Introduction

This study firstly focused on examining the sensitivity of a multivalued treatment effect. Multivalued treatment is often seen in practice, but it is rarely discussed for its causation association possibility in previous research. Second, a small simulation study of ITCV was conducted to examine to what degree the ITCV index is affected by test statistics and sample size. Third, whether the different propensity score weighting methods were sensitive enough to examine different levels of treatment, and how robust the inference can be made on each value treatment, were demonstrated using TALIS data.

To put the relationships in the causal inference framework, the three research questions of interest were:

1) Does teacher participation affect their job satisfaction?  
2) What must be the conditions in the alternative sample to invalidate the inference?  
3) Which propensity score weighting methods can help to identify the effect of teacher participation the most?

Methods

This study first focused on examining the sensitivity of a multivalued treatment effect. Second, a small simulation study of ITCV was developed to examine to what degree the ITCV index is affected corresponding to the test statistics and sample size. Third, whether three propensity score weighting methods were sensitive enough to examine different levels of treatments, and how robust the inference could make on each value treatment was demonstrated using TALIS data.

66,434 teachers in 367 schools of the 21 countries that participated in the Teaching and Learning International Survey (TALIS) 2008 dataset from the Organization for Economic Co-Operation and Development (OECD) were used. We involved all countries that were available for teacher level and country level.

The treatment variable is a 4-level Likert scale at never, seldom, quite often, and very often. These multiple levels were categorized into different dichotomous variables using different thresholds:

1) never, vs. seldom and above (TP1: 1 vs. 2, 3, 4);  
2) never and seldom vs. quite often and very often (TP2: 1, 2 vs. 3, 4);  
3) quite often and below vs. very often (TP3: 1, 2, 3 vs. 4).

Results

<table>
<thead>
<tr>
<th>Table 1</th>
<th>ITCV Indices of Teacher Participation and Job Satisfaction in the Univariate Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>n</td>
</tr>
<tr>
<td>TP1</td>
<td>66,434</td>
</tr>
<tr>
<td>TP2</td>
<td>66,434</td>
</tr>
<tr>
<td>TP3</td>
<td>66,434</td>
</tr>
</tbody>
</table>

Note: r is the threshold of the inference.

Table 3: ITCV Indices of Teacher Participation and Job Satisfaction in the Multivariate Condition

Discussion

The findings of this study validate the inference that teacher participation can affect their job satisfaction. The results also use an index for the threshold for this inference statement, and for examining the robustness of the sample. The indexed threshold necessary for the impact and necessary to invalidate the inference, and the threshold for sample replacement provide power arguments for internal validity and external validity, respectively, and they also show a clear warrant for sensitivity analysis. Because this is an international survey, the large sample size also makes this argument stronger.

Different cut points in the treatment variable were applied to create three dichotomous treatment variables, and the results of these were compared in this study. The results were within expectation, since the lowest cut point made the size of treatment group larger, which allowed for a stronger treatment effect. Similar results, in contrast, can be found when identifying the third treatment variable, when the highest cut point was applied. The treatment effect of the last treatment variable was the smallest. At any rate, the three dichotomous variables were significantly different from zero, indicating their significant effect at any level. Multivalued treatments have not been broadly explored in the social science area, while the Likert scale with levels of four, five, or above has greatly been used in related studies. Knowing how to deal with multivalued treatments instead of losing information by dichotomizing a treatment effect at a random cut point might be a valuable and warranted topic and a critical issue to investigate in future studies.