Membership Turnover And Collaboration Success In Online Communities
Sam Ranshotham and Gerald C. Kane

Maximilian Otto Baumann
23.11.2016
Outline

1. Research Question & Theoretical Background
2. Research Design
3. Proportional Hazard Models
4. Variables
   4.1 Dependent Variable
   4.2 Explanatory Variable
   4.2 Control Variables
5. Results
6. Discussion
How does Membership Turnover impact the Quality of Wikipedia Articles?

with respect to Content

- Generation
- Retention

Hypothesis:

“Membership Turnover influences the quality in a curvilinear fashion”
Turnover Effects

Negative Impact

- Expenditure of time and resources to train new members
- Loss of unique knowledge and skills
- Impact on others
Turnover Effects

Positive Impact

- Improving the working environment
- New views and information
- Knowledge is already „harvested“
Turnover Effects

Ambiguous Effect

- Tradeoff between new knowledge and acquired expertise

- “Keep the good, get rid of the bad”
“Analyze impact of membership turnover, on the time it takes an article gets promoted/demoted to/from featured article status, and control for possible confounding effects.”
Proportional Hazard Models
Dependent Variable

Featured Article Status

- between 2001 and 2008
- 2065 received featured article status
- 447 lost featured article status
Explanatory Variable

Experience

- for each revision
- number of edits author had on article

Average Experience

- sum over Experience
- divide by total number of edits on article
Explanatory Variable

Average Experience

- 3,720,826 revisions

- 118,474 monthly observations

- 736,054 distinct authors

- Bots excluded
Control Variables

- article length
- section depth
- external references
- internal references
- reading complexity (ARI score)
- multimedia intensity
- number of edits
Table 1. Proportional Hazard Analysis of the Effect of Membership Turnover

<table>
<thead>
<tr>
<th></th>
<th>Promotion</th>
<th></th>
<th>Promoted</th>
<th>Demoted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 4</td>
<td></td>
</tr>
<tr>
<td>± Article length</td>
<td>-0.169</td>
<td>-0.149</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>± Section depth</td>
<td>-0.001</td>
<td>-0.011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>± External references</td>
<td>-1.566**</td>
<td>-1.484**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>± Internal references</td>
<td>0.058</td>
<td>-0.058</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>± Reading complexity</td>
<td>-1.103</td>
<td>-1.224</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>± Number of edits</td>
<td>0.051</td>
<td>0.086</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>± Experience/Edit</td>
<td>-0.327**</td>
<td>-0.221**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Article length</td>
<td>0.108**</td>
<td>0.107**</td>
<td>-0.024</td>
<td>-0.023</td>
<td></td>
</tr>
<tr>
<td>Section depth</td>
<td>0.624**</td>
<td>0.487**</td>
<td>0.036</td>
<td>0.038</td>
<td></td>
</tr>
<tr>
<td>External references</td>
<td>0.307**</td>
<td>0.244**</td>
<td>-0.168**</td>
<td>-0.162**</td>
<td></td>
</tr>
<tr>
<td>Internal references</td>
<td>0.042</td>
<td>0.065**</td>
<td>0.420**</td>
<td>0.408**</td>
<td></td>
</tr>
<tr>
<td>Reading complexity</td>
<td>-0.010</td>
<td>-0.004</td>
<td>-0.073</td>
<td>-0.078</td>
<td></td>
</tr>
<tr>
<td>Media references</td>
<td>0.27*</td>
<td>0.040**</td>
<td>-0.013</td>
<td>-0.012</td>
<td></td>
</tr>
<tr>
<td>Number of edits</td>
<td>0.059</td>
<td>-0.002</td>
<td>0.144**</td>
<td>0.184**</td>
<td></td>
</tr>
<tr>
<td>Average experience</td>
<td></td>
<td></td>
<td>0.974**</td>
<td>-0.869**</td>
<td></td>
</tr>
<tr>
<td>(Experience/Edit)^2</td>
<td>-1.624**</td>
<td></td>
<td>0.323**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>64,554</td>
<td>64,554</td>
<td>47,364</td>
<td>47,364</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-12541</td>
<td>-12197</td>
<td>-2708</td>
<td>-2694</td>
<td></td>
</tr>
<tr>
<td>( \chi^2 )</td>
<td>1529.33</td>
<td>1623.65</td>
<td>243.83</td>
<td>250.56</td>
<td></td>
</tr>
<tr>
<td>Adjusted pseudo R^2</td>
<td>8.8</td>
<td>11.3</td>
<td>7.3</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>Effect size (f2) of experience</td>
<td>0.03</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo F values</td>
<td>1821.29*</td>
<td></td>
<td>239.21*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Critique

External Validity

- only articles that will be promoted
- no generalization of findings
- in Causal Inference Framework:
  
  Average Treatment Effect on the Treated
Critique

Nomination Procedure

- articles are nominated
- discussion could be source of turnover
- if an article is promising, editors could fix minor issues