Comparison of Approaches to Modelling Field Effects: Tohoku Evacuation Case Study

Jun JI[1], Jan-Dirk SCHMÖCKER [2], Prawira Fajarindra BELGIAWAN[3], Hiroki YAMAZAKI[4], and Nobuhiro UNO[5]

[1]Japan radio co., Ltd.
(Nakano Central Park East, 4-10-1Nakano, Nakano-ku, Tokyo 164-8570)
E-mail: ji.jun@jrc.co.jp
[2]Member of JSCE, Associate Professor, Dept. of Urban Management., Kyoto University
(C1-2-436, Kyoto daigaku-Katsura, Nishikyo-ku, Kyoto 615-8540)
E-mail: schmoecker@trans.kuciv.kyoto-u.ac.jp
[3]Researcher, Dept. of Urban Management., Kyoto University
(C1-2-437, Kyoto daigaku-Katsura, Nishikyo-ku, Kyoto 615-8540)
E-mail: belgiawan@trans.kuciv.kyoto-u.ac.jp
[4]Member of JSCE, Regional Futures Research Center
(1-5-17, Dojima, Kita-ku, Osaka 530-0003)
E-mail: yamazaki@refrec.jp
[5]Member of JSCE, Associate Professor, Dept. of Urban Management., Kyoto University
(C1-2-436, Kyoto daigaku-Katsura, Nishikyo-ku, Kyoto 615-8540)
E-mail: uno@trans.kuciv.kyoto-u.ac.jp

In the discrete choice literature field effects have recently attracted increased attention as a number of studies have demonstrated their importance to explain behavioural decisions also regarding travel behaviour. Studies have shown that mode choice and tendency of illegal parking are better explained if behaviour of others is considered. We test for field effects in the case for evacuation decisions, where they might be particular important. We discuss and compare formulations as common 2-stage model and one with BLP correction (Berry, Levinshon and Pakes method).

Key Words : *evacuation decision, field effect, correcting for endogeneity*

1. INTRODUCTION

In the discrete choice literature field effects have recently attracted increased attention as a number of studies have demonstrated their importance to explain behavioural decisions also regarding travel behaviour. Previous studies have shown that mode choice[1)2)] and illegal parking[3)] are better explained if behaviour of others is considered.

The pressure to adjust to other social influences, is also commonly referred to as norming effects. Schmöcker et al.[4)] use the term “mass effect” as a geneal term to describe the positive influence to adjust one’s choice to be in line with the observed choice of others. Dugundji and Walker[2)] use the term “field effect” to capture social influence in a discrete choice model. Goetzke and Rave[5)] use the term “social network effect” to incorporate norming effect into travel demand model. Abou-Zeid et al.[5)] discuss several further closely related terms such as herd behavior, peer effects, conformity, or fashion.

The norming effect mentioned in above literature are utilized to describe non-urgent, recurrent and/or long term decision making. However, also in urgent and unfamiliar situations such as in disaster situations the norming effect might become important for individual decisions. Teo et al.[7)] incorporate hypothesized norming effects (or in their term “compliance behavior”) in an agent-based simulation and suggest that it can influence the performance of a network in an evacuation situation significantly.

Inspired by the work of Teo et al.[7)] we aim to incorporate field effects in a discrete choice model for evacuation decisions using observed data. We realize that the incorporation of field effect itself raises some endogeneity issues (this is discussed in section 2) therefore several methods to correct the endogeneity are introduced, which is the focus of this paper.

The structure of this paper is as follows: After this introduction, the second section of this paper will discuss previous research on the incorporation of field effects and some methods to correct for endogeneity. In Section 3 we will discuss first characteristics of our study area, Tohoku, Japan, before explaining the survey among our respondents regarding their evacuation decision. After that we introduce the basic evacuation choice model (Section 4). In the following section we then advance the model by presenting models that include different formulations of the field effect. Finally, in Section 6 some conclusions and possible future research directions are outlined.

2. LITERATURE REVIEW

In the estimation of the behavioral models, there is potential endogeneity in incorporating field effect variables. The endogeneity problem arises when an explanatory variable is correlated with the unobserved factors, and such a situation leads to biased and inconsistent parameter estimates. The failure in capturing endogeneity usually is caused by errors in variables, simultaneous determination, and omitted attributes, among other causes.

The problem of dealing with field effect in discrete choice modelling is highlighted by Dugundji and Walker[2]. In their work, the average zonal model split is used as a proxy for field effect. Their starting point was the discussions related to social interactions in binary discrete choice model by Aoki[8], Brock and Durlauf[9] and Blume and Durlauf[10]. According to those three papers, the choice of a given agent for a particular alternative depends on the overall share of decision makers who choose that alternative, this is called "global interactions". According to Walker et al.[11] a field effect defined globally would be perfectly correlated with a set of alternative-specific constant therefore to capture social effect without the problem of perfect correlation with alternative-specific constant it is better to use "non-global interaction". This means that an agent's choice is influenced by only a subset of decision makers. Goetzke and Rave[6] tested similar field effects for explaining the modal share of bicycles in German cities. Their approach is an extension of Dugundji and Walker[2] in which they take the average model split in each municipality in German as a proxy to field effect.

Goetzke and Andrade[12] instead introduced a spatially autoregressive 2-stage least square method (2-SLS) instrumental variable method using a heteroscedasticity-corrected linear probability model for binary discrete choice models of walking. They found that the instrumental variable method works for discrete choice models.

Another attempt to correct for endogeneity is by utilizing the Berry Levinsohn Pakes (BLP) method[13)14)] for mode choice decision[11]. The BLP method removes the endogeneity from the choice model through the use of market-specific constants. Then by using instrumental variables in a linear regression, consistent estimates of the social influence effect are obtained. This consistent estimate of field effect parameters is then reintroduced in the choice model.

Despite these recent publications, endogeneity in discrete choice appears to be not yet studied well enough. The range of choice situations studies so far is limited. Further, in particular a systematic comparison of the different approaches appears to be not yet done.

In this paper we aim to narrow this gap by attempting to correct for the endogeneity of incorporating field effect in a choice model on evacuation decisions by using both the 2-SLS and BLP methods.

3. THE GREAT EAST JAPAN EARTHQUAKE DATA

(1) Data source

The Great East Japan earthquake was a magnitude 9.0 earthquake off the coast of Japan that occurred at 14:46 JST on March 11th, 2011. It was the most powerful earthquake ever recorded to have hit Japan, and the fourth most powerful earthquake in the world since modern record keeping began in 1900. The earthquake triggered powerful Tsunami waves that reached heights of up to 40.5 meters in Miyako in Tohoku's Iwate Prefecture and in the Sendai area Tsunami waves travelled up to 10 km inland.

The data used for this research are taken from the Reconstruction assistance survey archive website (http://fukkou.csis.u-tokyo.ac.jp) which is operated by the Center for Spatial Information Science (CSIS) at The University of Tokyo. The data we utilize in this study are based on a survey from the Ministry of Land, Infrastructure and Transport (MLIT). The survey was carried out in the form of a questionnaire survey by MLIT in Aug 2011. Around 10,603 samples from six prefectures in the hit region namely Aomori prefecture, Iwate prefecture, Miyagi prefecture, Fukushima prefecture, Ibaraki prefecture and Chiba prefecture were collected. After some data cleaning, 10,384 observations and 29 potential explanatory variables were extracted. Those variables can be grouped into sociodemographic ones (age, gender, occupation), their preparedness for the Tsunami (whether the person knows evacuation routes, knows the shelter location, has participated in Tsunami drills, etc.), whether and when the person has obtained a Tsunami warning, as well as the person's and his family's location during the earthquake.

(2) Data preparation

The dependent variable for our model is evacuation decision with binary responses, 1 if the respondents made a decision to evacuate, and 0 if otherwise. At first, to reduce the number of variables that can explain the decision to evacuate or not, we performed a Pearson correlation analysis to check for strong correlation between explanatory variables. Based on thi, we conducted a principal component analysis with varimax rotation to convert correlated variables into uncorrelated ones.

In the end, we obtain follwing variables that have no significant correlation with each other which are: “gender”, (hearing) “tsunami warning”, “family location duing earthquake (home, work, kindergarten), “living near harbour”, “preparation”, and “seen sign”. “Preparation” is a factor constructed by seven variables that correspond to preparation before tsunami, for example check location and participation in drills. “Seen sign” is a factor constructed by four variables including whether the person was aware of signboards with inundation warnings.

4. EVACUATION CHOICE MODEL WITHOUT FIELD EFFECT

In the following we describe how we estimate a base model for evacuation decisions. Let U_{in_m} denote the unobserved utility of person *n* in city *m* to choose option *i*, that is whether to evacuate or note. Further, let $P_{n_m}(i)$ denote the probability of person *n* in city *m* to choose option *i*. Assuming a logit distribution of the error terms we obtain:

$$P_{n_m}(i) = \frac{e^{V_{in_m}}}{e^{V_{in_m}} + e^{V_{jn_m}}} \tag{4a}$$

Our task is hence to obtain the observed part of the utility function by maximum likelihood observations. Ignoring any field effect we obtain

$$U_{in_m} = V\left(s_{n_m}; r_m; \beta\right) + \varepsilon_{in_m} \tag{4b}$$

In our case the dependent variable U_{in_m} describe the perceived utility to evacuate after the earthquake and possibly hearing a Tsunami warning. This standard binary binary logistic regression model was performed with person specific dependent variable S_{n_m} and city specific dependent variable r_m explaining the differences in evacuation decisions.

Person specific dependent variable S_{n_m} were gender, heard tsunami warning (from government), family location, living near harbour or not (individual), preparation before tsunami and having seen the tsunami sign or not.

City specific dependent variable r_m were tsunami warning time (after earthquake happened since 14:46), Population for each city (10 thousands), population density (number of people in every km^2 land area), flooded area density (flooded area in every km2 land area), harbour city or not (city) from 41 cities samples (The list of cities can be seen in Table 4).

Table 1 Base model result

Variable	Estimate	Stdz	t-stat
Alternative Specific Constant	**1.26**		**18.33**
Person specific			
Male	-0.05	-0.11	-1.88
Heard Tsunami Warning	**0.39**	**0.86**	**13.74**
Family at home	**0.13**	**0.15**	**2.57**
Family at work	**-0.11**	**-0.13**	**-2.15**
Family at kindergarten	0.19	0.11	1.73
Live near harbor	**0.24**	**0.29**	**4.20**
Preparation	**0.28**	**0.67**	**9.49**
Seen sign	**0.10**	**0.23**	**3.15**
City specific			
Tsunami warning time	**-0.01**	**-0.21**	**-3.94**
Population density	-0.01	-0.11	-1.72
Flooded area density	0.01	0.12	1.81
Harbor city [dummy]	0.00	0.09	1.56
Model Summary			
Sample Size N	10384		
AIC	10620		
R-squared	0.05		

Note: Bold p value<0.05; Italic p value <1; Stdz: standardized value

The model results are shown in Table 1. Parameter estimates show that tsunami warning has a positive significant impact on evacuation decisions, which means that people who heard the tsunami warning are more likely to evacuate. People who have family at home are also more likely to evacuate, especially compared to those who have family at work. People who live near the harbour, have prepared before tsunami and seen signboard information also have positive significate impact on evacuation decisions.

Tsunami warning time from government has a negative significant effect to evacuation decisions. This indicates that the quicker people obtain the warning from government after the earthquake the more people are likely to evacuate.

To allow comparability between parameter estimates we standardized them as can be seen in the middle column of each model. After standardizing, we observe that tsunami warning is the most important factor for evacuation decisions. Another variable that is also important is preparation before tsunami. For the case of tsunami warning time, if we look at the standardized result, we can see that with the increase of one standard deviation of tsunami warning time, which is 7 minutes, people are 21% less likely to evacuate. In other words, each minute the tsunami warning time arrives later, means that on average 3% less of the population are evacuating.

5. EVACUATION CHOICE MODEL WITH FIELD EFFECT

We now hypothesize that evacuation is contagious, that is, we presume that some people evacuating will affect further people to evacuate.

Here we introduce the use of a field effect variable to capture social influences in evacuation decision model, and then describe the issue of endogeneity that arises and propose a correction.

(1) Naïve model

Such a model might hence be defined as

$$U_{in_m} = V(s_{n_m}; r_m; \beta) + \gamma F_{im} + \varepsilon_{in_m} \tag{5a}$$

where F_{im} describes the percentage of persons choosing to evacuate in city m and γ denotes the "strength of influence" or field effect. The field effect that we introduce here has the potential of endogeneity, however in this model we do not perform any attempt to correct or reduce the endogeneity issues thus this model is called the Naïve model.

The result shown in Table 2 tells that Field effect has a positive significant impact on evacuation decision, which means that the more people evacuate in a city the more people are likely to evacuate. We note that in this model living near the harbor and the Tsunami warning time become not significant.

Table 2 Naïve model result

Variable	Estimate	Stdz	t-stat
Alternative Specific Constant	**-1.01**		**-4.76**
Person specific			
Male	-0.05	-0.11	-1.85
Heard Tsunami Warning	**0.35**	**0.77**	**12.25**
Family at home	**0.12**	**0.13**	**2.25**
Family at work	-0.09	-0.10	-1.75
Family at kindergarten	**0.22**	**0.12**	**1.99**
Live near harbour	0.04	0.05	0.71
Preparation	**0.28**	**0.65**	**9.13**
Seen sign	0.06	0.14	1.94
City specific			
Tsunami warning time	0.00	0.04	0.63
Population density	0.00	0.01	0.20
Flooded area density	0.00	-0.02	-0.23
Harbor city [dummy]	0.00	0.04	0.62
Field effect			
Percent evacuation in city	**3.14**	**0.78**	**11.23**
Model Summary			
Sample Size N	10384		
AIC	10532		
R-squared	0.06		

Note: Bold p value<0.05; Italic p value <1; Stdz: standardized value

For other explanatory variables, similar with the base model, looking at the standardized result, the variables of heard tsunami warning followed by preparation are still influential. But compare to the standardized value of field effect those variables are less influential. The standardized value of field effect with 0.78 is the largest compared to all other variables.

The problem of this model is that F_{im} will be correlated with both s_{n_m} and r_m, which is likely to lead to an upward biased estimation of γ and too low estimate for parameter β. Therefore in the sub-sequent model we utilize a method to solve this endogeneity issue.

(2) Two stage model

To account for these correlations, following Goetzke and Andrade[12] one approach would be to use a 2-stage model.

The first step is a binary logit model with

$$d_{n_m} = \theta^T s_{n_m} + \dot{\varepsilon}_{in_m} \tag{5b}$$

Where d_{n_m} takes the value of one if the person *n* in city *m* is deciding to evacuate and zero otherwise.

The parameters $\boldsymbol{\theta^T}$ are then used to obtain the fitted field effect for decision *i* = {evacuate} as the expected decision by firstly obtaining the estimated aggregate $P_n(i|\boldsymbol{\theta^T})$ using all person-specific variables. These probabilities are subsequently used to obtain the fitted value of evacuation decisions which is explained in Eq. 5c:

$$\hat{F}_{im} = \sum_{n \in m} P_n(i|\boldsymbol{\theta^T}) \tag{5c}$$

Then in step 2 this estimation is inserted in the binary choice model:

$$U_{in_m} = V(\boldsymbol{s}_{n_m}; \boldsymbol{r}_m; \beta) + \gamma \hat{F}_{im} + \varepsilon_{in_m} \tag{5d}$$

Eq. (5d) ties together the person specific variable with the city-specific variable and the social effects.

The result for this model can be seen in Table 3. Compared to the naïve model, seen sign information and flooded area density become significant. This might be because the endogeneity problem is corrected to some extent, however, in the first step the area specific variables are omitted in order to reduce multi-collinearity and to be able to distinguish geographic characteristics. That is $\hat{F}_{im}$ might be interpreted as the "expected evacuations independent of the geographic characteristics of the city"; these are controlled for separately with $\boldsymbol{r}_m$. Since the instrument $\hat{F}_{im}$ is constructed as in Eq. (5c) the endogeneity problem with respect to the personal characteristics $\boldsymbol{s}_{n_m}$ might be indeed reduced, though not completely. We therefore expect that in Eq. (5d), compared to Eq. (5a), the parameter γ is corrected downward but still possibly overestimated.

Looking at the standardize value for comparison, we can see that the influence of the field effect is indeed reduced to more than half compared to the naïve model. This makes the variable "heard tsunami warning" again the most influential variable for determining the decision to evacuate followed by preparation. Interesting to note is that the city specific tsunami warning becomes significant again.

Table 3 Two stage model result

Variable	Step 1		Step 2		
	Est	t-stat	Est	Stdz	t-stat
ASC	**1.17**	**20.6**	**1.10**		**13.0**
Person specific					
Male	**-0.05**	**-2.2**	*-0.04*	*-0.10*	*-1.7*
Heard Tsunami Warning	**0.37**	**13.4**	**0.39**	**0.87**	**13.9**
Family at home	**0.13**	**2.5**	**0.14**	**0.16**	**2.7**
Family at work	**-0.11**	**-2.2**	**-0.10**	**-0.12**	**-2.0**
Family at kindergarten	*0.20*	*1.8*	*0.21*	*0.12*	*1.9*
Live near harbor	**0.28**	**5.2**	**0.26**	**0.31**	**4.5**
Preparation	**0.30**	**10.3**	**0.27**	**0.65**	**9.2**
Seen sign	**0.10**	**3.5**	**0.10**	**0.23**	**3.2**
City specific					
Tsunami warning time			**-0.01**	**-0.16**	**-2.77**
Population density			0.00	-0.07	-1.00
Flooded area density			*0.01*	*0.13*	*1.88*
Harbor city [dummy]			0.00	-0.13	-1.47
Field effect					
Fitted evacuation (sum in city)			**0.21**	**0.30**	**3.30**
Model Summary					
Sample Size N	10384		10384		
AIC	10678		10651		
R-squared	0.05		0.05		

Note: Bold p value<0.05; Italic p value <1; Stdz: standardized value

(3) BLP Correction

An alternative approach would be the BLP approach described in Walker et al.[11]. The BLP procedure involves decomposing the error into two parts: the endogenous causing part and the random portion. The evacuation model then becomes

$$U_{in_m} = V(\boldsymbol{s}_{n_m}; \boldsymbol{r}_m; \beta) + \gamma \mathrm{F}_{im} + \ddot{\varepsilon}_{\mathrm{im}} + \dot{\varepsilon}_{\mathrm{in_m}} \quad (5e)$$

Where $\ddot{\varepsilon}_{im}$ is correlated with F_{im} and $\dot{\varepsilon}_{\mathrm{in_m}}$ is uncorrelated with F_{im}, $\boldsymbol{s}_{n_m}$ and $\boldsymbol{r}_m$. To isolate the endogenous-causing components F_{im} and $\ddot{\varepsilon}_{im}$ the terms are thus rearranged as follows

$$U_{in_m} = [\gamma \mathrm{F}_{im} + \ddot{\varepsilon}_{\mathrm{im}}] + V(\boldsymbol{s}_{n_m}; \boldsymbol{r}_m; \beta) + \dot{\varepsilon}_{\mathrm{in_m}} \quad (5f)$$

The first term $[\gamma \mathrm{F}_{im} + \ddot{\varepsilon}_{\mathrm{im}}]$ represents the unobservable and observable components of utility relevant to the individuals peer group m. It represents the average, utility of a given choice in a given group. The error term $\dot{\varepsilon}_{\mathrm{in_m}}$ is orthogonal to all explanatory variables and varies across decision makers.

The trick in the BLP procedure is now replace the peer group effect with specific constants α_{im} for each alternative i and each peer city m, the new utility equation is

$$U_{in_m} = \alpha_{\mathrm{im}} + V(\boldsymbol{s}_{n_m}; \boldsymbol{r}_m; \beta) + \dot{\varepsilon}_{\mathrm{in_m}} \quad (5g)$$

where

$$\alpha_{\mathrm{im}} = \gamma \mathrm{F}_{im} + \ddot{\varepsilon}_{\mathrm{im}} \quad (5h)$$

These constants capture the average effects of the peer group. There is no endogeneity issue in the choice model as written this way, and therefore the parameters α_{im} and β are estimated via usual choice modeling procedures. We are interested in the social effect as represented by the parameter γ, which is not estimated via the choice model.

To estimate the parameter γ, the next stage of BLP is required, which is to estimate via linear regression the city-specific constants as explained by the field effect variable.

While the endogeneity issue remains (F_{im} is correlated with $\ddot{\varepsilon}_{\mathrm{im}}$), it is more straight forward to correct for endogeneity in the linear model. For this we use a two-stage instrumental variables approach.

In the first stage, the field effect variable F_{im} is regressed on the instrumental variables I_{im} (correlated with the field effect variable F_{im} and uncorrelated with the error $\ddot{\varepsilon}_{\mathrm{im}}$) as follows:

$$\mathrm{F}_{im} = \theta_i + \theta_F I_{\mathrm{im}} + v_{\mathrm{im}} \quad (5i)$$

where v_{im} is a random error (orthogonal to I_{im}) and θ_i and θ_F are estimated parameters.

In the second stage, the city-specific constants are regressed on the fitted value of the field effect from the first stage, i.e. $\hat{F}_{im} = \hat{\theta}_i + \hat{\theta}_F I_{\mathrm{im}}$, as follows:

$$\alpha_{\mathrm{im}} = \gamma_i + \gamma_F \hat{F}_{im} + \ddot{\varepsilon}_{\mathrm{im}} \quad (5j)$$

As $\hat{F}_{im}$ is orthogonal to $\ddot{\varepsilon}_{\mathrm{im}}$, this regression results is a consistent estimate of γ_F, which captures the effect of the field effect variable on the utility. This can then be inserted back into the choice model (replacing the city-specific constants with Eq.(5j) so that the choice model captures the effect of the peer city. Note that the entire right-hand side, including the fitted value of the error, is included in the final choice model.

In summary, the BLP process removes the endogeneity from the choice model via the use of city-specific constants. The endogeneity is then dealt with in a linear regression setting (with instrumental variables) to obtain consistent estimates of the social influence effect. This consistent estimate of the field effect parameter is then reintroduced to the choice model to obtain a choice model that captures social influences.

a) Instrumental variable

The BLP procedure requires hence the definition of an appropriate instrumental variable.

The instrument is defined as a variable that is correlated with the endogenous variable but uncorrelated with the error. In a spatial context with endogenous zonal variables, there are a set of natural instruments, which are the values of the endogenous variables in the spatially adjacent zones. These values are hence assumed to be correlated with the problem variable and uncorrelated with the error.

Table 4 Instrument for the endogenous field effect

City	% eva in city	Surrounding Cities	Instru-ment
Hashikami	0.52	Hachinohe, Yono-cho	0.71
Minamisanriku	0.82	Ishinomaki, Kesennuma	0.79
Shiogama	0.79	Shichigahama, Tagajo	0.71
Iwaki	0.61	Hirono-cho, Kitaibaraki	0.64
Hitachinaka	0.86	Oarai-machi	0.76
Oarai-machi	0.76	Hitachinaka	0.86
Kamisu	0.29	Kashima, Choshi	0.66
Higashimatsushima	0.60	Ishinomaki, Matsushima-machi	0.80
Onagawa	0.83	Ishinomaki	0.73
Tagajo	0.59	Shichigahama, Shiogama, Sendai	0.76
Natori	0.52	Iwanuma, Sendai	0.62
Soma	0.74	Minamisoma, Shinchi	0.67
Hirono-cho	0.68	Iwaki	0.61
Kitaibaraki	0.60	Iwaki	0.61
Ishinomaki	0.73	Higashimatsushima, Minamisanriku, Onagawa	0.75
Hachinohe	0.58	Misawa, Hashikami	0.53
Iwanuma	0.59	Watari-cho, Natori	0.60
Kashima	0.48	Kamisu	0.29
Choshi	0.84	Kamisu, Asahi	0.42
Asahi	0.54	Choshi, Sousa	0.79
Yono-cho	0.84	Kuji, Hashikami	0.70
Kuji	0.88	Yono-cho, Noda-mura	0.82
Noda-mura	0.80	Kuji	0.88
Tanohata	0.75		
Miyako	0.79	Yamada	0.82
Yamada	0.82	Miyako, Otsuchi-cho	0.81
Otsuchi-cho	0.83	Yamada, Kamaishi	0.82
Kamaishi	0.82	Otsuchi-cho, Ofunato	0.80
Ofunato	0.77	Kamaishi, Rikuzentakata	0.78
Kesennuma	0.85	Rikuzentakata, Minamisanriku	0.78
Hitachi	0.87		
Shichigahama	0.83	Shiogana, Tagajo	0.69
Shinchi	0.75	Yamamoto-cho, Soma	0.75
Watari-cho	0.67	Iwanuma, Yamamoto-cho	0.68
Misawa	0.55	Hachinohe	0.58
Rikuzentakata	0.74	Ofunato, Kesennuma	0.81
Matsushima-machi	0.86	Higashimatsushima	0.59
Yamamoto-cho	0.77	Shinchi, Watari-cho	0.71
Sendai	0.66	Natori, Tagajo	0.56
Minamisoma	0.60	Soma	0.73
Sousa City	0.73	Asahi City	0.54

Note: The order of the cities is random

The first part of this assumption, that the instrument is correlated with the problem variable, is explicit in our case because the evacuation in a zone is likely to be correlated with the evacuation in the adjacent zones. This can be explained by the spatial continuity of both the transportation network and social structure. The second part of the assumption, that the instrument is uncorrelated with the error, is hard to prove. In our case we need to rely on our zonal definitions that the boundaries of community-defined reference groups are meaningful. Walker et al.[11] presumed that the predominant social influences are coming from individuals within the decision-makers' postal code and is marginal for those from other areas. This study chooses to use the average evacuation of the surrounding cities (Fig.1) as the instrument for the endogenous field effect term (Table 4).

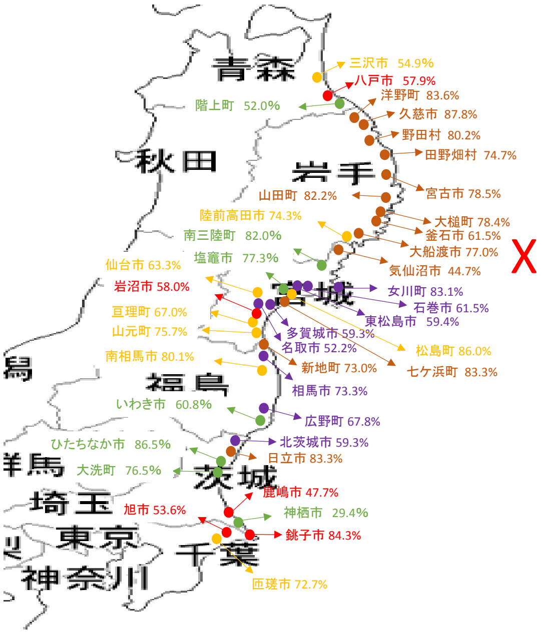

Fig.1 Proportion of evacuation in each city

b) Estimation Result

At the first stage, we run a choice model with incorporating city specific dummy to estimate constant variables for each person. There are 41 cities incorporated in the model thus we only use 40 city dummy variables with Sendai city as our reference. The result can be seen in Table 5 first step. In this model, we omitted city specific variables due to multicolineearity issues.

After obtaining 40 city specific constant, we perform a linear regression with country specific constants as the dependent variable and field effect as the explanatory variable. The field effect in this model is the uncorrected one that we use in the naïve model. The result can be seen in the top of Table 5. As we can see from the result, the field effect is significant with similar magnitude as the naïve model. We show this model in order to compare between the uncorrected endogeneity field effect and the corrected one.

To start with the BLP procedure, we perform the first stage of instrumental variable regression. In this model in order to correct the endogeneity of field effect variable, we use the uncorrected field effect in the naïve model as the dependent variable with the instrumental variable as the explanatory variable.

Our instrumental variable is the average evacuation decision of the adjacent location. Since there is a city which no adjacent location we add the instrumental variable dummy in which the response is 1 for a city with no adjacent city and 0 otherwise. The result can be seen in the middle part of Table 6 (model 2a).

Table 5 Evacuation model with BLP procedure

Variable	Step 1		Step 2		
	Est	t-stat	Est	Stdz	t-stat
Corrected constant	40 constant range from -0.49 to 2.07		40 constant range from -0.49 to 2.07		
Person specific					
Male	-0.03	-1.3	-0.03	-0.07	-1.3
Heard Tsunami Warning	**0.36**	**12.4**	**0.36**	**0.80**	**12.4**
Family at home	**0.24**	**4.8**	**0.24**	**0.27**	**4.8**
Family at work	0.02	0.4	0.02	0.02	0.4
Family at kindergarten	**0.25**	**2.2**	**0.25**	**0.14**	**2.2**
Live near harbor	0.22	1.5	0.22	0.26	1.5
Preparation	**0.28**	**9.1**	**0.28**	**0.66**	**9.1**
Seen sign	0.04	1.1	0.04	0.10	1.1
Field effect					
Fitted evacuation BLP			**No significant result as shown in Table 6 (Model 2-b)**		
Model Summary					
Sample Size N	10,384		10,384		
AIC	10,459.51		10,459.51		
R-squared	0.32		0.32		

Note: Bold p value<0.05; Stdz: standardized value

In this model, we found that the instrumental variable is significant only at the 10% level. Nevertheless, in the absence of any better field effect variable, we use this instrumental variable for the next step in which we calculate the fitted value for the field effect.

Table 6 Regression results for the city–specific constants

Variable	Estimate	T-stat
Model 1 (uncorrected spatial)		
dependent variable		
(City specific constant from choice model)		
Intercept	***-1.07***	***-3.40***
Field Effect	***3.11***	***7.01***
Model Summary		
Observation	40	
R-square	0.56	
Model 2a (corrected spatial)		
dependent variable (field effect)		
Intercept	***0.45***	***3.56***
Average Evacuation Adjacent Location (instrumental Variable)	0.34	1.93
Instrumental Variable dummy	***0.34***	***2.14***
Model Summary		
Observation	40	
R-square	0.11	
Model 2b (corrected spatial)		
dependent variable		
(City specific constant from choice model)		
Intercept	0.34	0.24
Fitted Field Effect	1.10	0.55
Model Summary		
Observation	40	
R-square	0.01	

Note: Bold p value<0.05; Italic p value <1

After obtaining the field effect, for the second stage we then perform a linear regression with city specific constants as the dependent variable. The independent variable is the fitted field effect. The result can be seen at the bottom of Table 6 (model 2b). We find that the fitted field effect is not significant. The reason for the insignificant result might be partly because we have not found an appropriate instrumental variable to correct for the endogeneity. But at the same time this also means that it is very difficult to explain whether a decision to evacuate partly because of influence of others. At present the result of our two stage model in Table 3, where we reduce the endogeneity effect, might be the closest one to explain the effect of others on evacuation decision.

6. CONCLUSION

As first attempt to incorporate "field effect", we perform a naïve model considering percentage of those who evacuate successfully in each city as proxy of field effect. Although this field effect variable has the significant impact to evacuation decisions, this variable might have endogeneity issues thus it is not easy to be measured.

In order to solve this issues, we perform a 2-stage model using Goetzke & Andrade's[12)] method and another model using BLP correction. Depending on the choice of model we find different answers on whether the field effect is indeed significant, highlighting the importance of carefully choosing the modelling approach. Taking all models results together, we suggest that the field effect will be somewhere between the estimates shown in Tables 5 and 6. That is the 2-stage approach might still slightly overestimate the field effect as shown when we fully correct for endogenity as in the BLP approach. However, we have limited trust in our BLP parameter estimates, due to the difficulty in finding a good instrumental variable.

Accordingly, data limitations are a primary shortcoming of this research and correspondingly, a good direction for future work. Geographically more detailed information about the respondents' location could help us to give more confident answers on the importance of the field effect.

REFERENCES

1) T. Morikawa. : A choice model with social interaction – application to voluntary restraint of car usage, *Proc. Of Int. Symp. on Structural change in transportation and communications in knowledge society: implications for theory, modeling and data,* Boston, MA, 1999.

2) Dugundji, E.R. and Walker, J. L. : Discrete choice with social and spatial network interdependencies: An empirical example using mixed generalized extreme value models with field and panel effects, *Transportation Research Record*, Vol. 1921, pp. 70-78, 2005.

3) Fukuda, D. and Morichi, S. : Incorporating aggregate behav-

ior in an individual’s discrete choice: An application to analyzing illegal bicycle parking behavior, *Transportation Research Part A*, Vol. 41, No. 4, pp. 313–325, 2007.
4) Schmöcker, J.-D., Hatori, T. and Watling, D. : Dynamic process model of mass effects on travel demand, *Transportation*, Vol. 41, No. 2, pp. 279-304, 2014.
5) Abou-Zeid, M., Schmöcker, J.-D., Belgiawan, P. F. and Fujii, S. : Mass effects and mobility decisions, *Transportation Letters*, Vol. 5, No. 3, pp. 115–130, 2013.
6) Goetzke, F. and Rave, T. : Bicycle use in Germany: Explaining differences between municipalities with social network effects, *Urban Studies*, Vol. 48, No. 2, pp. 427-437 , 2011.
7) Teo, J. S. E., Schmöcker, J.-D., Leon, F., Li, J. Y.-T., Ji, J., Atanasiu, G. and Taniguchi, E. : Agent-based evacuation model considering field effects and government advice, *Transportation Research Record*, Vol. 2532, pp. 129-140, 2015.
8) Aoki, M. : Economic fluctuations with interactive agents: dynamic and stochastic externalities, *Japanese Economic Review*, Vol. 46, No. 2, pp. 148-165, 1995.
9) Brock, W. A. and Durlauf, S. N. : Discrete choice with social interactions, *Review of Economic Studies*, Vol. 68, No. 2, pp. 235-260, 2001.
10) Blume, L and Durlauf, S. : Equilibrium concepts for social interaction models, *International Game Theory Review*, Vol. 5, No. 3, pp. 193-210, 2003.
11) Walker, J. L., Ehlers, E., Banerjee, I. and Dugundji, E. R. : Correcting for endogeneity in behavioral choice models with social influence variables, *Transportation Research Part A*, Vol. 45, No. 4, pp. 362–374, 2011.
12) Goetzke, F. and Andrade, P. M. : Walkability as a summary measure in a spatially autoregressive mode choice model: An instrumental variable approach, in *Progress in Spatial Analysis*, A. Páez, J. Gallo, R. N. Buliung and S. Dall'erba, Eds., Springer, pp. 217-229, 2010.
13) Berry, S., Levinsohn, J. and Pakes, A. : Automobile prices in market equilibrium, *Econometrica*, Vol. 63, No. 4, pp. 841-890, 1995.
14) Berry, S., Levinsohn, J. and Pakes, A. : Differentiated products demand systems from a combination of micro and macro data: The new car market, *Journal of Political Economy*, Vol. 112, No. 1, pp. 68-105, 2004.

(Received April 22, 2016)