Methodologies for Tax Credit Evaluation: The Iowa New Jobs Training Program

Zhong Jin
Fiscal and Policy Analyst
Iowa Department of Revenue
Iowa Tax Credit Tracking System

- In the 2005 Legislative Session the Iowa Department of Revenue was authorized to establish a program to track tax credit awards and claims and evaluate various tax credit programs.
- Another purpose of the tax credit program is to provide revenue estimators or legislators with estimates of potential and expected general fund revenue impacts.
- In the 2010 Legislative Session, the Tax Expenditures Review Committee was established and a five year schedule was established for evaluating all tax credits (SF 2380).
Iowa Tax Credit Tracking Program
Major Program Features

- Tax credit database (awards and claims)
- Contingent Liabilities Report
- Annual Tax Credit Claims Report
- Tax credit program evaluation studies
  - Earned Income Tax Credit
  - Historic Preservation Tax Credit
  - Biofuel (Retailers and Producers) Tax Credits
  - Research Activities Tax Credit
  - New Jobs Training Tax Credit
Common Elements of Tax Credit Evaluation Studies

- Legislative Review (Federal and States)
- Review of Related Literature
- Summary of Tax Credit Awards and Claims
- Empirical Analysis of Tax Credit Program Impacts
Empirical Analysis of Tax Credit Program Impacts

- Define Research Questions
- Collect Data
- Determine Appropriate Methodologies
  - Propensity score matching
  - Difference-in-differences
  - Survival analysis
- Model Specification
- Statistical Estimation
- Results Discussion
- Recommendations
Example: New Jobs Training Tax Credit Program (260E) Impact on Wage

- Tax Credit Program Description: State provides withholding and income tax credits to companies to create new jobs and train new employees through community colleges.
- Research Question 1: Did the 260E program increase trainees’ wages?
- Data (1996-2006):
  - individual information (wage, age, job, address) comes from individual income tax records and W2
  - business information (NAICS code, address, size) comes from the Iowa Workforce Development Agency
  - 260E program trainees information comes from community colleges
Example: Methodology

- Establish the trainee group (new jobs and the 260E training) and the control group (new jobs but no 260E training)
- Methodology: Difference-in-differences model estimates impacts of “treatment (training)” by comparing the difference of trainees’ wages before training and after training relative to the difference of wages of individuals in the control group before new jobs and after new jobs
  - Training impact = (trainees’ wages after training - trainees’ wages before training) - (control group’s wages after new jobs - control group’s wages before new jobs)
Example: Methodology Cont.

- Not a controlled experiment
- To address the issue of selection bias
  - Selection bias: Non-trainees may be systematically different from trainees for reasons unrelated to the training. Therefore, there could be non-tax credit effects in the estimated training impact (statistical bias).
Example: Control Group Selection

- Two step process to select a control group and control for selection bias
- Step 1: Determine pools of potential candidates (control for observable selection bias)
  - Living in the same area (control for rural and urban difference)
  - Working in the same industry (control for industry difference)
  - Changing jobs at the same year as trainees in the trainee group (control for timing)
Example: Control Group Selection

- **Step 2: Propensity score matching (control for unobservable selection bias)**
  - Propensity score matching is a methodology attempting to provide unbiased estimation of treatment effects.
  - Purpose: Make the control group ‘similar’ to the trainee group before training/new jobs, so that the 260E training is the only factor contributing to the treatment (training) effect in the difference-in-differences model.
Example: Propensity Score Matching

- Estimate the binary choice function (train=1 or 0)=f(X)+\varepsilon, where X is a set of factors affecting the individual’s chance of participating in the 260E program.
- Calculate the predicted prob(train) using estimated coefficients and X (including age, sector, time, number of jobs).
- For every trainee, choose individuals with the closest predicted prob(train) from the pool of the potential control group.
- Use repeated observations to balance the sample sizes.
### Example: Summary Statistics, Wage Impact Estimation

<table>
<thead>
<tr>
<th>Wage</th>
<th>Trainee Group</th>
<th>Control Group</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wage in third year prior</td>
<td>$28,394</td>
<td>$27,292</td>
<td>$1,102</td>
</tr>
<tr>
<td>Wage in second year prior</td>
<td>$28,514</td>
<td>$27,237</td>
<td>$1,277</td>
</tr>
<tr>
<td>Wage in prior year</td>
<td>$30,364</td>
<td>$27,783</td>
<td>$2,581</td>
</tr>
<tr>
<td>Wage in transition year</td>
<td>$34,100</td>
<td>$28,430</td>
<td>$5,669</td>
</tr>
<tr>
<td>Wage in following year</td>
<td>$37,301</td>
<td>$30,860</td>
<td>$6,441</td>
</tr>
<tr>
<td>Wage in second year after</td>
<td>$38,102</td>
<td>$31,474</td>
<td>$6,628</td>
</tr>
<tr>
<td>Wage in third year after</td>
<td>$37,596</td>
<td>$32,808</td>
<td>$4,788</td>
</tr>
</tbody>
</table>

Trainee Group: 44,576 observations, Control Group: 49,924 observations
Sample Period: 1996-2006
Example: Model Specification

- \[ \text{Wage} = \alpha + \beta_1 (\text{Trainee/Control Group}) + \beta_2 (\text{Before/After Dummy Variable}) + \beta_3 (\text{Trainee/Control Group})* (\text{Before/After Dummy Variable}) + \beta_4 (\text{Other Factors}) + \epsilon \]
  - Wage is the annual wage of an individual
  - Impact of trainee/control group: \( \beta_1 \)
  - Impact of before/after training/changing job: \( \beta_2 \)
  - Impact of training effect: \( \beta_3 \)
  - Other factors: Age, age\(^2\), number of jobs held by individuals, trend and year dummy variables, college dummy variables, and industry dummy variables
## Example: Wage Impact Estimation

### Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Impact</td>
<td>2,476***</td>
<td>409</td>
<td>5.19</td>
</tr>
<tr>
<td>Hiring/training Dummy</td>
<td>-589</td>
<td>357</td>
<td>-0.65</td>
</tr>
<tr>
<td>Training/control Dummy</td>
<td>2,230***</td>
<td>327</td>
<td>13.56</td>
</tr>
</tbody>
</table>

***: significant at 1% level. Adj. R square=0.1506. 94,500 Observations
Example: Wage Impact Results

Discussion

- Compared to individuals in the control group, the 260E programs were found to have increased the average wage of trainees by $2,476 per year up to four years after the training.
- Average wages of new jobs in companies participating in the 260E program are higher than average wages of companies in the control group by $2,230.
- Employment impact is not estimated, which means we assumed that new jobs of the training group would be created even without the 260E program.
Example: Distribution of the 260E Contracts by Average Wage Growth (One Year) Rates
Example: Impact on Individual Income Tax Revenue

- Higher wages of trainees contributed to modestly higher state individual income tax payments.
- For every trainee, the increased individual income tax payment = Marginal tax rate* $2,476, which was about 5.2 percent of a trainee’s tax liability.
Example: Impact of the 260E Program on Tenure

- Research question 2: Did the 260E program increase trainees’ tenure?
- Trainee’s tenure: Period between the time that an individual received new job training and the time that the individual left the company
- Methodology: Survival Analysis
Example: Methodology for Tenure Analysis

- Dependent variable T: the observed length of time working for the same employer
- Dependent variable is censored because some individuals had not changed jobs during the sample period (1996-2006), but could change jobs after 2006 (unobservable).
Example: Survival Analysis

- Independent variables: Dummy variables for community colleges, dummy variable identifying the trainee group/control group
- Function form: $T = f(X) + \epsilon$, $X$ is a vector of independent variables
- Probability distribution function of $T$ is $\text{Prob}(T \leq t)$, where $t=11$ in this study
- Same control group as the wage impact estimation
### Example: Summary Statistics, Tenure Estimation

<table>
<thead>
<tr>
<th>Job Tenure (Years)</th>
<th>Trainee Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Employees</td>
<td>Number of Employees</td>
</tr>
<tr>
<td>0</td>
<td>1,512</td>
<td>114</td>
</tr>
<tr>
<td>1</td>
<td>648</td>
<td>3,510</td>
</tr>
<tr>
<td>2</td>
<td>1,056</td>
<td>2,274</td>
</tr>
<tr>
<td>3</td>
<td>1,303</td>
<td>1,570</td>
</tr>
<tr>
<td>4</td>
<td>579</td>
<td>613</td>
</tr>
<tr>
<td>5</td>
<td>427</td>
<td>529</td>
</tr>
<tr>
<td>6</td>
<td>394</td>
<td>349</td>
</tr>
<tr>
<td>7</td>
<td>312</td>
<td>349</td>
</tr>
<tr>
<td>8</td>
<td>412</td>
<td>218</td>
</tr>
<tr>
<td>9</td>
<td>197</td>
<td>260</td>
</tr>
<tr>
<td>10</td>
<td>79</td>
<td>205</td>
</tr>
<tr>
<td>11</td>
<td>60</td>
<td>27</td>
</tr>
</tbody>
</table>

**Average Job Tenure**
- Trainee Group: 3.18 years
- Control Group: 2.88 years

**Total Employees**
- Trainee Group: 6,979
- Control Group: 10,018
Example: Survival Analysis Estimation Results

- On average, trainees’ job tenure is 6.4 months longer than comparable non-participating workers.
- The 260E programs helped employers retain skilled employees and keep workers who received training in Iowa.
Conclusion

- Is the treatment/control group comparison approach appropriate?
- How to select a control group?
- How to interpret the comparison results?
Link to the 260E Evaluation Study Report

- Questions and comments?