Estimating Effects of Working Conditions on Labor Surplus: Application to Rural Myanmar[*]

Su Thet Hnin[†], Airi Kato[‡], Shinji Kaneko[§], Keisuke Kawata[¶], Yuichiro Yoshida[‖]

February 19, 2017

Abstract

This paper introduces a new empirical approach to evaluate the welfare impacts of hypothetical labor market policies. We combine and extend a new design of conjoint experiments (Hainmueller, et., al 2014) and an empirical welfare analysis (Bhattacharya 2015), which allows us to nonparametrically estimate the causal effects of the change of labor market characteristics on the distribution of individual labor surplus. As an illustration, we apply the approach in rural area of Myanmar, where the labor demand is rapidly expanding, but there are no regulations for labor markets. Our estimation results show that Myanmar workers are sensitive for the risk of injury and it's compensation. Moreover, by introducing worker's compensation and reducing the risk of injury, the labor surplus distribution is significantly improved as statistically and economically.

Keywords: empirical welfare analysis, full-randomized conjoint experiment, labor surplus, Myanmar

[*]We thank to Shinpei Sano, Teppei Yamamoto, Ryosuke Okazawa, Katuya Takii, and participants in the workshop in GRIPS, the OEIO, Chiba University, and Osaka University.

[†]Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, Japan 739-8529.

[‡]Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, Japan 739-8529

[§]Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, Japan 739-8529.

[¶]Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, Japan 739-8529, E-mail: keisuke@hiroshima-u.ac.jp.

[‖]Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, Japan 739-8529.

1 Introduction

The main purpose of this paper is to empirically examine welfare effects of working conditions, based on the labor surplus. More specifically, we estimate how affected the labor surplus by changing working conditions, such as more safety working environment, providing workplace insurance, and changing working location. To do so, the paper proposes a simple empirical approach which combines and extends an empirical welfare analysis (Bhattacharya 2015) and a new design of conjoint experiments (Hainmueller, et., al 2014). The approach allows us to estimate the causal effects of working conditions on the distribution of labor surplus. Moreover, our approach does not require huge sample size and any restrictive assumptions on the nature of preference heterogeneity and income effects, which can be then implemented in many area where both data availability and background information are limited.

In both developed and developing countries, "working condition" is consisted with many attributes including wage, risk of injury, and working location. Labor economists have then examined the welfare implications of such attribute for instance, how affects the labor surplus by changing each working condition. While the methods of theoretical welfare analysis are well developed, it's empirical methods are still contravention. This paper contributes to offer a new empirical approach of "comparative statistics" on labor surplus.

In the conventional approach of the empirical welfare analysis, estimation of the welfare indicator is mainly based on a random utility model (see Hensher et al 2005 and Train 2009). In recent years, a recent growing field in the literature is nonparametric welfare-analysis which has been made under general preference heterogeneity. In the discrete choice settings, Bhattacharya (2015) proposes simple identification results to recover the marginal distribution of equivalent/compensating variation resulting from a price-change. Moreover, Bhattacharya (2016) extends Bhattacharya (2015) and shows that we can identify equivalent/compensating variation for important additional scenarios; simultaneous price-change of multiple alternative, elimination of a choice-alternative, and choice among non-exclusive options. Our approach is also extension of Bhattacharya (2015) to explicitly allow multiple attributes of choice-alternative. Additionally, this paper newly introduces a causal quantity, *the marginal component effect on surplus distribution*, which is based on the counterfactual framework and can evaluate the welfare effect of each job-attribute. More technically, the *marginal component effect on surplus distribution* is the marginal effect of each attribute on the distribution of labor surplus, which can be simply identified in limited sample size.

Similar to Bhattacharya (2015, 2016), our identification results do not depend on any parametric assumptions on preference but require some identification assumptions on data set. Data needs to include (i) exogenous variation of attributes and (ii) attributes of unchoice alternative. Therefore, even through our results can potentially apply for the study with observed data, it may be then not easy to find data satisfying those identification assumptions. However, our approach can be applied the data generated by a conjoint survey experiment (Hainmueller, et., al 2014). They present a new design of conjoint experiment which is based on the standard potential outcome framework (Neyman 1923 and Rubin 1974). In their approach, full randomization of attribute levels ensures nonparametric identification of the average marginal effect of many attributes without any untestable assumption on the individual preference. Hainmueller, et., al. (2014) can then provide valid estimators, but their original approach cannot directly apply to estimate any welfare indicators including "surplus" because only effects on choice probabilities can be estimated[1]. The present paper newly shows that by combining the rational choice model, Hainmueller, et., al. (2014)'s approach is useful even for empirical welfare analysis.

As an illustration of theoretical results, we estimate the *marginal component effect on surplus distribution* in the Inlay lake located in rural Myanmar. Around the lake, the labor demand in construction sector is rapidly increased, but there are no studies for design of labor regulation. We conduct the survey in which respondents rank three options in terms of their personal preference: working in one of two jobs and one that maintains the status quo (not working). Our estimators show that the reducing the risk of injury and providing worker's compensation have significant welfare effects, which are larger than the welfare effects of other attribute as working location.

The structure of paper is as follows. Section 2 introduces the conceptual frameworks and shows identification results. Section 3 shows the illustration of our identification results with conjoint data. In the section, we first present our conjoint design, background of survey area, and estimation strategies. After that, the estimation results are shown. Finally, Section 4 presents conclusions.

2 Conceptual Framework

The section introduces a causal and welfare quantity, *the marginal component effect on surplus distribution*, and shows a practical identification result. We start from the statistical framework as potential outcome to define the causal effects on choice probabilities, and then the rational choice framework is introduced to define and identify *the marginal component effect on surplus distribution.*

Through the section, let consider workers facing the choice problem whether working in a job or not. A job is characterized by wage, $w \in \Phi_w$, and L types of non-pecuniary attributes, $\boldsymbol{a} = \{a_1, ..., a_L\} \in \Phi_{\boldsymbol{a}}$. The researcher can observe realizations of the decision at individual level with job-characteristics.

2.1 Potential outcome framework

Our causal inference is based on Rubin's potential outcome framework (Neyman 1923; Rubin 1974). Let $Y_i(w, \boldsymbol{a})$ denotes a potential choice outcome; $Y_i(w, \boldsymbol{a}) = 1$ if a worker i chooses to work in a job with wage w and non-pecuniary attributes $\boldsymbol{a}$, and $Y_i(w, \boldsymbol{a}) = 0$ if not. Following Hainmueller, et., al. (2014), the *average conditional component effect* (ACCE) of an attribute l is defined by

$$\bar{\pi}_l(a_1, a_0 | w, \boldsymbol{a}_{-l}) = E\left[Y_i(w, a_1, \boldsymbol{a}_{-l}) - Y_i(w, a_0, \boldsymbol{a}_{-l})\right], \tag{1}$$

where $\boldsymbol{a}_{-l}$ is a vector of non-pecuniary attributes excluding attribute l. Because $E\left[Y_i(w, a_1, \boldsymbol{a}_{-l})\right]$ can be interpreted as a potential choice probability to work,$\bar{\pi}_l(a_1, a_0 | w, \boldsymbol{a}_{-l})$ can be then interpreted as the increase in the choice probability if the

[1] In the literature on the conjoint-survey experiment, there are many applications of Hainmueller, et., al. (2014)'s design. For instance, Bechtel and Scheve (2013), Gampfer, et., al (2014), and Bernauer and Gampfer (2015) are for international environmental agreements. However, in our best knowledge, there are no papers on job preference.

value of an attribute l is changed from a_0 to a_1 given values of wage and other attributes.

We next introduce identification assumptions. Let suppose that a data observes an individual job choice; Y_{ij} denotes the (observable) choice indicator: $Y_{ij} = 1$ if a worker i chooses to work in a job j, and $Y_{ij} = 0$ if not. Additionally, job j's wage W_{ij} and non-pecuniary attributes $\boldsymbol{A}_{ij} = \left\{A_{1ij}, ..., A_{Lij}\right\}$ can be also observable. We assume that the working conditions offering a worker are randomized and full supported.

Randomization $\{Y_i(w, \boldsymbol{a})\}_{w,\boldsymbol{a}} \perp\!\!\!\perp \left\{W_{ij}, \boldsymbol{A}_{ij}\right\}$ for all i and j, and $0 < p(w, \boldsymbol{a}) \equiv \Pr(W_{ij} = w, \boldsymbol{A}_{ij} = \boldsymbol{a})$ for all $w \in \Phi_w$ and $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$.

The first part of *Randomization* requires that the potential outcomes are statistically independent on each realized wage and non-pecuniary attribute, which must be satisfied if both wage and non-pecuniary attribute are randomized. The second part of the assumption states that the randomization scheme must assign a positive probability to all the possible combinations of wage and non-pecuniary attributes. Note that while it is hard to satisfy *Randomization* in observed data, as shown latter, a new design of conjoint experiment can ensure the assumption.

Hainmueller, et., al. (2014) shows that *Randomization* allows to simply and simply identify $\bar{\pi}_l(a_1, a_0|w, \boldsymbol{a}_{-l})$ as

$$\hat{\pi}_l(a_1, a_0|w, \boldsymbol{a}_{-l}) = E\left[Y_{ij}|W_{ij} = w, A_{lij} = a_1, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right] - E\left[Y_{ij}|W_{ij} = w, A_{lij} = a_0, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right], \tag{2}$$

where $\boldsymbol{A}_{-lj}$ is a vector of non-pecuniary attributes excluding $l$the attribute. In practice, however, the estimation of the ACCE is difficult because the numbers of observations that belong to the conditioning set in the right-hand side of equations (2) are very small.

The average marginal component effect (AMCE) is introduced as

$$\bar{\pi}_l(a_1, a_0, p(w, \boldsymbol{a}_{-l})) = \sum_{w,\boldsymbol{a}_{-l}} p\left(W_{ij} = w, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right) \times E\left[Y_i(w, a_1, \boldsymbol{a}_{-l}) - Y_i(w, a_0, \boldsymbol{a}_{-l})\right], \tag{3}$$

where $p\left(W_{ij} = w, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right)$ is the joint distribution of wage and other attributes. The AMCE can be then interpreted as the increase in the external probability if a value of an attribute l is changed from a_0 to a_1, averaged over all the possible values of wage and other attributes given their joint distribution $p\left(W_{ij} = w, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right)$.

By using the identification result (2), the AMCE can be simply identified as

$$\hat{\pi}_l(a_1, a_0, p(w, \boldsymbol{a}_{-l})) = \sum_{w,\boldsymbol{a}_{-l}} p\left(W_{ij} = w, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right)$$

$$\times\left\{E\left[Y_{ij}|W_{ij} = w, A_{lij} = a_1, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right] - E\left[Y_{ij}|W_{ij} = w, A_{lij} = a_0, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right]\right\}. \tag{4}$$

As discuss in Hainmueller, et., al. (2014), the AMCE is easily estimated even in limited sample size. Therefore, in the next section, we report only AMCE, nor ACCE.

Finally, we define the AMCE conditional on only wage;

$$\bar{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|w) = \sum_{\boldsymbol{a}} p\left(\boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right) \times E\left[Y_i(w, a_1, \boldsymbol{a}_{-l}) - Y_i(w, a_0, \boldsymbol{a}_{-l})\right], \tag{5}$$

and then identified as

$$\hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|w) = \sum_{\boldsymbol{a}_{-l}} p\left(\boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right)$$
$$\times \left\{E\left[Y_{ij}|W_{ij} = w, A_{lij} = a_1, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right] - E\left[Y_{ij}|W_{ij} = w, A_{lij} = a_0, \boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right]\right\}. \tag{6}$$

The AMCE conditional on wage is a quit important causal quantity because as shown in Section 2.2, all causal welfare quantity can be connected with the AMCE. Note that even through we need to conditional on not only an attribute l but also wage level, the right-hand side in equation (6) is still estimable in our limited sample size.

2.2 Rational choice framework

Next, we introduce a rational choice framework that shows that the identified causal effects on choice probabilities (equation 6) can be connected with the distribution of labor surplus.

2.2.1 Preference

Assuming that individuals have *complete* and *transitive* preferences on working conditions[2]. Let $\succeq_i$ denotes preference of an individual i; $\{w, \boldsymbol{a}\} \succeq_i \phi$ if she/he would like to work in a job with w and $\boldsymbol{a}$ where ϕ denotes status-quo (not working). Note that same as Bhattacharya (2015), our framework does not require any strong assumptions on the utility functions; it is nonparametrically specified and allows preference-heterogeneity and income effects. We just need four nonparametric assumptions on preference;

Monotonicity For any $w, w' \in \Phi_w$ and $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $\{w, \boldsymbol{a}\} \succeq_i \{w', \boldsymbol{a}\}$ if and only if $w \geq w'$.

Continuity For any $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $\succeq_i$ is a continuous over w.

Boundary For any $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $\phi \succeq_i \{0, \boldsymbol{a}\}$.

Rationality: For any $w \in \Phi_w$ and $\boldsymbol{a} \in \Phi_{\boldsymbol{a}}$, $Y_i(w, \boldsymbol{a}) = 1$ if and only if$\{w, \boldsymbol{a}\} \succeq_i \phi$.

Monotonicity requires that all workers prefer jobs with higher wages given non-pecuniary attributes. Combining with transitivity yields $\{w, \boldsymbol{a}\} \succeq_i \phi$ if $\{w', \boldsymbol{a}\} \succeq_i \phi$ and $w \geq w'$. This is a key assumption to obtain a main identification result.

Both *Continuity* and *Boundary* are technical assumptions. *Continuity* is needed only to ensure the existence of individual labor surplus (as shown latter). *Boundary* requires that no workers prefer to work in a job without wage, which is needed only to provide upper bounds of individual labor surplus.

[2] *Transitivity* requires that $\{w', \boldsymbol{a}'\} \succeq_i \{w'', \boldsymbol{a}''\}$ if $\{w, \boldsymbol{a}\} \succeq_i \{w', \boldsymbol{a}'\}$ and $\{w', \boldsymbol{a}'\} \succeq_i \{w'', \boldsymbol{a}''\}$.

Finally, *Rationality* requires that an individual chooses to work if and only if she/he prefers to work, which leads the following equation;

$$E\left[Y_i(w, \boldsymbol{a})\right] = \Pr\left[\{w, \boldsymbol{a}\} \succeq_i \phi\right]. \tag{7}$$

Above equation means that *Rationality* can connect the choice probability and worker's preference. Moreover, the choice probability can be interpreted as the preference probability.

2.2.2 Labor surplus

First, we show the identification result of the distribution of labor surplus from a labor market with wage ω and non-pecuniary attributes $\boldsymbol{a}$. An individual surplus from the labor market, $LS_i(\omega, \boldsymbol{a})$, is defined as

$$LS_i(\omega, \boldsymbol{a}) = 0 \text{ if } \phi \succeq_i \{\omega, \boldsymbol{a}\},$$

$$\{\omega - LS_i(\omega, \boldsymbol{a}), \boldsymbol{a}\} \sim_i \phi \text{ if } \{\omega, \boldsymbol{a}\} \succ_i \phi. \tag{8}$$

It is straightforward to show that *Continuity* and *Boundary* ensure the existence and uniqueness of $LS_i(\omega, \boldsymbol{a})$ for any $\omega \in \Phi_w$ and $\boldsymbol{a} \in \Phi_A$.

Because it is quit difficult to estimate the individual labor surplus, we focus to estimate related with the distribution of the individual labor surplus. Let $\Pr\left[LS_i \leq X|\omega, \boldsymbol{a}\right]$ denotes the cumulative distribution function of the individual labor surplus if wage and values of non-pecuniary attributes are w and $\boldsymbol{a}$.

From *Monotonicity* and the definition of labor surplus (8), for any $X \in (0, \omega)$,

$$LS_i \leq X \iff \phi \succeq_i \{\omega - X, \boldsymbol{a}\},$$

which allows us to rewrite $\Pr\left[LS_i \leq X|\omega, \boldsymbol{a}\right]$ as

$$\Pr\left[LS_i \leq X|\omega, \boldsymbol{a}\right] = \Pr\left[\phi \succeq_i \{\omega - X, \boldsymbol{a}\}\right]. \tag{9}$$

Above equation means that the share of individuals receiving lower surplus than X is equal to the share of individuals who do not prefer to work a job.

Combining with equations (7) and (9) yields an identification result for the surplus distribution. If X is satisfied $\omega - X \in \Phi_w$,

$$\Pr\left[LS_i \leq X|\omega, \boldsymbol{a}\right] = 1 - E\left[Y_i(\omega - X, \boldsymbol{a})\right], \tag{10}$$

which means that the distribution of individual labor surplus can be rewritten by sorry the choice probability. Moreover, equation (10) shows that the distribution of labor surplus is identifiable because the right-hand side of equaiton (10) can be identified by $E[Y_{ij}|W_{ij} = \omega - X, \boldsymbol{A}_{ij} = \boldsymbol{a}_{ij}]$. In practice, however, the estimation of $E\left[Y_i(\omega - X, \boldsymbol{a})\right]$ is difficult because of small sub-sample size problem as mentioned above. In the next section, we then introduce more practical causal quantity.

2.3 Marginal component effect on surplus distribution

We next define the conditional effect of an attribute l on the distribution of labor surplus, which is the change of labor surplus distribution if an attribute is changed. The cumulative distribution function of labor surplus with $a_l = a_1$ and $a_l = a_0$ can be defined as $\Pr\left[LS_i \leq X|\omega, a_1, \boldsymbol{a_{-l}}\right]$ and $\Pr\left[LS_i \leq X|\omega, a_0, \boldsymbol{a_{-l}}\right]$, respectively. The conditional component effect on surplus distribution is (CCE_SD) defined as

$$\bar{\Pi}_l(X|\omega, \boldsymbol{a}_{-l}) = \Pr\left[LS_i \leq X|\omega, a_1, \boldsymbol{a_{-l}}\right] - \Pr\left[LS_i \leq X|\omega, a_0, \boldsymbol{a_{-l}}\right]. \tag{11}$$

$\bar{\Pi}_l(X|\omega, \boldsymbol{a}_{-l})$ can be interpreted as the increase in the share of labor surplus less than X if the value of an attribute l is changed from a_0 to a_1 given wage and other attributes in the labor market.

We next show the identification result of the conditional component effect on surplus distribution $\bar{\Pi}_l(X|\omega, \boldsymbol{a}_{-l})$. From equation (10), each distribution of labor surplus with $a_l = a_0$ and a_1 can be rewritten as

$$\Pr\left[LS_i \leq X|\omega, a_1, \boldsymbol{a_{-l}}\right] = 1 - E\left[Y_i(\omega - X, a_1, \boldsymbol{a}_{-l})\right],$$

and

$$\Pr\left[LS_i \leq X|\omega, a_0, \boldsymbol{a_{-l}}\right] = 1 - E\left[Y_i(\omega - X, a_0, \boldsymbol{a}_{-l})\right].$$

Combining with the definition of the ACCE (equation 1) yields

$$\bar{\Pi}_l(LS_i \leq X|\omega, \boldsymbol{a}_{-l}) = E\left[Y_i(\omega - X, a_0, \boldsymbol{a}_{-l})\right] - E\left[Y_i(\omega - X, a_1, \boldsymbol{a}_{-l})\right] = -\bar{\pi}_l(a_1, a_0|\omega - X, \boldsymbol{a}_{-l}). \tag{12}$$

Therefore, from the identification result of ACCE (equaiton 2), $\bar{\Pi}_l(X|\omega, \boldsymbol{a}_{-l})$ can be identified as

$$\hat{\Pi}_l(LS_i \leq X|\omega, \boldsymbol{a}_{-l}) = -\hat{\pi}_l(a_1, a_0|\omega - X, \boldsymbol{a}_{-l}). \tag{13}$$

Above equation means that the conditional component effect on surplus distribution can be identified by the ACCE which is also simply identified by the difference-in-means estimator. Equation (13) then shows a simple but nonparametric identification result to recover the conditional component effect on surplus distribution, however, as mentioned above, the small sub-sample size problem is generally serious.

Therefore, to use in practical situation, we introduce an alternative causal quantity, the marginal component effect on surplus distribution (MCE_SD), which is defined as

$$\bar{\Pi}_l(LS_i \leq X|\omega) = \sum_{\boldsymbol{a}_{-l}} \bar{\Pi}_l(LS_i \leq X|\omega, \boldsymbol{a}_{-l}) \times p(\boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}).$$

The MCE_SD can be interpreted as the increase in the cumulative distribution function if the value of an attribute l is changed from a_0 to a_1, averaged over all the possible values of other attributes given their joint distribution $p\left(\boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}\right)$.

By using equations (5) and (12), $\bar{\Pi}_l(LS_i \leq X|\omega)$ can be rewritten as

$$\bar{\Pi}_l(LS_i \leq X|\omega) = -\sum_{\boldsymbol{a}_{-l}} \bar{\pi}_l(a_1, a_0|\omega - X, \boldsymbol{a}_{-l}) \times p(\boldsymbol{A}_{-lij} = \boldsymbol{a}_{-l}) = -\bar{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|\omega - X).$$

If X satisfies $\omega - X \in \Phi_w$, equation (6) provides the identification results as

$$\hat{\Pi}_l(LS_i \leq X|\omega) = -\hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|\omega - X). \tag{14}$$

Note that the minor modification of equation (14) yields

$$\hat{\Pi}_l(X' \leq LS_i \leq X|\omega) = \hat{\Pi}_l(LS_i \leq X|\omega) - \hat{\Pi}_l(LS_i \leq X'|\omega)$$

$$= -\hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|\omega - X) + \hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|\omega - X'). \tag{15}$$

for any X and X' satisfying $\omega - X \in \Phi_w$ and $\omega - X' \in \Phi_w$.

Equations (14) and (15) mean that the MCE_SD can be identified only by the AMCE. Therefore, the MCE_SD has some practical advantages. First, it allows us to evaluate the welfare impact of each non-pecuniary attributes, which has rich implications for labor market policies. Second is that these effects can be easily estimated by the difference-in-means estimators conditional on just wage and the level of an interest attribute. In the next section, we then report estimated MCE_SD by using equation (15).

3 Illustration

3.1 Survey design

Each respondent is represented with 3 choice tasks in our conjoint survey experiment. In each choice task, a respondent puts the ranking of preferred alternatives in a choice set including three alternatives, job A, job B, and the status quo (or not working one of the first two jobs).

Both jobs A and B are characterized by five attributes; (1) working location, (2) nationality of ownership, (3) risk of injury, (4) worker's compensation, and (5) monthly wage payment. The levels of each attribute are shown in Table 1.

[Table 1 around here]

Working location represents the location of working place, with two levels; Taunggyi and Nyaung Shwe. Nyaung Shwe is the nearest town of the Inly lake, while Taunggyi has a distance of 27 km from Nyaung Shwe. The risks of injury are characterized by two dimensions. The first dimension is a probability of injury; high probability (457 in 100,000 workers) or low probability (229 in 100,000 workers), while the second dimension is recoverable or not. Common assumptions of the proposed job opportunities are 1) 6 months' continuous work period with monthly payment, 2) 8 hours per day from 8 am to 5 pm for 6 days per week where Sunday is holiday, and 3) simple lodging is provided if you cannot commute every day at free of charge.

Finally, the survey is conducted from December 27, 2016, to January 25, 2017 by 327 households. After conjoint survey experiments, a survey for respondent's background characteristics is also conducted, which show the household characteristics are as follows. Slightly more than half of all respondents are female (51.38%) and their average age is 45 years. Their family size ranges from 1 to 11, with an average size of 4.7.

[Figure 1 around here]

The total annual household income ranges from 360,000 Kyats (300 USD) to 21.6 million Kyats (18,000 USD), with an average of 3.6 million Kyats (3,000 USD). Moreover, the average per capita income is 840,000 Kyats (700 USD), with a range of between 158,000 Kyats (131 USD) and 5.4 million Kyats (4,500 USD). Figure 1 shows the income distribution by kernel density. Similar to developed countries, the Myanmar income distribution has long tail.

Finally, nearly 60% of the household heads completed up to a primary level of education (59.33%), 21.10% completed up to middle school, and 3.98% and 1.83% are high school and collage/university graduates, respectively, whereas 13.76% are illiterate. In all, 43.43% of household heads are engaged in farming, and typically tomato production.

3.2 Estimation strategy

The identification results in Section 2 can be applied for our conjoint data. To identify the AMCE and MCE_SD, let specified Φ_W and Φ_A as in Table 1 and $p(w, \boldsymbol{a}_{-l})$ as the uniform distribution.

3.2.1 Estimation of AMCE

Based on respondent's ranking of preferred alternatives, we define a choice indicator of respondent i as a job j in her/his kth choice task, $\bar{Y}_{ijk}$, $\bar{Y}_{ijk} = 1$ is assigned to any policy alternatives with a higher ranking than that of the status quo, and is $\bar{Y}_{ijk} = 0$ otherwise.

Hainmueller, et., al (2014) shows that the AMCE can be simply estimated if following two assumptions hold.

Stability For any i, k, and k',

$$\bar{Y}_{ijk} = \bar{Y}_{ijk'} \text{ if } T_{ijk} = T_{ijk'}.$$

Perfect-Randomization $T_{ijkl} \perp\!\!\!\perp \left\{W_{ijk}, \boldsymbol{T}_{ijk-l}\right\}$ and $W_{ijk} \perp\!\!\!\perp \boldsymbol{T}_{ijk}$ for all i, j, k, and l.

Stability requires that the choice indicator always take on the same value as long as all job characteristics take same values. Under *Stability,* $\bar{Y}_{ijk}$ is equal to Y_{ij}. *Perfect-Randomization* ensures that the level of an attribute is independent from other attribute's level. Our conjoint data clearly satisfy *Perfect-Randomization* because each attribute level is pure-randomly determined.

Under the *Perfect-Randomization*, equation (4) can be rewritten as

$$\hat{\pi}_l(a_1, a_0, p(w, \boldsymbol{a}_{-l})) = E[Y_{ij}|A_{ijl} = a_1] - E[Y_{ij}|A_{ijl} = a_0].$$

From equation (6), the AMCE conditional on wage is also simply identified as

$$\hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|w) = E[Y_{ij}|W_{ij} = w, A_{ijl} = a_1] - E[Y_{ij}|W_{ij} = w, A_{ijl} = a_0].$$

Above equations imply that both $\hat{\pi}_l(a_1, a_0, p(w, \boldsymbol{a}_{-l}))$ and $\hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|w)$ can be obtained by using simple difference-in-mean estimators.

Moreover, to conveniently estimate difference-in-mean , a la Hainmueller, et., al. (2014), we estimate the following population model;

$$Y_{ij} = \beta_0 + \sum_{w=2}^{D_w} \beta_w \times W_{iijld} + \sum_{l=2}^{7} \sum_{d=2}^{D_l} \beta_{ld} \times A_{ijld} + u_{ij}, \tag{16}$$

where β_0 is a constant term, D_w and D_l are the number of levels of a wage and an attributes l, respectively, and u_{itj} denote the error terms. The AMCE of wage and attribute l are captured by β_w and β_{ld}, which is coefficient of the level d of wage and attribute l, respectively. Similarly, the AMCE conditional on wage can also be estimated by

$$Y_{ij} = \beta_0 + \sum_{w=2}^{D_w} \beta_w \times W_{iijlw} + \sum_{l=2}^{7} \sum_{d=2}^{D_l} \beta_{ld} \times A_{ijld} + \sum_{l=1}^{7} \sum_{w=2}^{D_w} \sum_{d=2}^{D_l} \beta_{wld} \times W_{ijlw} \times A_{ijld} + u_{ij}, \tag{17}$$

where β_{wld} is coefficient of interaction terms between wage and attribute l.

Note that the unit of analysis in the regression is each job in each task of each respondent. Therefore, even though respondents are sampled randomly from the population, the observed choice outcomes within a respondent may be correlated, which may mislead the statistical inference results. For example, respondents have unobservable characteristics that affect their answer in every task, which generates a correlation of choice outcome within a respondent. To avoid the bias from such correlation in the error terms, we use the cluster robust standard error at the respondent level in all regressions, as suggested by Hainmueller, et., al (2014).

3.2.2 Estimation of MCE_SD

By using the estimated AMCE conditional on wage, the MCE_SD can be also estimated by using difference-in-mean estimators. Through this section, ω is specified as $1,500,000$ Kyats. Because $\Phi_W = \{500,000,\ 1,000,000\ 1,500,000\}$, we can estimate four partitions of WTP distribution as

$$\hat{\Pi}_l(1,000,000 < LS_i \leq 1,500,000|1,500,000) = \hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|500,000),$$

$$\hat{\Pi}_l(500,000 < LS_i \leq 1,000,000|1,500,000) = -\hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|500,000) + \hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|1,000,000),$$

$$\hat{\Pi}_l(0 < LS_i \leq 500,000|1,500,000) = -\hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|1,000,000) + \hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|1,500,000),$$

and

$$\hat{\Pi}_l(LS_i = 0|1,500,000) = 1 - \hat{\pi}_l(a_1, a_0, p(\boldsymbol{a}_{-l})|1,500,000).$$

3.3 Estimation results

3.3.1 Estimated AMCE

[Figure 2 around here]

Figures 1 report the AMCE and the 95% confidence intervals in population model (@@). Each solid circle in the figure represents a point estimator, while the horizontal bar is the 95% confidence interval.

The figure shows that workers seriously consider the risk of injury. Especially, removing unrecoverable injury (loss of arm) has largest positive impacts on choice probability. Even to compare an impact of raising wage by triple, the impact of removing unrecoverable injury is still larger.

Related the risk of injury, worker's compensation also has larger impacts than wages. By providing the compensation, the choice probability is increased as 20%. These estimators consistently show that Myanmar workers seriously worry about the risk of injury and, consequently, getting poor.

Compare to risk of injury and it's compensation, nationality of ownership and working location have smaller AMCE. The only exception is that Myanmar people prefer to work in domestic firms.

3.3.2 Estimated average marginal component effects on surplus

[Figure 3, 4, 5, and 6 around here]

In each location, compensation, and risk of injuries, the estimation results of average marginal component effects on surplus are reported in Figure 3, 4, 5, and 6. Each figure reports the estimated AMCE on the WTP distribution between (i) 150,000 Kyats to 100,000 Kyats, (ii) 100,000 Kyats to 50,000 Kyats, (iii) 50,000 Kyats to 0 Kyats, and (iv) 0 Kyats. Same as in Figure 1, each solid circle represents a point estimator, while the horizontal bar is the 95% confidence interval.

First, Figure 3 reports AMCE_S of working location is so weak because because all estimators are statistical insignificant. Meanwhile, Figure 4 clearly shows that worker's compensation has positive impacts on the WTP distribution. By providing worker's compensation, AMCE_S on highest partition (between 150,000 Kyats to 100,000 Kyats) is about 0.25, which implies that the share of WTP between 150,000 Kyats to 100,000 Kyats is increased as 25 percent, while AMCE_S on lowest partition (0 Kyats) is about -0.13, which implies that the share of workers without any positive WTP is decreased as 13 percent. Those findings are consistent with the estimated AMCE (see Figure 1). Figure 1 reported that the AMCE of working location is statistically significant but economically weaker than worker's compensation.

4 Conclusion

This study proposed a new approach of nonparametric welfare analysis with the discrete choice data. The approach can estimate the causal effect of each job characteristics on the distribution of labor surplus without any restrictive assumptions on preference distribution. Moreover, by combining a new conjoint survey experiment designed by Hainmueller, et., al. (2014), our approach is easily implemented.

We apply the approach to examine welfare implications of working conditions in Inlay Lake, Myanmar, based on a randomized conjoint experiment and a nonparametric welfare analysis. We have shown that reducing the risk of injury and it's compensation have strong positive effects on the labor surplus distribution.

Our approach can use to obtain implications for future policy plans to improve the social efficient. However, our study has an important limitation in terms of the external validity of the randomized conjoint analysis. Even though Hainmueller, Hangartner, and Yamamoto (2015) provides evidence for this validity in developed countries, no studies focus on developing countries. Therefore, additional studies are needed to test the external validity of the conjoint experiments in developing countries.

References

[1] Bechtel, M. M., & Scheve, K. F. (2013). Mass support for global climate agreements depends on institutional design. Proceedings of the National Academy of Sciences, 110(34), 13763-13768.

[2] Bernauer, T., & Gampfer, R. (2015). How robust is public support for unilateral climate policy?. Environmental Science & Policy, 54, 316-330.

[3] Bhattacharya, D. (2015). Nonparametric welfare analysis for discrete choice. Econometrica, 83(2), 617-649.

[4] Bhattacharya, D. (2016). Welfare Analysis for Discrete Choice: Some New Results. Manuscript, Depart.

[5] Gampfer, R., Bernauer, T., & Kachi, A. (2014). Obtaining public support for North-South climate funding: Evidence from conjoint experiments in donor countries. Global Environmental Change, 29, 118-126.

[6] Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. Political Analysis, 22(1), 1-30.

[7] Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. Proceedings of the National Academy of Sciences, 112(8), 2395-2400.

[8] Hausman, J. A., & Newey, W. K. (2016). Individual heterogeneity and average welfare. Econometrica, 84(3), 1225-1248.

[9] Hninn, S. T., Kawata, K., Kaneko, S., & Yoshida, Y. (2016). A nonparametric welfare analysis on water quality improvement of the floating people on Inlay Lake via a randomized conjoint field experiment (No. 6-2). Hiroshima University, Graduate School for International Development and Cooperation (IDEC).

[10] Neyman, J (1923). On the application of probability theory to agricultural experiments: Essay on principles, section 9. (translated in 1990). Statistical Science 5, 465–80.

[11] Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of educational Psychology, 66(5), 688.

[12] Train, K. E. (2009). Discrete choice methods with simulation. Cambridge university press.

[13] Wiswall, M., & Zafar, B. (2016). Preference for the Workplace, Human Capital, and Gender (No. w22173). National Bureau of Economic Research.

Appendix: Full scenario

Now we will ask your preferences on hypothetical temporal job opportunities for simple construction works with different conditions. You will be requested to make rankings among three choice-sets including two job opportunities, option A and option B and "not to choose" (option C). Among the three, the rankings are 1) the most preferable choice, 2) the 2nd preferable choice and 3) the least preferable choice. Common assumptions of the proposed job opportunities are 1) 6 months' continuous work period with monthly payment, 2) 8 hours per day from 8 am to 5 pm for 6 days per week where Sunday is holiday, and 3) simple lodging is provided if you cannot commute every day at free of charge.

The job opportunities are characterized by five attributes such as location, company nationality, risk of injury, availability of insurance and payment. The locations are two places, one is in Taunggyi, and the other is in Nyaung Shwe. Company nationality is five types such as Myanmar, Korea, Thailand, China and UK. The risks are four types, high probability of recoverable injury after three months (bone fracture of dominant arm), low probability of recoverable injury after three months (bone fracture of dominant arm), high probability of unrecoverable injury (loss of dominant arm), and low probability of unrecoverable injury (loss of dominant arm). The higher risk is constructed as 10 times higher probability of incident occurrence in the United States in construction sector in 2014. Lower risk is ... (please explain according to the graph). When the insurance is available, the insurance compensates the payment for the period when you cannot work due to the injury by the day within 6 months. When the insurance is not available, the payment is made by the day for only period you work. The payment amount is shown in monthly basis that continues for 6 months. You will be asked to make rankings at three times repeatedly with different combinations of job opportunities with different conditions.

Attribute	Level 1	Level 2	Level 3	Level 4	Level 5
Working location	Taunggyi	Nyaung Shwe			
Nationality of ownership	Myanmar	Korea	Thailand	China	UK
Risk of injury	Fracture of arm with high probability	Fracture of arm with low probability	Loss of arm with high probability	Loss of arm with low probability	
Worker’s compensation	Covered	Uncovered			
Monthly wage payment	500,000 Kyats	1,000,000 Kyats	1,500,000 Kyats		

Table 1: Attribute list

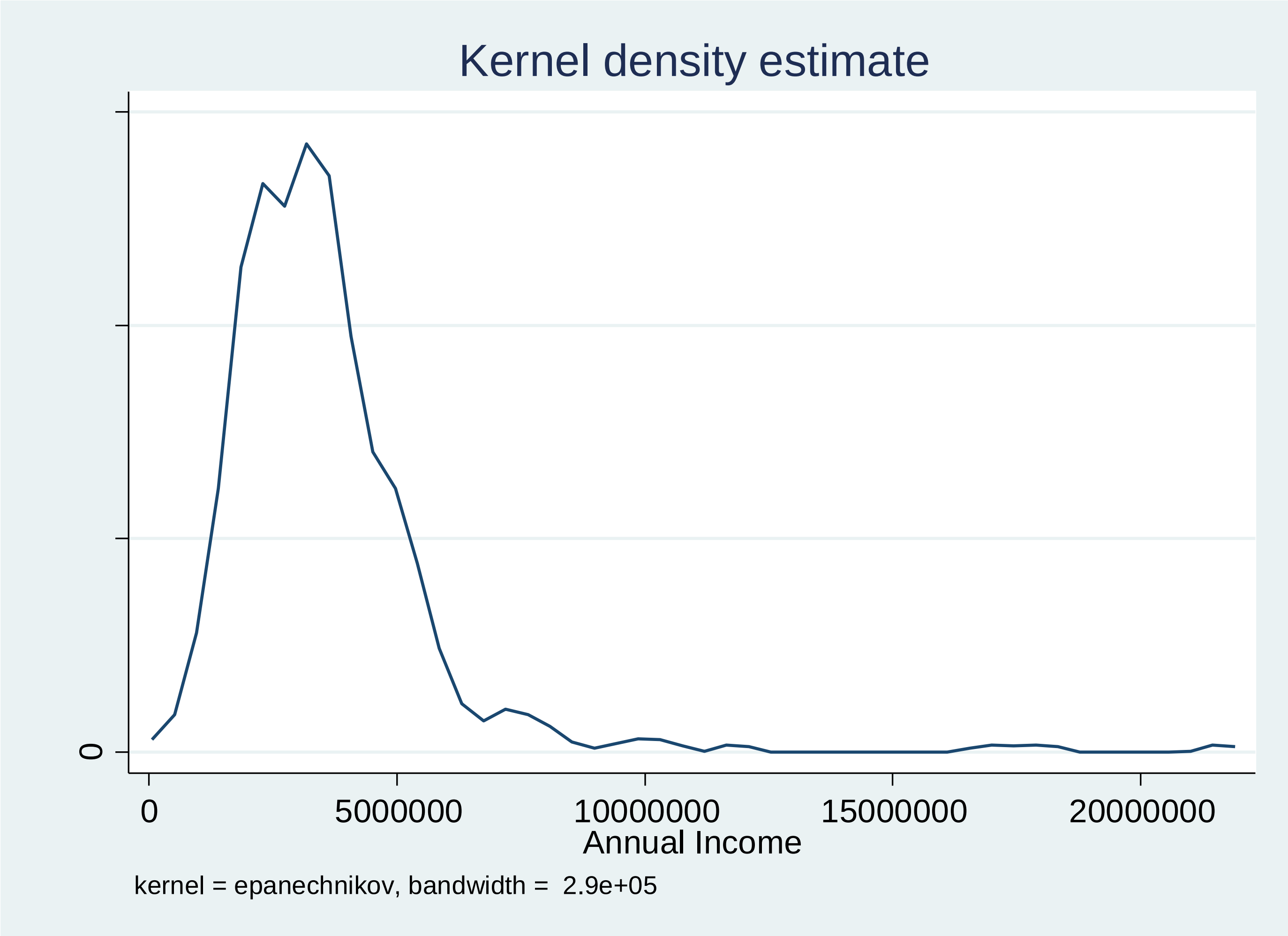

Figure 1. Income distribution

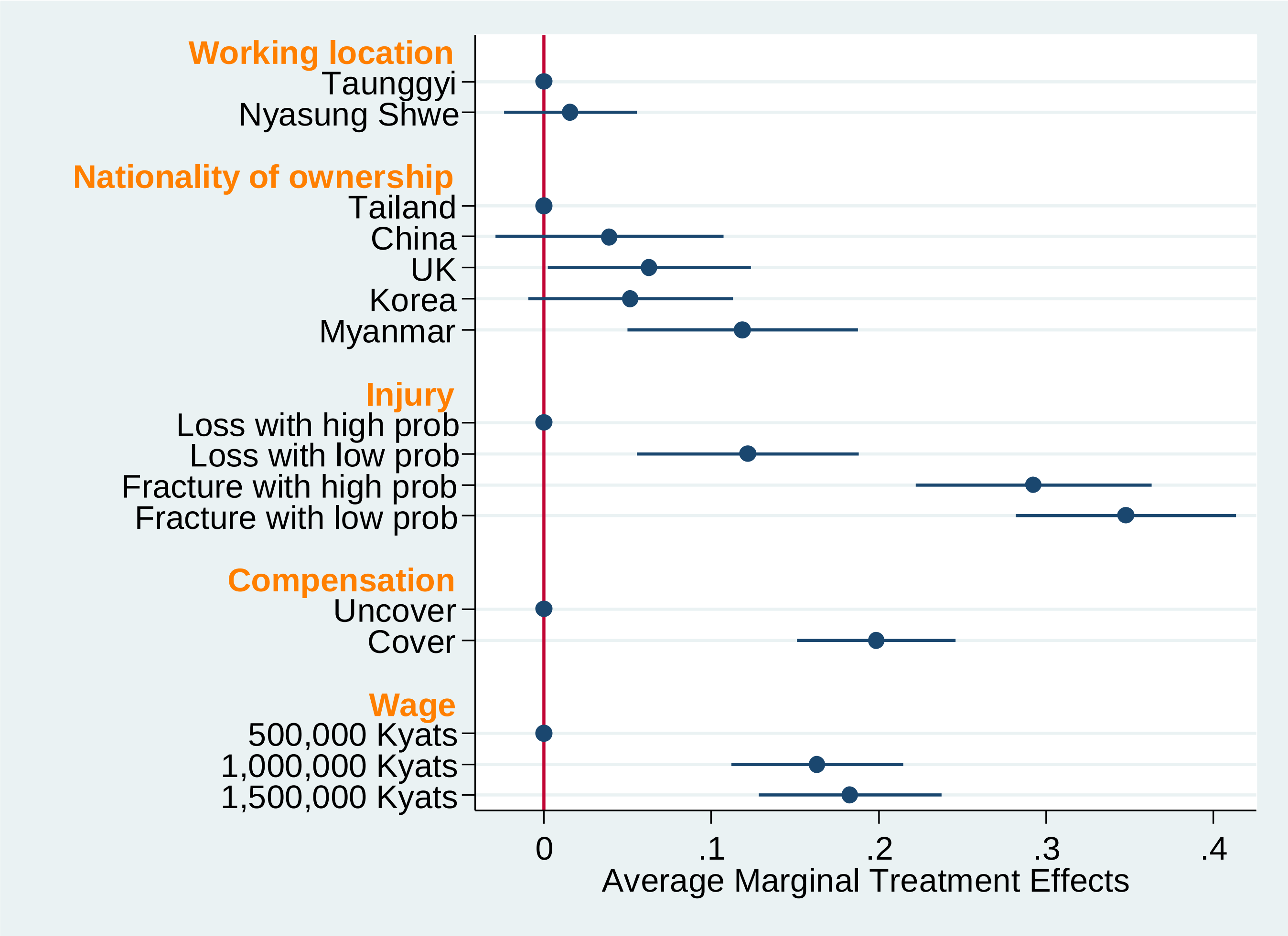

Figure 2. Estimated average marginal treatment effects

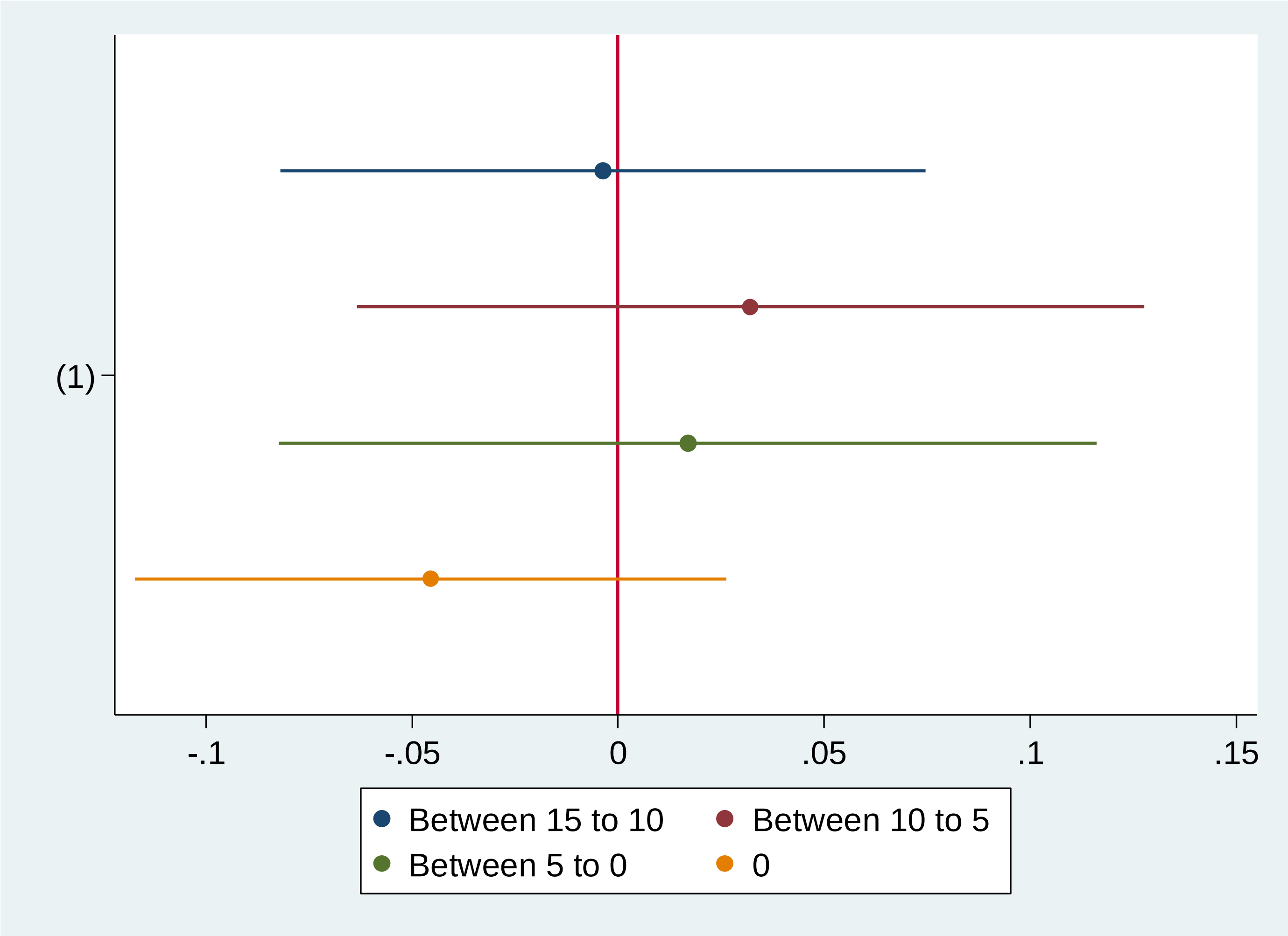

Figure 3. Estimated average marginal treatment effects on surplus distribution: Location

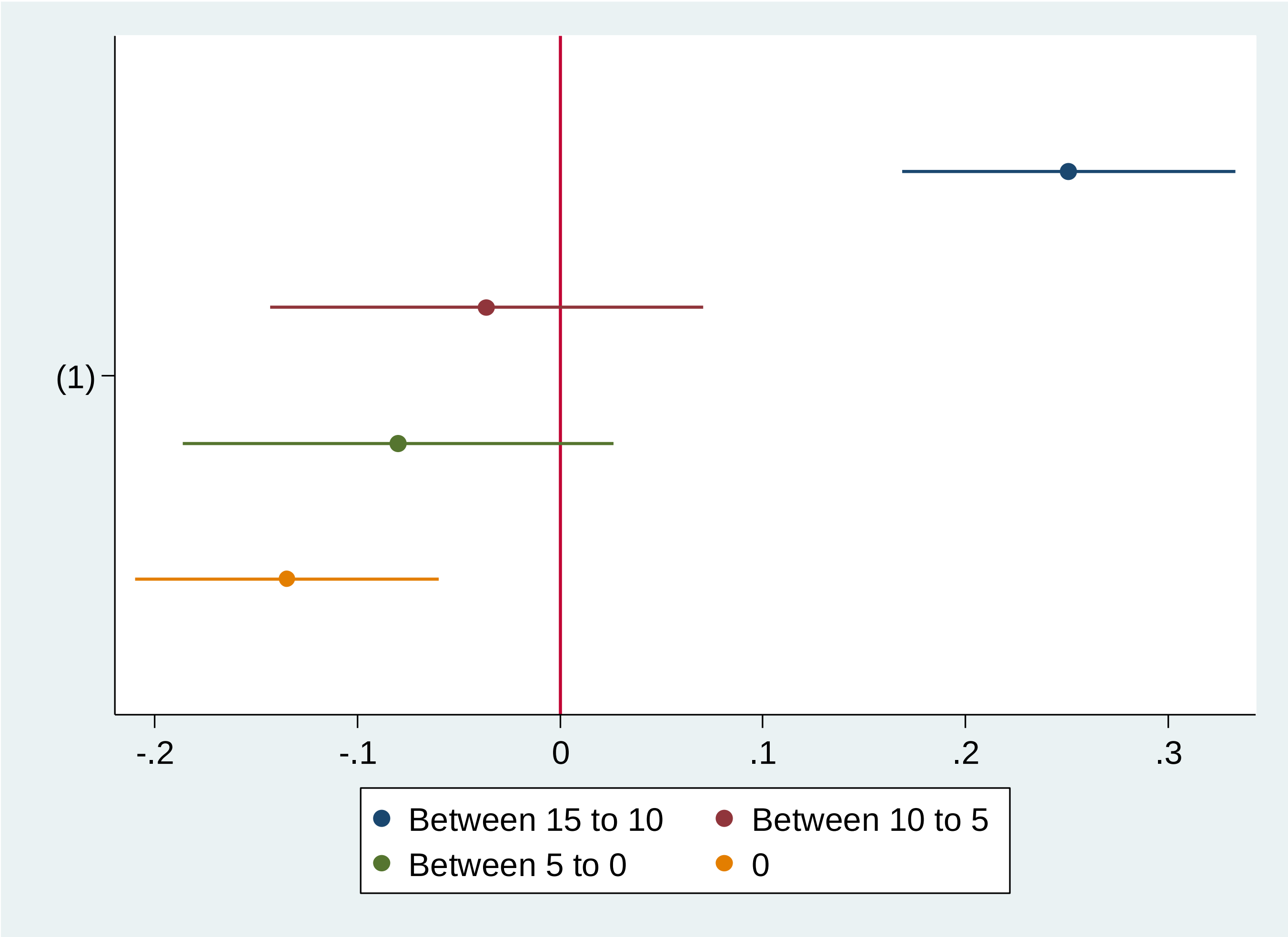

Figure 4. Estimated average marginal treatment effects on surplus distribution: Compensation

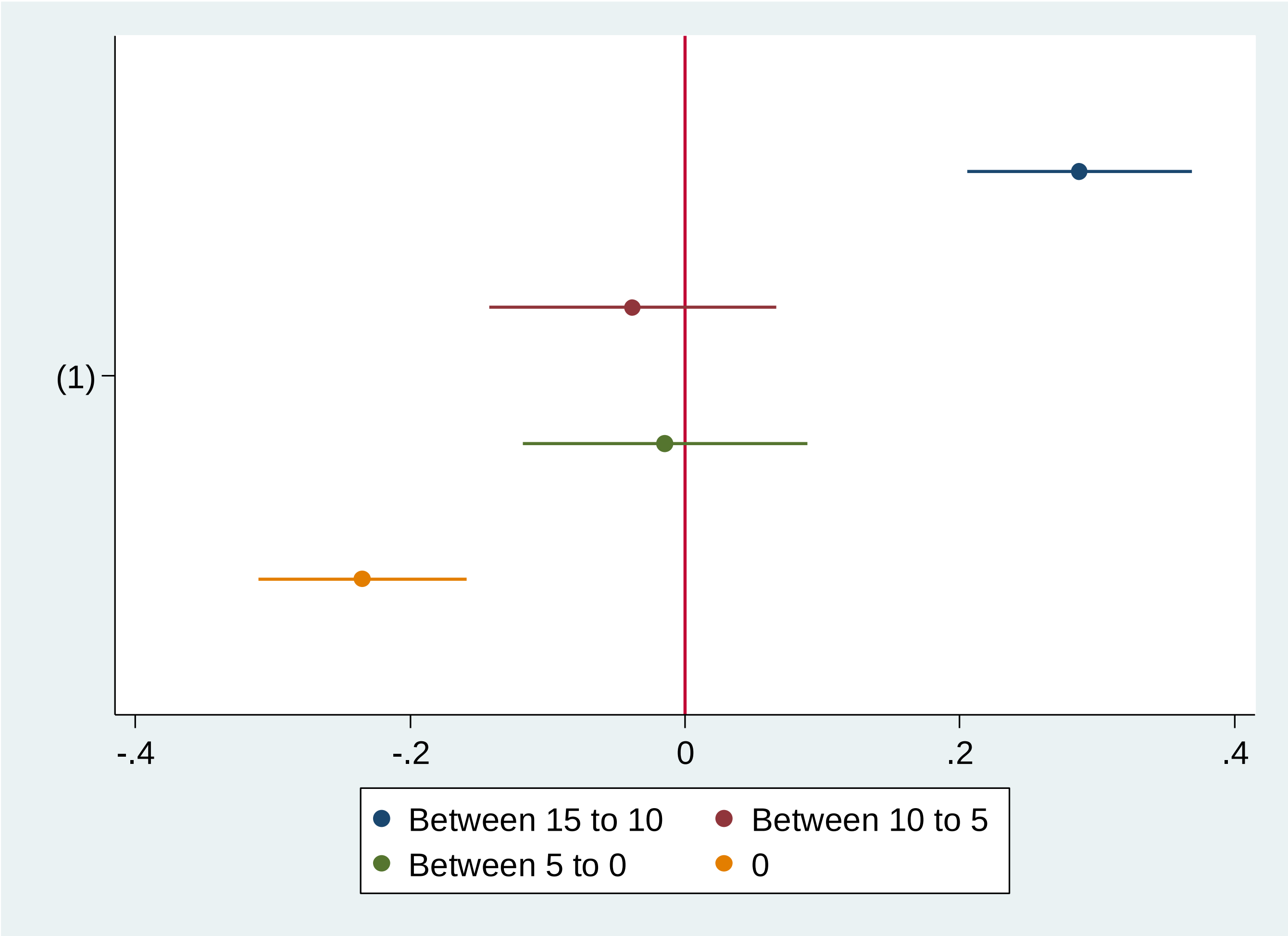

Figure 5. Estimated average marginal treatment effects on surplus distribution: Removing the risk of unrecoverable injury

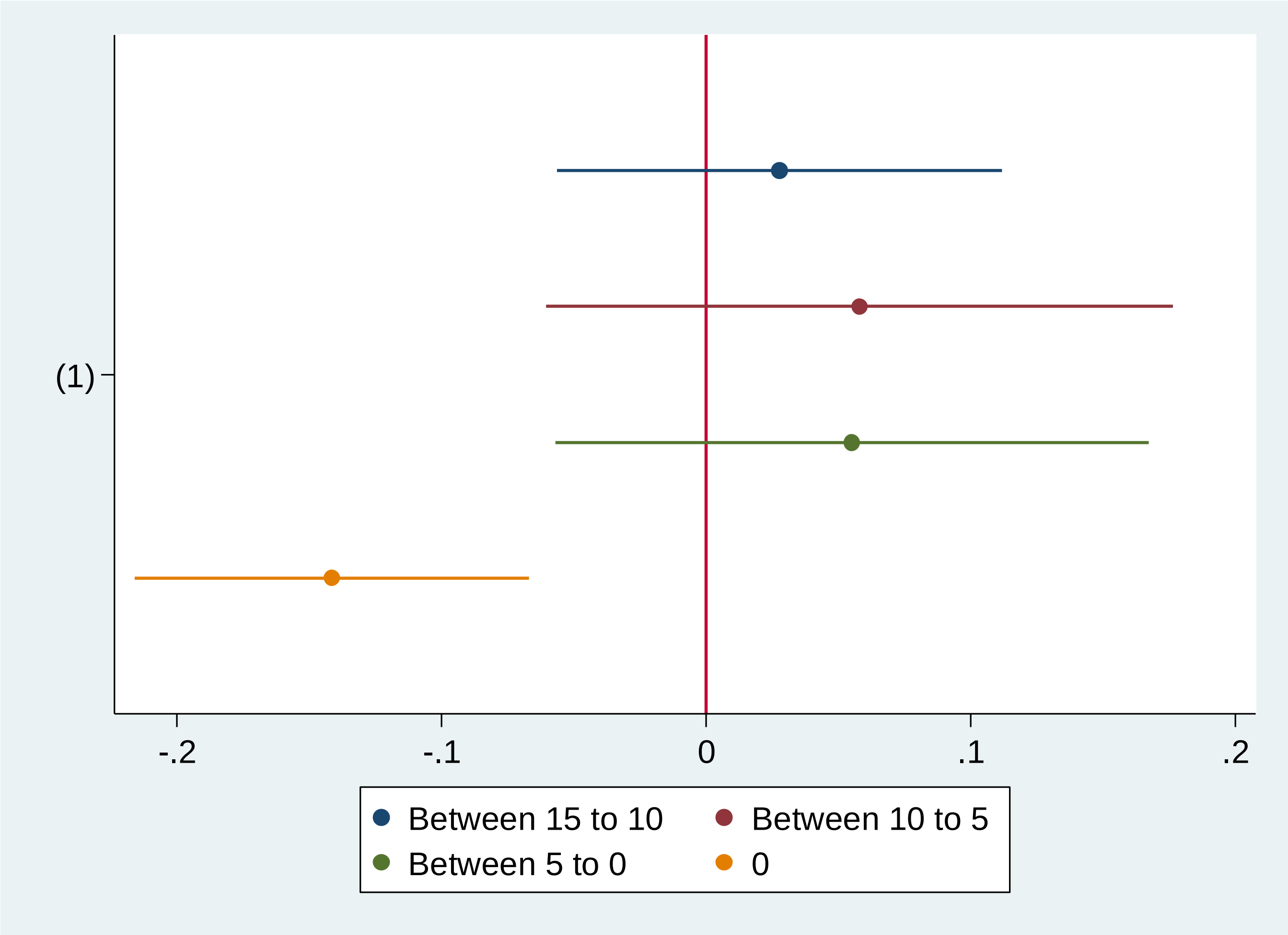

Figure 6. Estimated average marginal treatment effects on surplus distribution: Reducing the probability of injury