Toward Automated Policy Analysis Using Judicial Corpora

Elliott Ash, ETH Zurich
Daniel L. Chen, Toulouse School of Economics

HKU, 29 June 2018
“Promises and Challenges of Data-Driven Legal Research”
Analyzing Judicial Policies

- A large literature hand-codes cases by their policy choices and looks at downstream impacts.
  - (Currie and MacLeod 2008; Belloni et al 2012; Ash and Chen 2017)
  - Necessarily focuses on relatively small sets of decisions
  - Arguably relies on subjective judgments about what policy is being decided and the direction of the decision.

- A parallel literature in law and computation has used unsupervised learning methods to classify large corpora of cases by topic based on language and other features.
  - (e.g. Livermore, Riddell, and Rockmore, 2016, Ash and Chen 2018)
  - The resulting topics may or may not coincide with policy classes
Analyzing Judicial Policies

- A large literature hand-codes cases by their policy choices and looks at downstream impacts.
  - (Currie and MacLeod 2008; Belloni et al 2012; Ash and Chen 2017)
  - Necessarily focuses on relatively small sets of decisions
  - Arguably relies on subjective judgments about what policy is being decided and the direction of the decision.

- A parallel literature in law and computation has used unsupervised learning methods to classify large corpora of cases by topic based on language and other features.
  - (e.g. Livermore, Riddell, and Rockmore, 2016, Ash and Chen 2018)
  - The resulting topics may or may not coincide with policy classes
Analyzing Judicial Policies

- A large literature hand-codes cases by their policy choices and looks at downstream impacts.
  - (Currie and MacLeod 2008; Belloni et al 2012; Ash and Chen 2017)
  - Necessarily focuses on relatively small sets of decisions
  - Arguably relies on subjective judgments about what policy is being decided and the direction of the decision.
- A parallel literature in law and computation has used unsupervised learning methods to classify large corpora of cases by topic based on language and other features.
  - (e.g. Livermore, Riddell, and Rockmore, 2016, Ash and Chen 2018)
  - The resulting topics may or may not coincide with policy classes
Analyzing Judicial Policies

- A large literature hand-codes cases by their policy choices and looks at downstream impacts.
  - (Currie and MacLeod 2008; Belloni et al 2012; Ash and Chen 2017)
  - Necessarily focuses on relatively small sets of decisions
  - Arguably relies on subjective judgments about what policy is being decided and the direction of the decision.

- A parallel literature in law and computation has used unsupervised learning methods to classify large corpora of cases by topic based on language and other features.
  - (e.g. Livermore, Riddell, and Rockmore, 2016, Ash and Chen 2018)
  - The resulting topics may or may not coincide with policy classes
This Paper

- We propose and implement a semi-supervised clustering model, trained on hand-coded dataset of thousands of cases making one of 22 policy choices.
  - Using opinion text features as predictors, our model can predict policy classes with 91% accuracy in held out test sample.
  - We then take the model to the broader set of unlabeled cases and ask whether it can identify new groups of cases by shared policy impact.
This Paper

- We propose and implement a semi-supervised clustering model, trained on hand-coded dataset of thousands of cases making one of 22 policy choices.
  - Using opinion text features as predictors, our model can predict policy classes with 91% accuracy in held out test sample.
  - We then take the model to the broader set of unlabeled cases and ask whether it can identify new groups of cases by shared policy impact.
This Paper

- We propose and implement a semi-supervised clustering model, trained on hand-coded dataset of thousands of cases making one of 22 policy choices.
  - Using opinion text features as predictors, our model can predict policy classes with 91% accuracy in held out test sample.
  - We then take the model to the broader set of unlabeled cases and ask whether it can identify new groups of cases by shared policy impact.
Outline

Background
Ash-Chen 2018: Case Vectors
Ash-Chen-Ornaghi 2018: Implicit Bias in Judge Language
Ash-Morelli-Osnabrugge 2018: Policy Topic Model

Automated Policy Classification

Concluding Thoughts
Previous Work: Impacts of Judicial Decisions

Example papers looking at impacts of judicial decisions:
- Belloni et al (2012), eminent domain decisions affect housing prices
- Chen and Sethi (2017), sexual harassment protections reduce gender wage inequalities
- Ash and Chen (2017), religious liberties cases affect religiosity
- Ash, Chen, and Holbein (2018), assembly rights cases affect protest rates

These papers rely on hand-coded data sets, which are costly to collect.
- What if a machine could learn all the policies for us?
Previous Work: Impacts of Judicial Decisions

- Example papers looking at impacts of judicial decisions:
  - Belloni et al (2012), eminent domain decisions affect housing prices
  - Chen and Sethi (2017), sexual harassment protections reduce gender wage inequalities
  - Ash and Chen (2017), religious liberties cases affect religiosity
  - Ash, Chen, and Holbein (2018), assembly rights cases affect protest rates

- These papers rely on hand-coded data sets, which are costly to collect.
  - What if a machine could learn all the policies for us?
Recent advances in natural language processing have stemmed from using dense vectors to represent language relations:

- **Topic models for encoding relations between documents (e.g. LDA, Blei 2003)**

Livermore, Riddell, and Rockmore (2016) apply LDA to understand the content of Supreme Court decisions.
Embeddings Models

- Word embeddings and document embeddings encode relations between words, phrases, and documents in geometric space (e.g. Mikolov et al 2013, Le and Mikolov 2014).

1. Finding the degree of similarity between two words.
   ```python
   model.similarity('woman','man')
   0.73723527
   ```

2. Finding odd one out.
   ```python
   model.doesnt_match('breakfast cereal dinner lunch';.split())
   'cereal'
   ```

3. Amazing things like woman+king-man = queen
   ```python
   model.most_similar(positive=['woman','king'],negative=['man'],topn=1)
   queen: 0.508
   ```

4. Probability of a text under the model
   ```python
   model.score(['The fox jumped over the lazy dog'.split()])
   0.21
   ```

- Ash and Chen (2018) train document vectors on a corpus of U.S. Circuit Court and Supreme Court cases.
Outline

Background

Ash-Chen 2018: Case Vectors
Ash-Chen-Ornaghi 2018: Implicit Bias in Judge Language
Ash-Morelli-Osnabruegge 2018: Policy Topic Model

Automated Policy Classification

Concluding Thoughts
Ash-Chen 2018: Case Vectors by Circuit

Figure 1: Centered by Topic-Year, Averaged by Judge, Labeled by Court
Ash-Chen 2018: Case Vectors by Decade

Figure 2: Centered by Court-Topic, Averaged by Court-Year, Labeled by Decade
Ash-Chen 2018: Case Vectors by Topic

Figure 3: Centered by Judge-Year, Averaged by Topic-Year, Labeled by Topic
Ash-Chen 2018: Case Vectors by Party

Figure 4: Centered by Court-Topic-Year, Averaged by Judge, Labeled by Political Party
<table>
<thead>
<tr>
<th>Circuit Judge Name</th>
<th>Similarity</th>
<th>Rank</th>
<th>Circuit Judge Name</th>
<th>Similarity</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>POSNER, RICHARD A.</td>
<td>1.000</td>
<td>1</td>
<td>TONE, PHILIP W.</td>
<td>0.459</td>
<td>16</td>
</tr>
<tr>
<td>EASTERBROOK, FRANK H.</td>
<td>0.663</td>
<td>2</td>
<td>SIBLEY, SAMUEL</td>
<td>0.459</td>
<td>17</td>
</tr>
<tr>
<td>SUTTON, JEFFREY S.</td>
<td>0.620</td>
<td>3</td>
<td>SCALIA, ANTONIN</td>
<td>0.456</td>
<td>18</td>
</tr>
<tr>
<td>NOONAN, JOHN T.</td>
<td>0.596</td>
<td>4</td>
<td>COLLOTON, STEVEN M.</td>
<td>0.445</td>
<td>19</td>
</tr>
<tr>
<td>NELSON, DAVID A.</td>
<td>0.592</td>
<td>5</td>
<td>DUNIWAY, BENJAMIN</td>
<td>0.438</td>
<td>20</td>
</tr>
<tr>
<td>CARNES, EDWARD E.</td>
<td>0.567</td>
<td>6</td>
<td>GIBBONS, JOHN J.</td>
<td>0.422</td>
<td>21</td>
</tr>
<tr>
<td>FRIENDLY, HENRY</td>
<td>0.566</td>
<td>7</td>
<td>BOGGS, DANNY J.</td>
<td>0.420</td>
<td>22</td>
</tr>
<tr>
<td>KOZINSKI, ALEX</td>
<td>0.563</td>
<td>8</td>
<td>BREYER, STEPHEN G.</td>
<td>0.414</td>
<td>23</td>
</tr>
<tr>
<td>GORSUCH, NEIL M.</td>
<td>0.559</td>
<td>9</td>
<td>GOODRICH, HERBERT</td>
<td>0.412</td>
<td>24</td>
</tr>
<tr>
<td>CHAMBERS, RICHARD H.</td>
<td>0.546</td>
<td>10</td>
<td>LOKEN, JAMES B.</td>
<td>0.410</td>
<td>25</td>
</tr>
<tr>
<td>FERNANDEZ, FERDINAND F.</td>
<td>0.503</td>
<td>11</td>
<td>WEIS, JOSEPH F.</td>
<td>0.408</td>
<td>26</td>
</tr>
<tr>
<td>EDMONDSON, JAMES L.</td>
<td>0.501</td>
<td>12</td>
<td>SCALIA, ANTONIN (SCOTUS)</td>
<td>0.406</td>
<td>27</td>
</tr>
<tr>
<td>KLEINFELD, ANDREW J.</td>
<td>0.491</td>
<td>13</td>
<td>BOUDIN, MICHAEL</td>
<td>0.403</td>
<td>28</td>
</tr>
<tr>
<td>WILLIAMS, STEPHEN F.</td>
<td>0.481</td>
<td>14</td>
<td>RANDOLPH, A. RAYMOND</td>
<td>0.397</td>
<td>29</td>
</tr>
<tr>
<td>KETHLEDGE, RAYMOND M.</td>
<td>0.459</td>
<td>15</td>
<td>MCCONNELL, MICHAEL W.</td>
<td>0.390</td>
<td>30</td>
</tr>
</tbody>
</table>

Most similar circuit court judges to Richard Posner based on document vectors.
Outline

Background
- Ash-Chen 2018: Case Vectors
- Ash-Chen-Ornaghi 2018: Implicit Bias in Judge Language
- Ash-Morelli-Osnabruegge 2018: Policy Topic Model

Automated Policy Classification

Concluding Thoughts
Implicit Bias in Judge Language: Ash-Chen-Ornaghi (2018)

Positive (left) and negative (right) language:
- energetic, efficient, sincere, helpful, thoughtful, confident, reasonable, dependable, strong, reliable, organized, natural
- irritable, complaining, moody, impatient, stubborn, impulsive, unstable, self-centered, sarcastic

Innocent (left) and guilty (right) language:
- peaceable, anxious, realistic, cooperative, unafraid, independent, practical, reserved, unmoved, quiet, obliging, unselfish, beneficent
- coarse, cruel, selfish, irresponsible, irresponsible, untrustworthy, disorderly, impulsive, dominating, pretentious

Male names are more associated to “positive” language and female names are more associated to “innocent” language.
No racial relation to “positive” language, but black-sounding names are more associated to “guilty” language
Hispanic versus Guilty/Innocent
post-1970 corpus

Hispanic versus Guilty/Innocent
democratic vs republican judges

▶ Hispanic surnames more associated to “guilty”, and stronger association for Democrats
Outline

Background
Ash-Chen 2018: Case Vectors
Ash-Chen-Ornaghi 2018: Implicit Bias in Judge Language
Ash-Morelli-Osnabruegge 2018: Policy Topic Model

Automated Policy Classification

Concluding Thoughts
Policy Topic Model: Ash, Morelli, and Osnabruegge (2018)

- Comparative Manifesto Project Corpus:
  - 44,020 annotated English-language political statements from hundreds of political party platforms since 1980.
  - Each statement gets a topic code, e.g. “decentralization”, “education”
    - 45 topics, normalized to 19 broader, more interpretable topics
  - Regularized logit model trained to predict these topics from statement text (53% out-of-sample accuracy)
Policy Topic Model: Ash, Morelli, and Osnabruegge (2018)

- Comparative Manifesto Project Corpus:
  - 44,020 annotated English-language political statements from hundreds of political party platforms since 1980.
  - Each statement gets a topic code, e.g. “decentralization”, “education”
    - 45 topics, normalized to 19 broader, more interpretable topics
- Regularized logit model trained to predict these topics from statement text (53% out-of-sample accuracy)
Policy Topic Model: Ash, Morelli, and Osnabruegge (2018)

- Comparative Manifesto Project Corpus:
  - 44,020 annotated English-language political statements from hundreds of political party platforms since 1980.
  - Each statement gets a topic code, e.g. “decentralization”, “education”
    - 45 topics, normalized to 19 broader, more interpretable topics
  - Regularized logit model trained to predict these topics from statement text (53% out-of-sample accuracy)
## Ash-Morelli-Osnabruegge 2018: Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration</td>
<td>270</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>77</td>
<td>12</td>
<td>71</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>78</td>
<td>11</td>
<td>36</td>
<td>3</td>
<td>11</td>
<td>0</td>
<td>101</td>
</tr>
<tr>
<td>Agriculture</td>
<td>5</td>
<td>114</td>
<td>0</td>
<td>3</td>
<td>24</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>23</td>
<td>1</td>
<td>56</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Culture</td>
<td>8</td>
<td>3</td>
<td>155</td>
<td>5</td>
<td>22</td>
<td>21</td>
<td>22</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>19</td>
<td>4</td>
<td>56</td>
<td>3</td>
<td>18</td>
<td>7</td>
<td>17</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>Decentralization</td>
<td>27</td>
<td>2</td>
<td>6</td>
<td>73</td>
<td>16</td>
<td>11</td>
<td>24</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>32</td>
<td>7</td>
<td>20</td>
<td>0</td>
<td>13</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>Economics</td>
<td>59</td>
<td>12</td>
<td>5</td>
<td>4</td>
<td>715</td>
<td>13</td>
<td>29</td>
<td>16</td>
<td>19</td>
<td>2</td>
<td>11</td>
<td>4</td>
<td>138</td>
<td>30</td>
<td>96</td>
<td>1</td>
<td>49</td>
<td>2</td>
<td>126</td>
</tr>
<tr>
<td>Education</td>
<td>7</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>17</td>
<td>461</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>75</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>37</td>
<td>3</td>
<td>69</td>
</tr>
<tr>
<td>Freedom</td>
<td>51</td>
<td>0</td>
<td>9</td>
<td>20</td>
<td>41</td>
<td>13</td>
<td>642</td>
<td>19</td>
<td>3</td>
<td>16</td>
<td>12</td>
<td>6</td>
<td>126</td>
<td>34</td>
<td>16</td>
<td>4</td>
<td>14</td>
<td>7</td>
<td>89</td>
</tr>
<tr>
<td>Internationalism</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>33</td>
<td>2</td>
<td>45</td>
<td>245</td>
<td>2</td>
<td>14</td>
<td>11</td>
<td>1</td>
<td>46</td>
<td>9</td>
<td>19</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>Labor Groups</td>
<td>11</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>36</td>
<td>7</td>
<td>12</td>
<td>2</td>
<td>92</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>20</td>
<td>3</td>
<td>13</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>Military</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>24</td>
<td>27</td>
<td>2</td>
<td>118</td>
<td>5</td>
<td>0</td>
<td>32</td>
<td>2</td>
<td>10</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Nat'l Way of Life</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>38</td>
<td>3</td>
<td>32</td>
<td>17</td>
<td>2</td>
<td>6</td>
<td>88</td>
<td>6</td>
<td>61</td>
<td>21</td>
<td>19</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>44</td>
</tr>
<tr>
<td>Non-Econ Demo Grps</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>15</td>
<td>15</td>
<td>30</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>95</td>
<td>36</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>104</td>
</tr>
<tr>
<td>Other Topic</td>
<td>38</td>
<td>14</td>
<td>17</td>
<td>13</td>
<td>78</td>
<td>48</td>
<td>86</td>
<td>25</td>
<td>20</td>
<td>12</td>
<td>20</td>
<td>26</td>
<td>1263</td>
<td>39</td>
<td>71</td>
<td>20</td>
<td>48</td>
<td>7</td>
<td>159</td>
</tr>
<tr>
<td>Party Politics</td>
<td>20</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>44</td>
<td>2</td>
<td>48</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td>80</td>
<td>183</td>
<td>25</td>
<td>4</td>
<td>10</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>Quality of Life</td>
<td>21</td>
<td>17</td>
<td>9</td>
<td>4</td>
<td>102</td>
<td>9</td>
<td>13</td>
<td>14</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>112</td>
<td>16</td>
<td>684</td>
<td>0</td>
<td>42</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>Target Groups</td>
<td>11</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>17</td>
<td>24</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>28</td>
<td>4</td>
<td>1</td>
<td>67</td>
<td>8</td>
<td>1</td>
<td>63</td>
</tr>
<tr>
<td>Tech &amp; Infra</td>
<td>22</td>
<td>8</td>
<td>5</td>
<td>8</td>
<td>57</td>
<td>31</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>85</td>
<td>1</td>
<td>62</td>
<td>1</td>
<td>413</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>Trad'l Morality</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>9</td>
<td>21</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>25</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Welfare</td>
<td>33</td>
<td>6</td>
<td>11</td>
<td>3</td>
<td>93</td>
<td>32</td>
<td>62</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>9</td>
<td>26</td>
<td>142</td>
<td>37</td>
<td>29</td>
<td>11</td>
<td>23</td>
<td>7</td>
<td>1173</td>
</tr>
</tbody>
</table>

| Total Predicted        | 611    | 194   | 261   | 151      | 1428  | 710   | 1206   | 404    | 182   | 185    | 229       | 186   | 2458           | 423     | 1190   | 131 | 717    | 104     | 2244       |
Log party politics topic share by year. Error spikes give 90% confidence intervals.

- After moving to proportional representation (dashed vertical line), party politics speeches increased significantly.
Outline

Background
Ash-Chen 2018: Case Vectors
Ash-Chen-Ornaghi 2018: Implicit Bias in Judge Language
Ash-Morelli-Osnabruegge 2018: Policy Topic Model

Automated Policy Classification

Concluding Thoughts
What if we don’t know all the policy classes?

- One could try LDA, but topics are not equivalent with policies.
- Ideally, the model could learn the structure of existing policy classes and taken to unlabeled data to learn the new classes.
This project

- What if we don’t know all the policy classes?
  - One could try LDA, but topics are not equivalent with policies.
  - Ideally, the model could learn the structure of existing policy classes and taken to unlabeled data to learn the new classes.
Data

- 326,000 cases from Federal Circuit Courts.
  - 7,600 have been hand-labeled into 22 policy classes:
  - All cases hand-labeled to 82 legal areas
    - Administrative Law, 'Bankruptcy Law', Energy Law, etc.

- Featurization:
  - Doc2Vec used to embed cases as document vectors (Le and Mikolov 2014).
    - DBOW specification, window size = 12, minimum gram frequency = 50, dimensions = 100
326,000 cases from Federal Circuit Courts.
7,600 have been hand-labeled into 22 policy classes:

All cases hand-labeled to 82 legal areas
- Administrative Law, 'Bankruptcy Law', Energy Law, etc.

Featurization:
- Doc2Vec used to embed cases as document vectors (Le and Mikolov 2014).
  - DBOW specification, window size = 12, minimum gram frequency = 50, dimensions = 100
Document Embeddings Recover Policy Classes

t-SNE plot in two dimensions, labeled cases.
Policy Label Classification

- Inputs:
  - Document vectors

- Labels:
  - Policy classes (11 labels with at least 50 member cases)

- Model:
  - Regularized logistic (parameters selected by cross validation)

- Validation:
  - Classification accuracy on 20% held out sample
Out of Sample Classification Accuracy: 91%
K-Means Clustering in Doc2Vec Space

- **Validation:**
  - First, we ask whether these clusters recover policy classes in labeled data.
  - Second, we ask whether the clusters tend to replicate topics in unlabeled data.
- Model parameters chosen to improve performance on these tasks.
  - Tested 185 models.
K-Means Clustering in Doc2Vec Space

- Validation:
  - First, we ask whether these clusters recover policy classes in labeled data
  - Second, we ask whether the clusters tend to replicate topics in unlabeled data
- Model parameters chosen to improve performance on these tasks.
  - tested 185 models
K-Means Clusters Recover 15 out of 22 Policy Classes

GM for 1537 test sample
K-Means Clusters Recover Topics in Unlabeled Data
Promises of the method

- The clusters discovered in the unlabeled data could then be analyzed for what policy decision is being made, if any.
  - This would probably have to be done manually.
- Once we have the policies, we can use random assignment of judges for exogenous variation in those policies, and then look at downstream societal impacts.
  - This could be done without interpreting the policies.
  - We could test for impacts first, and then analyze/interpret the most impactful policy classes.
Promises of the method

- The clusters discovered in the unlabeled data could then be analyzed for what policy decision is being made, if any.
  - This would probably have to be done manually.
- Once we have the policies, we can use random assignment of judges for exogenous variation in those policies, and then look at downstream societal impacts.
  - This could be done without interpreting the policies.
  - We could test for impacts first, and then analyze/interpret the most impactful policy classes.
Promises of the method

- The clusters discovered in the unlabeled data could then be analyzed for what policy decision is being made, if any.
  - This would probably have to be done manually.
- Once we have the policies, we can use random assignment of judges for exogenous variation in those policies, and then look at downstream societal impacts.
  - This could be done without interpreting the policies.
  - We could test for impacts first, and then analyze/interpret the most impactful policy classes.
Some Limitations/Caveats

- Costly to validate (probably requires manual checking)
- Assumes cases make a single policy decision
- Probably won’t work on other courts.
- Many more...
Outline

Background
  Ash-Chen 2018: Case Vectors
  Ash-Chen-Ornaghi 2018: Implicit Bias in Judge Language
  Ash-Morelli-Osnabruegge 2018: Policy Topic Model

Automated Policy Classification

Concluding Thoughts
Promises and challenges

- This research explores the potential for automated policy analysis in the judiciary.
  - These methods may help us understand better the causes and consequences of judicial decision-making.
  - But method is costly to validate, limiting large-scale applications.
- New corpora and new computational models have promise for leading to a richer understanding of law and the legal system, but major challenges remain – traditional methods still needed.
Promises and challenges

▶ This research explores the potential for automated policy analysis in the judiciary.
  ▶ These methods may help us understand better the causes and consequences of judicial decision-making.
  ▶ But method is costly to validate, limiting large-scale applications.

▶ New corpora and new computational models have promise for leading to a richer understanding of law and the legal system, but major challenges remain – traditional methods still needed.
Promises and challenges

- This research explores the potential for automated policy analysis in the judiciary.
  - These methods may help us understand better the causes and consequences of judicial decision-making.
  - But method is costly to validate, limiting large-scale applications.
- New corpora and new computational models have promise for leading to a richer understanding of law and the legal system, but major challenges remain – traditional methods still needed.