The Role Partnerships between Community Banks and Fintech Companies Played in PPP Loan Distribution

Gio DeFrank, Junghee Mun, Kristina Posner
<table>
<thead>
<tr>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis Overview</td>
<td>Related Literature and Research</td>
<td>Data Sources</td>
<td>Data Collection and Cleaning Methodology</td>
</tr>
<tr>
<td>05</td>
<td>06</td>
<td>07</td>
<td></td>
</tr>
<tr>
<td>Model &amp; Methodology: Impact of Fintech Partnerships on Distribution of Loans to Minorities</td>
<td>Model &amp; Methodology: Impact of Fintech Partnerships on Distribution of Loans to Low and Moderate Income areas</td>
<td>Further Discussion &amp; Conclusion</td>
<td></td>
</tr>
</tbody>
</table>
01 Hypothesis Overview
Our Hypothesis:
Community Banks who partnered with financial technology (Fintech) firms were ultimately more successful in distributing PPP loans to small businesses with racial minorities and underbanked populations.
Related Literature and Research
Current Literature Overview

- Community Banks played an outsized role distributing PPP loans to small businesses during the pandemic
- Fintechs are more likely to reach minority and underbanked populations than conventional banks
- Anecdotal evidence suggests Community Bank and Fintech partnerships played a large role in distributing PPP loans to disadvantaged populations during the pandemic

(Community banks play outsized role, 2020) (Erel & Liebersohn, 2020)
Model Summary and Justification

- Implemented logistic regression to evaluate the impact a Fintech Partnership had on the odds of a PPP loan being distributed to minority & LMI labeled businesses
- Our data violated half of the required assumptions of logistic regression

Random Forest

- Implemented random forest model to validate/test the results of the logistic regression model
- Random forest models do not have assumptions
- Random forest is an intuitive and easy-to-understand model
- Achieved accuracy rate of 67% for predicting minority owned businesses and 70% accuracy for predicting businesses located in LMI areas
Data Sources

- PPP Loan Data
  - Approximately 8 million rows
  - Provided loan level information
  - Formed the core of our analysis

- FDIC Community Bank Designation Data
  - Requirements that define Community Banks
  - Additional information about each bank's financials

- FDIC Yearly “How America Banks” Survey
  - FDIC data about America's unbanked population
  - Percent unbanked vs percent banked
  - Confidence intervals for provided statistics
Data Collection and Cleaning Methodology
Our Data-Cleaning/Collection Methodology

**Step 1**
- Used Python script to web-scrape Google for articles that mention Fintech & Community Bank partnerships

**Step 2**
- Cleaned, combined, and summarized data sources using Alteryx

**Step 3**
- Built a classification neural network model to predict the probability of a business being minority owned. Used predicted values to fill in NAs within the minority column
Web Scraping Process

Utilized SerpAPI to web-scrape Google

Read each article, validated the partnership, and generated a list of 41 banks that had Fintech partnerships

Collected over 100 links that mentioned Fintech + Community Bank partnerships

Identified 237K loans that were likely distributed by Community Bank + Fintech partnerships

237K
Alteryx: Preprocessing and Data-Cleaning

- Cleaned Data
  - Combined various data sources
  - Cleaned data
  - Filled in NAs
  - Summarized data
  - Split data into train and test sets
Classification Neural Network Model

- Problem:
  - The minority column contained a large amount of empty rows

- Feedforward model:
  - Experimented with different number of layers and nodes
  - Experimented with various learning and momentum rates

- Final model:
  - 6 hidden layers (varying number of nodes within each layer)
  - Learning rate of 0.2
  - Momentum rate of 0.7
  - Accuracy of 72%
Minority Model Methodology
Logistic Regression Assumptions:

- #1: Linearity of the Logit
- #2: Absence of Multicollinearity
- #3: Lack of Strongly Influential Outliers
- #4: Independence of Errors
Classification
Random Forest
Minority Model

- Utilized the number of trees with highest accuracy: **300**
- Overall accuracy of **66.83%**

<table>
<thead>
<tr>
<th>Number of Trees</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>66.73%</td>
</tr>
<tr>
<td>200</td>
<td>66.76%</td>
</tr>
<tr>
<td>300</td>
<td><strong>66.83%</strong></td>
</tr>
<tr>
<td>400</td>
<td>66.77%</td>
</tr>
<tr>
<td>500</td>
<td>66.77%</td>
</tr>
</tbody>
</table>

Figure 6
Importance Report: Minority Model Results

- Fintech Partnership has the lowest importance in predicting PPP Borrower as Minority
- Fintech Partnership doesn’t contribute to increasing Minority outreach
Low and Moderate Income (LMI) Model Methodology
Classification
Random Forest
LMI Model

- Utilized the number of trees with highest accuracy: **100**
- Overall accuracy of **69.864%**

<table>
<thead>
<tr>
<th>Number of Trees</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>69.864%</td>
</tr>
<tr>
<td>200</td>
<td>69.859%</td>
</tr>
<tr>
<td>300</td>
<td>66.856%</td>
</tr>
<tr>
<td>400</td>
<td>55.859%</td>
</tr>
<tr>
<td>500</td>
<td>66.857%</td>
</tr>
</tbody>
</table>

Figure 8
Importance Report: LMI Model Results

- Fintech Partnership has the lowest importance in predicting PPP Borrower as LMI
- Fintech Partnership doesn’t contribute to increasing LMI outreach
Further Discussion & Conclusion
Further Possibilities

1. Further data gathering & cleaning
   - Lack of publicly available data
   - Limited options of metrics and predictors

2. Web scraping
   - “Brute forced” list of Fintechs from scrapings
   - Potential user error and data leakage

3. Minority prediction using ANN classifier
   - Dependent on predictive library’s validity
   - Accuracy cap: 72%

4. Neo-bank classification
   - Dependent on FDIC’s ruling update
   - Good precedent for future evolution of banks
Implications

— Fintech partnership was not a significant factor for community banks handling PPP loans
— However, may help acceleration with banks’ digitalization
— Caution for hasty digitalization without establishing security measures and/or compliance protocols
— Safety measures and guidelines should be mandated by governmental bodies
Conclusion

Hypothesis
- Community bank’s partnership with fintech did not help to reach minority and LMI population with PPP loans in 2020

Fintech Partnership
- Help to diversify the banking ecosystem
- Need for adequate regulations is ever-present
- More research is needed

Best Model
- Random forest classification
- Accuracy: 70%
