by Industry (2018)
Continued

	Manufactory	Construction	Catering	Transportation	Finance	Science	Service
age	0.0117*** (0.0031)	-0.0049 (0.0058)	-0.0017 (0.006)	-0.0134 (0.0092)	0.0099 (0.0079)	0.0012 (0.0041)	0.0098 (0.0062)
_cons	4.07*** (0.140)	4.96*** (0.291)	4.40*** (0.246)	4.53*** (0.621)	4.19*** (0.425)	4.03*** (0.363)	4.95*** (0.254)
N	937	125	394	96	177	158	395
R ²	0.363	0.188	0.248	0.215	0.173	0.371	0.142

Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3. Summary Statistics of Job Attributes and Worker Attributes by Industry

Variables	Definition	Manufac- ture Obs. 2,257	Construc- tion Obs. 315	Catering Obs. 1,055	Transpor- tation Obs. 263	Finance Obs. 418	Science Obs. 413	Service Obs. 1,204
wage_alt	yearly wage income, unit: 1000 TWD in 2014 values	500.53 (264.22)	545.57 (154.25)	454.33 (147.79)	519.98 (167.09)	607.97 (228.02)	492.48 (175.29)	517.53 (204.61)
risk_alt	industry mortality rates, unit: 1/1,000,000	23.70 (0.00)	114.50 (8.57)	11.32 (1.65)	48.47 (0.20)	4.03 (1.32)	25.17 (3.64)	8.75 (2.66)
workhour_alt	industry average monthly total work hours, unit: hour	175.74 (0.30)	166.13 (0.15)	162.34 (1.54)	172.74 (0.30)	167.05 (1.09)	172.26 (0.30)	161.10 (1.49)
umem_alt	industry unemployment rates, unit: %	2.84 (0.02)	4.58 (0.00)	3.53 (0.09)	3.19 (0.00)	2.83 (0.32)	4.31 (0.20)	2.69 (0.14)

Note: standard deviation in parentheses.

wage_alt is constructed as follows: for the industries that are chosen by workers, *wage_alt* is their real wage reported in the PSFD, for industries not chosen, *wage_alt* is predicted from the first stage of wage regression by industry.