DARTMOUTH UNDERGRADUATE JOURNAL OF POLITICS, ECONOMICS AND
WORLD AFFAIRS

VOLUME 1 | ISSUE 2

MASTHEAD

Editor-in-Chief
Jacob Meister

President
Mason Wegener

Executive Editors
Sahil Abbi
Emma Plumb

Associate Editors
Jonathan Briffault
Isabella Brooks
Addison Dick
Theodore Hill-Weld
Andrew Sosanya

Submissions Director
Maxwell Brautigam

Design Director
Alexander Montague
LETTER FROM THE EDITOR

We exist in a moment of unprecedented connection to our fellow human, and the effects are both awesome and disturbing. Near infinite growth in global interaction and communication means that our relatively small journal can reach thousands, that humanitarian efforts can reach millions, and that the well-intentioned may connect to spread positivity around the globe. However, exposure to new ideas and new ways of life simultaneously incites conflict in those unprepared for potentially paradigm-altering consequences. Racism, economic hardship, and systematic violence are just a few localized results of this increasingly globalized world.

Nevertheless, we must not turn away from new ideas simply because they are not readily compatible with the status quo. Our communities, countries, and world will continue to evolve, whether we wish for that evolution or not. The only responsible recourse, therefore, is to rigorously explore, develop, and interrogate those ideas so that we may shape the world as we wish to see it. In 1825, Dartmouth College alumnus Daniel Webster foretold, “It is a wonderful fellowship of individual intelligences which will make up the minds and opinions of the age.” What follows is such a fellowship, comprised of undergraduate students from across the country and distilled into eight articles which we hope may help make the minds of the age.

I would like to acknowledge the entire DUJPEW staff for their efforts over the past two years. This journal would not exist without these close friends and collaborators: Mason, for your steadfast leadership and guidance; Emma, for your unyielding and indispensable support; and Sahil, for your willingness and ability to fill any role asked of you. This issue in particular would not exist without our new staff, who will hopefully look to the past two editions for inspiration and guidance in the many yet-to-come.

Likewise, I must thank the network of mentors, professors, and readers who continue to place their faith in our relatively new enterprise, engaging and connecting us to new scholarship across the globe. But most of all, I would like to thank the eight extraordinary authors who provided the critical pieces necessary for the present edition. This journal is a platform for your work, and I will be forever grateful for your willingness to engage with the world and expand the limits of human knowledge.

Jacob Meister ’20
Editor-in-Chief

*Dartmouth Undergraduate Journal of Politics, Economics and World Affairs*
ARTICLES

Paul-Angelo dell’Isola- Columbia University

27 Playing the Race Card: Did Racism Cause Anti-Immigrant Sentiment in Brexit?
Lucy Hu- University of Pennsylvania

55 Ecotourism and the Middle East: Lessons in Development
Andrew Ver Steeg- Amherst College

88 Faith in Democracy: Korean Churches as Engines of Pro-Democracy Protest
Will Matheson- Harvard University

111 Till Death Do Us Part: Will Longstanding Rivalry Impede the Ethnic Coalition of ISIS and Al Qaeda?
Bianca Pergher- American University

143 The Cost of the Culturati: Studying the Neighborhood Stability Impact of Cultural District Designations
Prakash Mishra- University of Pennsylvania

197 Token Representation?: The Impact of Female Reservations in Panchayati Raj Institutions in State and National Elections
Surbhi Bharadwaj- Yale University

243 Empowerment, Resistance and the Pill: A Feminist Analysis of Contraception in the Developing World
Abigail Trombley- Wake Forest University
The ‘Security Pacific Letter’:
Estimating the Causal Effect of Securitization on Banks’ Systemic Exposure

Paul-Angelo dell’Isola
Abstract

This paper aims to test the hypothesis of the ‘Safe Asset narrative’ which states that banks became manufacturers of pseudo safe assets to meet a global shortage of safe assets in the pre-crisis period. In this narrative, securitization is the mechanism which enables banks to become underwriters of safe assets. This paper takes this hypothesis to the data and attempts to estimate the causal effect of securitization on banks’ systemic exposure. In particular, this paper exploits a regulatory change that occurred in 1987 when the OCC expanded the scope of assets US national banks could securitize. By using state-chartered banks as a control group and estimating a diff-in-diff model, I find that securitization significantly increased banks’ systemic exposure. I then provide evidence on changes of banks’ balance sheet features to pinpoint a direct channel through which securitization may have increased banks’ systemic exposure.

This paper aims to test the hypothesis of the ‘Safe Asset narrative’ which states that banks became manufacturers of pseudo safe assets to meet a global shortage of safe assets in the pre-crisis period. In this narrative, securitization is the mechanism which enables banks to become underwriters of safe assets. This paper takes this hypothesis to the data and attempts to estimate the causal effect of securitization on banks’ systemic exposure. In particular, this paper exploits a regulatory change that occurred in 1987 when the OCC expanded the scope of assets US national banks could securitize. By using state-chartered banks as a control group and estimating a diff-in-diff model, I find that securitization significantly increased banks’ systemic exposure. I then provide evidence on changes of banks’ balance sheet features to pinpoint a direct channel through which securitization may have increased banks’ systemic exposure.
### Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>8</td>
</tr>
<tr>
<td>Literature Review &amp; Theoretical Motivation</td>
<td>9</td>
</tr>
<tr>
<td>Identification: The ‘Security Pacific Letter’</td>
<td>13</td>
</tr>
<tr>
<td>Data &amp; Methodology</td>
<td>16</td>
</tr>
<tr>
<td>The Effect of Securitization on Banks’ Systemic Exposure</td>
<td>19</td>
</tr>
<tr>
<td>Discussion: Identification of a Potential Direct Channel</td>
<td>22</td>
</tr>
<tr>
<td>Conclusion</td>
<td>24</td>
</tr>
<tr>
<td>Notes</td>
<td>26</td>
</tr>
</tbody>
</table>
Introduction

Both the subprime mortgage crisis and Eurozone sovereign-debt crisis can be understood as crises of pseudo-safe assets engineered by the private sector in the US and issued by fiscally-weak countries in Europe. The ‘Safe Asset narrative’ assumes that a global shortage of safe assets facilitated the excessive issuance of debt instruments during the pre-crisis period. In such a scheme, the development of securitization may have been a channel through which the safe-asset shortage materialized in the real economy. Indeed, securitization allowed banks to engineer pseudo-safe assets by transforming pools of illiquid assets (mortgage, commercial and retail loans) into liquid (tradable) securities. The resulting assets were labelled as “safe” prior to the crisis even though they could not insure against systemic risk.

This paper aims to test the hypothesis that banks engaged in securitization as a response to a safe assets shortage, thus becoming manufacturers of ‘safe assets’ and underwriters of insurance against systemic risk. In particular, I deliver causal evidence on the effects of securitization on banks’ systemic exposure. Assessing the impact of securitization on banks’ exposure to systemic risk presents an empirical challenge as securitization indirectly fueled the growth of a national real estate mortgage market, thus indirectly exposing banks to systemic shocks. I overcome this identification challenge by estimating the causal impact of a regulatory change that affected US banks independently from their connections to the real estate market. By exploiting a regulatory change to the range of permissible assets that national banks could securitize and using state-chartered banks as a control group, I find that securitization significantly increased banks’ systemic exposure.

Having established a causal link between securitization and banks’ systemic exposure, I provide further evidence on structural changes to banks’ balance sheets so as to identify the
channel through which national banks increased their exposure to systemic risk. In doing so, I test the empirical prediction of the ‘Safe Assets narrative’. In the new ‘Originate and distribute’ banking model that developed after 1990, banks were allowed to originate assets and redistribute the associated risks through securitization instead of keeping those assets on their balance sheets. I find that national banks significantly decreased the share of their loans financed by deposits, which implies that banks increased their reliance on alternative sources of funding to finance their lending activities. This evidence is consistent with my initial hypothesis and pinpoints a channel through which securitization may have increased banks’ systemic exposure. Had securitization been aimed at merely diversifying the banks’ funding structure, we would not observe such increase in systemic exposure.

The rest of the paper is structured as follows: Section 2 reviews the existing economic literature on the ‘Safe Asset narrative’ and provide theoretical motivation for my empirical hypothesis. Section 3 presents my identification strategy and Section 4 describes the data. Section 5 and 6 present the results of my diff-in-diff estimation and further evidence of changes to banks’ balance sheet features. Section 7 concludes.

Literature Review & Theoretical Motivation

The ‘Safe Asset Narrative’ of Financial Crises

A recent strand of economic literature on the demand and supply of monetary aggregates has emerged to explain the cyclical over-issuance of debt instruments leading to debt crises. In this framework, monetary aggregates – the so-called ‘safe assets’- encompass money and money-like financial instruments that are used as cash and transacted without much concern for adverse selection in financial markets. Unlike the ‘Minskyan approach’ which emphasizes the role of
agents’ expectations about the value of liquidity, the ‘Safe Asset narrative’ focuses on the global asymmetric issuance of safe assets and assumes that a global shortage of safe assets during the pre-crisis period facilitated the excessive issuance of debt instruments in advanced economies who were suppliers of those scare assets. In this framework, the subprime mortgage crisis can be comprehended as a crisis of pseudo-safe assets engineered by the private sector in the US to meet the needs of a globally integrated financial system.

In spite of its intuitive explanatory power, this narrative is difficult to test empirically. In this regard, Gorton, Lewellen and Metrick provide some evidence of a structural change to the manufacturing of safe assets. While the share of safe assets has remained constant over the last 30 years, the authors point at the change in the components of this share. Governments and banks – traditional manufacturers of safe assets through sovereign debt issuance and banks’ deposits certificates – have been replaced by the “shadow banking” sector, which now produces a substantial fraction of those assets. Indeed, money market mutual funds shares, commercial papers, repurchase agreements and securitized debt are all money-like instruments that can be used as collateral in financial markets.

*The ‘Safe Asset Narrative’ in Practice: Securitization and Systemic Risk*

Focusing on the rise of structured finance is therefore key to test the validity of the ‘Safe Asset narrative’. Indeed, this narrative implicitly assumes that the development of a new ‘Originate and distribute’ banking model was a response to the growing demand for safe assets from global investors. In this model, banks originate loans and resell them to outside investors instead of keeping those assets on their balance sheets. This process effectively distributes the associated risks to the rest of the financial sector and release banks from the constraint of keeping illiquid
assets on their balance sheet. This new model essentially offers alternative funding sources to the banks who are no longer dependent on their customers’ deposits. If the development of the ‘Originate and distribute’ funding model was a response to the safe asset shortage, the banks’ manufacturing of new financial instruments would be tantamount to the underwriting of insurance against systemic risk. In this regard, securitization was a channel trough which banks financed the origination of illiquid loans through the issuance of liquid, thus tradable securities. It is both a risk transfer and liquidity transformation mechanism whose resulting securitized products were bought for their perceived safety prior to the crisis.  

According to the ‘Safe Asset narrative’, the development of securitization should therefore be concomitant to an increase of banks’ systemic exposure. In their survey of the economic literature on the effect of growing interconnectedness and increased contagion risk in financial networks, Paul Glasserman and H. Peyton Young highlight the unsolved question of whether more interconnectedness tend to amplify or dampen systemic shocks. The ambiguous role played by securitization is at the heart of this tradeoff. On the one hand, banks can diversify their funding, thus better insuring themselves against systemic risk. On the other hand, securitization creates new obligations and increases the dominance of funding liquidity on market liquidity. The bank lending activity is more subject to adverse shocks as its financing depends on third agents’ perception of the resulting security liquidity.  

Glasserman and Young illustrate the groundwork for models of interconnected balance sheets with a basic intuitive scheme. Figure 1 show the stylized balance sheet of a bank $i$ who has two categories of assets. ‘Outside assets’ $c_i$ are claims on non-financial entities such as households or corporations while ‘In-network assets’ $p_{ik}$ are claims on any financial entity $k$ such as banks or...
asset managers. Likewise, the bank has ‘Outside liabilities’ $b_l$ towards depositories and ‘In-network liabilities’ $p_{lk}$ towards financial entities.

I used their stylized representation of banks’ balance sheets to contrast the ‘traditional banking’ model with the ‘Originate and distribute’ model. The two figures represent the structural features of a bank’s balance sheet in the two distinctive banking models. In the traditional banking model, the bank finances the origination of its outside assets with outside liabilities, thus somehow isolating its traditional lending business from market fluctuations. In the ‘Originate and distribute model’, the bank combines outside liabilities and in-network liabilities to finance its outside assets. Implicitly, this structure increases the bank’s financing liquidity and allows the bank to increase the size of its outside assets. Consequently, this stylized representation of a bank’s balance sheet sheds light on the new risk of asymmetry caused by securitization. The origination of bank assets is made increasingly dependent on the liquidity of asset-backed securities and their perceived safety.

Estimating the effect of securitization on banks’ systemic exposure is therefore key to explore the tradeoff between better risk diversification and increased dominance of the safe asset liquidity on the banks’ origination business. Indeed, securitization may have been a channel
through which the safe asset global shortage materialized. Before the financial crisis of 2007, the general consensus was that securitization supported financial stability. This view was prevalent among policy makers such as the Office of the Comptroller of the Currency (OCC), which recognized the benefits of securitization:

The need for liquidity and the ability to engage in sound asset-liability management practices is all the more important to the maintenance of a safe and sound banking system.

Securitization was aimed to cater investors’ need for safe assets as the payments tranching allowed to meet several liquidity needs:

Structural credit enhancement and diversified asset pool free investors of the need to obtain a detailed understanding of the underlying loans.

Consequently, the development of securitization can be viewed as a ‘safe asset’ engine. If securitization was adopted to create a safe asset, thus turning banks into underwriters of systemic risk insurance, we should observe an increase in treated banks’ systemic exposure. Testing this hypothesis could shed light on the role played by bank regulation in enabling this structural change.

Identification: The ‘Security Pacific Letter’

Identification

This paper aims to estimate the causal effect of securitization as a new funding source on banks’ systemic exposure. Disentangling the effect of securitization as a new funding source from contemporaneous changes in banks’ systemic exposure represents an empirical challenge. Indeed, the effect of securitization is twofold. On the asset side, securitization can indirectly increase a bank’s exposure to systemic risk if the greater amount of assets originated are highly correlated with systemic events. In the US, securitization fueled the growth of mortgage-backed securities.
Together with other government policies that encouraged home ownership through various tax subsidies and home finance programs, securitization turned the real-estate market into a national one and strengthened the transmission line between a shock to the real economy and a bank crisis. Analyzing the role played by securitization in the increase of banks’ systemic exposure could therefore lead to overestimation. The aim of my identification strategy is thus to test the direct effect of securitization as an alternative funding source for banks, independently from its indirect effects on the underlying assets markets. I overcome this identification challenge by exploiting a regulatory change regarding securitization which is independent from the story of the real estate mortgages in the US.

The ‘Security Pacific Letter’

In this section, I argue that the 1987 ‘Security Pacific Letter’ constitutes a valid natural experiment to test the causal effect of securitization on banks’ systemic exposure by differentiating between national commercial banks and state-chartered banks responses.

In the 1980s, the process of securitizing bank assets became much easier. In a 1987 letter to Security Pacific National Bank – the so-called ‘Security Pacific Letter’ –, the OCC expanded the scope of permissible assets national bank could resell as securitized products. In spite of the applicable Glass Steagall Act, the OCC allowed the underwriting of securitized assets by stating that national banks could securitize and sell ‘any of their […] lawfully acquired assets’. The supervisor of federal-chartered banks assumed that the pass-through certificates – representing claims on the trust holding the underlying illiquid assets - were not ‘securities’ within the meaning of Glass-Steagall Act and that securitization did not consist in the securities underwriting of the securities business. The Comptroller’s ruling was challenged by the Securities Industry...
Association which argued that the OCC’s decision allowed national banks to deal in securities, thus creating a breach to the GSA which clearly separated the banking business from the securities business. The U.S. Court of Appeals upheld the OCC ruling in 1989. The court stated that:

*If the activity constitutes the “business of banking,” then the Glass-Steagall Act prohibitions ... do not apply.* (885 F. 2d 1034, 1048)

Hence, the court’s decision sanctioned the sale by a national bank of any type of securitized assets, including the banks’ own consumer credit card receivables, automobile and boat loans, commercial loans and leases. Following this court ruling, the OCC started approving many securitization programs of national banks after 1989.

I propose to use this regulatory cutoff to isolate the effect that securitization had on national banks, independently from the underlying securitized assets. This regulatory change represents an empirical setting where I can differentiate between a treatment and control group, thus satisfying an essential condition of causal inference in observational studies. In this setting, state-chartered banks constitute an ideal control group as they were subject to the Glass-Steagall Act (GSA) but not to the OCC ruling. Indeed, the GSA provisions applied to all FDIC insured banks, be they state-chartered or federally-supervised while OCC rulings only applied to national banks. Given that both state-chartered banks and national banks were subject to the GSA, I propose to use state-chartered banks as a control in this natural experiment. My identification strategy therefore consists in estimating the change in a bank’s systemic exposure following the regulatory change, based on whether or not the FDIC insured bank was a state-chartered bank (control) or a national bank (treatment).
Data & Methodology

The banks’ betas are used as proxies for their systemic exposure. A bank beta ($\beta_i$) arises from co-movements of the stock price with the market and therefore captures the systemic component of a firm’s stock price as opposed to its idiosyncratic component.

I obtained the lists of all FDIC insured banks from 1982 to 1992 from the Federal Deposit Insurance Corporation website. Estimating my model using FDIC-insured banks only – be they state-chartered (control) or OCC-chartered (treatment) – was key to ensure that they were subject to the same regulations prior to 1987. The FDIC dataset indicates whether each affiliated bank was supervised by the OCC or had a state charter. I matched this dataset with the banks’ betas time series that I retrieved from the Bloomberg database. To do so, I used the FDIC variable indicating whether the individual bank was held by a holding company and obtained a list of state-chartered banks and OCC-chartered banks at the consolidated level. Given that each bank in the control group needs to be state-chartered and that the outcome variable is observed at the level of the holding company (‘consolidated’) only, matching the data at the consolidated level avoids identifying treated banks as state-chartered (control) in situations where the holding company of different state-chartered banks has an OCC-charter and should therefore be considered as treated.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FDIC banks</td>
<td>13,204</td>
</tr>
<tr>
<td>State-chartered banks</td>
<td>9,757 (73%)</td>
</tr>
<tr>
<td>OCC-chartered banks</td>
<td>3,447 (27%)</td>
</tr>
</tbody>
</table>

Table 1: FDIC dataset
Using a market-based measure of systemic risk puts a constraint on my sample. Indeed, many of the affiliated banks, be they national or state-chartered, were privately held. In order to obtain the most accurate measure of national and state-chartered banks’ systemic risk, I decided to remove all holding companies’ whose subsidiaries were not only banks but also insurance companies or other financial firms (‘diversified financials’). Indeed, the other lines of businesses of such companies would have created uninformative noise to my response variable. I also checked that each of my sample bank did not change its regulatory status during my period of interest (1982-1992).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Publicly traded banks (1982-1992)</td>
<td>482</td>
</tr>
<tr>
<td>State-chartered banks</td>
<td>264</td>
</tr>
<tr>
<td></td>
<td>(54%)</td>
</tr>
<tr>
<td>OCC-chartered banks</td>
<td>218</td>
</tr>
<tr>
<td></td>
<td>(46%)</td>
</tr>
</tbody>
</table>

Table 2: Bloomberg dataset

Lastly, I retrieved data on some of my sample banks’ total assets, total loans, total deposits and total liabilities from Mergent Online, a database of corporate information covering US and foreign companies from 1982 onward. A first overview of the differences in banks’ systemic exposure before and after the 1987 regulatory cutoff date is provided below. The changes in the distribution of the OCC-chartered banks’ betas should serve as a preliminary guess regarding the sign and magnitude of the causal effect of securitization.
Figure 3: Differences in means before and after 1987 across the treatment and control group

To assess the internal validity of my empirical strategy, I first used my sample to test the parallel trend assumption, which is critical to any diff-in-diff identification strategy. Ensuring that both the control and treatment group have parallel trends in their outcomes values prior to the treatment should support the appropriateness of the chosen control group as a counterfactual. I therefore plotted the moving average of each bank’s stock beta time series, covering the period from 1982 to 1992. Figure 3 shows the moving average trend of each bank group – national and state-chartered banks.

Consistent with my identification assumption, Figure 3 shows that state-chartered and national banks’ betas followed the same trend prior to 1987 when the OCC released the ‘Security
Pacific Letter’. I formally tested the robustness of this identification assumption in section 5.C.

Figure 4: Per-group moving average of banks’ betas.

The Effect of Securitization on Banks’ Systemic Exposure

Model Specification

I used the banks panel data to estimate the following diff-in-diff model:

\[ \beta_{it} = \alpha_i + s_t + \theta D_{it} + \epsilon_{it} \] (1)

In this model, each bank \( i \)’s beta at time \( t \) – thereafter denoted \( \beta_{it} \) – is regressed on its entity fixed effects \( \alpha_i \), time fixed effects \( s_t \) and a treatment dummy \( D_{it} \).

\[ D_{it} = \begin{cases} 1, & \text{bank } i \text{ is a national bank after 1987} \\ 0, & \text{otherwise} \end{cases} \]

Note that the coefficient of interest in this regression model is \( \theta \), whose sign and magnitude indicates whether the regulatory shock increased or decreased national banks’ systemic exposure.
Results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>OLS</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Treatment $\theta$</td>
<td>0.067***</td>
<td>0.067***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.683*</td>
<td>0.683*</td>
</tr>
<tr>
<td></td>
<td>(0.407)</td>
<td>(0.407)</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Entity fixed effects</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>27,639</td>
<td>27,639</td>
</tr>
<tr>
<td><strong>R2</strong></td>
<td>0.491</td>
<td>0.490</td>
</tr>
<tr>
<td><strong>Adjusted R2</strong></td>
<td>0.480</td>
<td>0.479</td>
</tr>
<tr>
<td><strong>Residual Std. Error</strong></td>
<td>0.404</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(df = 27067)</td>
<td></td>
</tr>
<tr>
<td><strong>F Statistic</strong> (df = 571; 27067)</td>
<td>45.672***</td>
<td>45.672***</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1: The ‘Security Pacific Letter’ effect on national banks’ systemic exposure
Table 1 reports the main results from the estimated model. Column 1 shows the point estimates derived using Ordinary Least Squares (OLS). Column 2 shows the point estimates of the same model estimated with standard errors clustered at the bank level.

The two estimated coefficients have high statistical significance (one percent) and are robust to clustered standard errors. The point estimates on the treatment dummy variable are pretty consistent across all estimated models and robust to outlying values. The above estimated model thus suggests that the OCC decision increased national banks’ systemic exposure by 0.07. The tables consequently support the view that exposure to securitization increased banks’ systemic exposure.

Robustness of Parallel Trend Assumption

In order to assess the robustness of my identification assumption, I estimated the following model:

$$\beta_{it} = \alpha_t + s_t + \sum_i \rho_i (G_i \times s_t) + \epsilon_{it} \tag{2}$$

In this model, I regress banks’ betas on time and entity fixed effects and an interaction term between time fixed effects $s_t$ and a dummy variable $G_i$ indicating whether the bank $i$ belonged to the treatment group (national bank) or control group (state-chartered bank).

$$G_i = \begin{cases} 1, & \text{bank } i \text{ is a national bank} \\ 0, & \text{bank } i \text{ is a state chartered bank} \end{cases}$$

This model aims at testing the parallel trend assumption, which is a necessary condition to estimate causality in a difference-in-differences empirical setting. In particular, estimating the difference in the mean response of the treatment and control group implies that the assignment of treatment should be insignificant before the treatment date. The coefficients of interest in this specification are therefore the series of coefficients on the interaction term between group
assignment and time fixed effects $\rho_t$. In the hypothesis that state-chartered banks constitute a valid control group, the series of $\rho_t$ should equal 0 before the regulatory cutoff date and be significantly different from 0 after the treatment.

The time series of $\rho_t$ and their 99% confidence intervals are plotted in figure 4.

![Coefficients on Group Dummy](image)

*Figure 4: Coefficients on the Group Dummy*

Figure 5: Time series of the Group assignment statistical significance

Figure 4 supports the parallel trend assumption of this empirical setting. The point estimates are statistically insignificant before 1987 and present statistical significance after 1990.

**Discussion: Identification of a potential direct channel**

My results consequently point at an indirect effect of securitization on banks’ systemic exposure. However, this empirical setting does not allow me to pinpoint the direct channel through which the regulation impacted banks’ balance sheet and therefore their systemic exposure. To identify the potential channel explaining the estimated indirect effect, I used the panel data on banks’ balance sheets retrieved from Mergent online. Based on random sample of 50 banks, I
found data on only 25 banks’ balance sheets. I then computed the Loan to Deposits ratio of each bank in 1986 and 1992. I propose this ratio as a proxy for the share of a bank’s lending activity that is financed through deposits. In a traditional banking model, this ratio should therefore be close to 1 as the bank only uses outside liabilities (deposits) to fund the origination of outside assets (loans). Table 3 shows the growth rate of the Loan to Deposits Ratios for both national and state-chartered banks between 1986 and 1992.

<table>
<thead>
<tr>
<th></th>
<th>Growth rate of the ratio $\frac{\text{Total Loans}<em>{t}}{\text{Total Deposits}</em>{t}}$ in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Banks</td>
<td>(-16%)</td>
</tr>
<tr>
<td>State-chartered banks</td>
<td>(-5%)</td>
</tr>
</tbody>
</table>

Table 3: Growth rate of Loan to Deposits ratios

I decided not to run any formal diff-in-diff model because of the small sample size. Instead, this data is only aimed at giving a hint of the direct mechanism at stake and should serve as a basis for future research using this empirical setting. Both national and state-chartered banks decreased their loan to deposits ratios. Hence, the share of outside assets financed by outside liabilities decreased for both groups of banks, thus explaining a common upward trend of in their systemic exposure. Furthermore, the difference in percentage growth shed light on the effect of securitization on banks’ systemic exposure. The decline in the share of outside assets funded by outside liabilities between 1986 and 1992 is significantly higher for national banks. This result suggests that national banks were able to increase the size of their loan portfolio through In-network liabilities, thus increasing their exposure to systemic risk.
These results are consistent with the ‘Safe Asset narrative’. In particular, the causal increase in banks’ systemic risk as a result of securitization validates a key prediction of the safe asset framework. Indeed, had the banks engaged in securitization for other reasons than meeting a shortage of safe assets through the engineering of liquid securities, we would not necessarily have observed an increase in their systemic risk. Securitization would have merely resulted in a change in the funding structure of the banks and the better diversification of their funding sources may have resulted even in lower systemic risk. Instead, my results pinpoint at a mechanism that is accounted for in the ‘safe asset narrative’. Securitization did not only result in a change in the funding structure of banks but also incentivized them to originate new outside assets so as to sell pseudo safe securities to in-network investors. Hence, the increase in banks’ systemic risk puts in evidence a change in banks’ incentives and their engagement in a new business aimed at catering a growing demand for safe assets. Instead of attempting to measure changes in global investors’ demand for safe assets in order to validate the ‘safe asset narrative’, my approach thus looks for evidence on the sell-side of new incentives created by a safe asset shortage.

**Conclusion**

This paper presents new evidence on the causal effect of bank deregulation on banks’ systemic exposure. By using state-chartered banks as a control group and exploiting a regulatory change to the permissibility of securitization, I find that securitization significantly increased banks’ systemic exposure.

These findings provide further empirical evidence for the validity of the ‘Safe Asset narrative’. Consistently with the intuition of this narrative, banks seemed to have increased the underwriting of safe assets through securitization, thus mechanically increasing their exposure to
systemic risk. While the estimated causal effect shows that securitization has an indirect effect on banks’ systemic exposure, I use banks’ balance sheet data to pinpoint the structural changes at the bank level that drove this result. I find that banks that were able to securitize their assets significantly increased the share of in-network liabilities used to finance the origination of assets. Thus, securitization seems to act as an alternative funding source for banks, thereby increasing their systemic exposure by reinforcing the dominance of funding liquidity on their lending business.
Notes

6 Ibid.
9 Ibid.
Playing the Race Card:
The Role of Racism in Anti-Immigrant Sentiment in Brexit

Lucy Hu
Abstract

Following the 2016 referendum on continued UK membership of the European Union, many attempts were made to explain its result. There has been consensus that the issue of immigration played a primary role in the Leave campaign and Brexiteers’ minds. The reasons for this anti-immigrant sentiment have been explored, with economic and cultural concerns at the fore of the literature. Critically, currently missing from the debate is whether racism played a substantial role in causing anti-immigrant sentiment in the context of Brexit. This article uses new public opinion data from 2018 to investigate the extent to which racism played a part in the Leave vote. It found that racism was an important predictor of referendum vote choice, even when economic concerns were held constant. Among all levels of financial satisfaction, people respond to their sociological concerns when making a political determination about immigration. Despite efforts from elites at the fore of the Leave campaign to rid the debates of racism, exclusively economic arguments proved to be a façade for private racist attitudes of many Leave voters. While concern over cultural pluralism is likely a complementary factor, this article finds the link between anxieties over skin color and anti-immigrant sentiment.
## Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>30</td>
</tr>
<tr>
<td>Theory and History</td>
<td>32</td>
</tr>
<tr>
<td>Brexit Explained</td>
<td>38</td>
</tr>
<tr>
<td>Data Analyses</td>
<td>42</td>
</tr>
<tr>
<td>Discussion</td>
<td>46</td>
</tr>
<tr>
<td>Conclusion</td>
<td>51</td>
</tr>
<tr>
<td>Notes</td>
<td>53</td>
</tr>
</tbody>
</table>
Introduction

On Thursday June 23, 2016, 17.4 million people in the United Kingdom voted for their nation to leave the European Union, sparking an immediate political frenzy. The Great British Pound dropped to a 31-year low shortly after polls closed on Thursday evening. On Friday morning at 8:15 a.m., Prime Minister David Cameron announced his resignation after a six-year premiership.

The UK’s relationship with the EU has consistently been questioned by both British elites and the public — and other European nations — since its first application to join in 1961 and eventual accession in 1973. The first referendum on membership of the European Community (as it was then known), held in 1975, revealed that 67 percent of voters supported the UK’s accession. The debate nevertheless continued, and by 2015, the public mood was clear — immigration reform was a priority, according to opinion polls. From June 2015 to February 2016, Cameron and his team renegotiated the terms of the UK’s membership, hoping to convince the public that remaining within the EU yielded more benefits than costs. Whether Cameron’s discussions with Brussels were enough was now up to Britons. The June referendum was officially announced on February 20, 2016.

In early 2016, the media, business world and financial markets assumed that the status quo would likely endure after June. Since the results of the referendum, an impressive post-mortem has been conducted on the Leave vote. Plenty of literature has sought to explain the surprise result by investigating campaign strategies of both sides, trends in opinion polling, and the political environment surrounding the debate. There has been consensus that the largest driver of the Leave vote, both in terms of campaign strategy and public opinion, was the prevalent feeling of the need for immigration control. Leading up to 2016, elites pushed messages of the perceived economic
detriments that accompany immigration. Analyzing elite rhetoric belies potential racist motivations in anti-immigration.

This article contributes to understandings of the sociological aspects of anti-immigration sentiment in the context of the Leave vote. Specifically, it is an investigation of how British public opinion on race manifests into a pervasive anti-immigration societal attitude, placed in the context of culture, identity, and xenophobia. In other words, this article seeks to explain to what extent the Leave vote was caused by issues of racial identity, an explanation that has been underplayed in deference to economic arguments.

Current literature suggests that anti-immigrant sentiment within the EU membership debates is caused primarily by fear of cultural pluralism and the perception of the economic burden that immigrants bring. A gap in the literature exists in accounting for the role of racism (discrimination based on skin color, separate from cultural xenophobia) in the Brexit debates. In the context of immigration, this article evaluates the role of racial concerns, in comparison to economic concerns, in pushing Britons towards the Leave vote.

The next section outlines theoretical and historical frameworks for understanding public opinion on identity, culture, and race, as well as whether these opinions contribute to anti-immigrant sentiment more than economic concerns. It also explains some media effects in causing anti-immigrant attitudes. Section three provides a review of current literature on explaining the result of the 2016 referendum, including a focus on explaining opposition to immigration. Section four dives into empirical analyses of public opinion data, evaluating the comparative explanatory power of race and economic concerns within the 2016 referendum debates. Section five discusses these data, evaluating them against existing literature and frameworks of understanding, as well as previous opinion polling.
Theory and History: Public Opinion on Identity Issues in Immigration

Scholarship on public opinion on immigration has been widespread. This review focuses on a few areas of identity politics concerns within anti-immigrant sentiment. Firstly, it looks at research on British attitudes towards different ethnic groups and immigrants from differing regions. Secondly, it reviews research on people’s fear of cultural pluralism and concerns over the potential threat of immigrants to national identity. This article seeks to compare the importance of identity politics and economic concerns in shaping anti-immigrant sentiment, so the third section of this chapter gauges existing studies that execute this multivariate analysis.

Ethnic Hierarchy

Researchers have conducted and analyzed various public opinion polls investigating discrepant attitudes towards different immigrant groups. Specifically, relevant results compare immigrants’ region of origin and race. Shamit Saggar, a leading public policy scholar on British race politics, found that negative sentiments are not purely based on color, citing a 2001 ICM poll which showed that every age group of respondents emphatically disapproved of having Romanian immigrants as neighbors. Almost every age group approved of having Chinese neighbors. The poll also revealed that almost every group of respondents approved of white South African asylum seekers as neighbors, while almost every group disapproved of Afghan and Iraqi asylum seekers, suggesting a complex mix of attitudes relating to differing perceptions based on national origin.1

While Saggar found opposition to certain white Eastern European immigrants, diminishing the importance of skin color, a wealth of scholarship points to the importance of the white versus non-white divide. Public opinion scholar and University of Manchester professor Robert Ford (2011) pointed to historical evidence emphasizing that non-white immigration is far more opposed...
by the public than white immigration. Early settlements of Polish wartime refugees and workers from Mediterranean Europe and Ireland caused little public resistance compared to the hostility faced by South Asian and Caribbean migrants. ‘Race riots’ occurred in Notting Hill (1958), Toxteth and Brixton (1981), and Bradford and Burnley (2001) in areas with concentrations of black and Asian immigrants and their descendants,\(^2\) while there have never been riots of this scale against white immigrants. By 2008, the population of West European migrants had surpassed that of Indian-born migrants, which did not result in public interest, suggesting the higher acceptability of white immigrants.\(^3\)

Historical evidence in terms of policy also reveal ethnic hierarchies in the minds of the public. While Western European migrants have been uncontroversial, Eastern European arrivals have attracted significant media attention and are monitored by the Home Office.\(^4\) Within non-white migration, a ranking ladder is also apparent. The 1950s saw Westminster’s defense of migration from the West Indies; this population was regarded as more ‘British’ than South Asians.\(^5\) Recently, Muslims have been seen as more ‘threatening’ than Afro-Carribbeans.\(^6\)

Ford also used data from six British Social Attitudes surveys to provide further evidence of an ethnic hierarchy, with each white ethnic group above the non-white groups. The surveys, conducted between 1983 and 1996, reveal that the least opposed group of immigrants are Australians and New Zealanders, followed by Western Europeans, Eastern Europeans, and then (in order of preference) immigrants from Hong Kong, Africa, the West Indies, and South Asia.\(^7\)

Nationalism scholar Eric Kaufmann (2014) compared Eastern Europeans with non-white groups from other regions, and found that while white British ‘habituate’ to both Eastern European and non-European immigrant groups, assimilation occurs far more rapidly among European immigrants than minorities. Forty-six percent of people classified as children of ‘Mediterranean
Commonwealth’ immigrants on the 1971 Census considered themselves as dominant white British by 2011. Meanwhile, there is a majority attitude that visible minorities are not accepted as fully ‘one of us.’

Cultural Unity

As for cultural concerns, replicated research has highlighted that perceptions of immigrant groups differ depending on the group’s perceived threat to cultural unity and national identity. Saggar (2003) cited the 1997 British Election Study which revealed resistance to cultural pluralism. Of white respondents, 71 percent agreed that ‘it is far better for immigrants to adapt and blend into society,’ rather than keep their customs and traditions. Kaufmann (2014) noted that ethnic minorities are perceived as the greatest threat when immigrants are close but not too close (the ‘halo effect’). Diversity in one’s city or local authority adds to threat perceptions because of immigrants’ potential to introduce large-scale change, but Kaufmann’s study showed that positive contact with minorities mitigated fears.

That Anglo-Saxon settlement happened in England over a thousand years ago, coupled with the limited immigration prior to 1948, consolidated a ‘native’ English ethnic group. Most western European ethnic majorities use their nation’s name to describe their ethnic group, providing an indigenousness that acts as an obstacle to introducing multiculturalism into their national identities.

Are Racial and Cultural Concerns Stronger than Economic Ones?

The two types of identity-politics-based concerns that voters may hold regarding immigration are race (or skin color) and culture — the ideas, social behavior and customs that
differ among societies. Another factor that may lead to anti-immigration is economic concerns, for example immigrants’ perceived burden on national resources, or the added competition for jobs.

Multivariate experiments are particularly useful to answer the question of why people harbor anti-immigrant sentiments. Studies have compared the relative importance of racism, concerns over multiculturalism, the economy, and other influences. There is a multitude of research that justifies anti-immigrant sentiment in economic terms, for example a survey that found high levels of resentment towards asylum seekers and recent migrants, as opposed to minorities, over the concern of the use of public resources.\textsuperscript{12}

However, multiple studies indicate that a preference for cultural unity is the strongest predictor of anti-immigrant sentiment. An experimental study found that Dutch hostility to immigrants is amplified by describing them in cultural rather than economic terms.\textsuperscript{13} While there exists research that explains public opinion on immigration in terms of economic concerns, most studies which compare economy and identity explanations find that cultural concerns more strongly influence opinions on immigration.\textsuperscript{14} Another study directly tested the variables of race against welfare concerns and economic competition, and found race to be the key determinant.\textsuperscript{15}

Ford (2011) once again used historical evidence to argue against the notion that hostility to non-white immigration is driven by fear of increased economic competition for jobs and government resources. Far-right mobilizations by the National Front in the 1970s,\textsuperscript{16} and the British National Party since 1999,\textsuperscript{17} have been on non-white immigrants. Irish, European and Australian immigrants have often migrated to Britain in larger numbers but Ford discussed the lack of political movements against these groups. Although Ford’s article predates the 2016 referendum by five years, he argued that there is strong historical evidence of non-white migrants garnering more passionate and organized opposition than white migrants.\textsuperscript{18}
University of Bergen professor Elisabeth Ivarsflaten’s (2005) influential study used multivariate analysis to pit cultural unity concerns against other factors in determining individual preferences for immigration policy. The study found that a preference for cultural unity matters more than all other variables taken together, including concerns about the national economy and racism. In the model of maximum impact, where respondents feel strongly about one explanatory variable but average on other concerns, belief in cultural unity is five times more influential than concerns about the national economy. Racism is also influential in this model (meaning that strong feelings of racism can have a significant impact on people’s preferences), but in the model showing current impact of each variable observed today, racism is not more influential than concerns about the economy.¹⁹

Other scholars have discussed the racialization of economic concerns. Immigration arouses a sense of resentment from the white working class, for whom inequality and perceptions of unfairness are compounded by fear of multiculturalism and the visible changes in the community around them.²⁰

Erroneous Perceptions

It is necessary to note that a significant contributor to anti-immigrant sentiment may be a lack of awareness of the proportion and size of immigrants, as well as minorities, in the population. A 2000 MORI/Reader’s Digest poll found that the average estimate of the size of the ethnic population was 26 percent of the UK population, at a time when the actual figure was seven percent. The poll also revealed that the average estimate for the proportion of migrants and asylum seekers was 51 percent of the population, when in reality it was four percent. Among the public, ethnic minorities, immigrants and asylum seekers are often viewed as one group. Several polls
reveal that more intolerant population subgroups are more likely to have erroneous views of migrant and ethnic populations.\(^{21}\)

Scott Blinder, a political psychology scholar of US and UK politics, found in a 2015 study that both EU and non-EU nationals appear prominently in mental images of immigrants. Sixty-two percent of respondents pictured non-EU nationals as immigrants, while 51 percent pictured EU nationals, which overstated the 2010 EU-immigrant proportion of 30 percent.\(^{22}\) Two significant influencers of public opinion are media effects and elite influence in the form of policymaking, with the former having the potential to heavily impact perceptions of the numbers and characteristics of immigrants.

Past literature has suggested that incorrect estimates of immigrant proportions may stem from the large quantity of media coverage of immigrants or the tendency to inflate the prominence of salient minority groups.\(^{23}\) In terms of the composition of different immigrant types, media framing has been found to be an effective tool in influencing perceptions. Blinder’s (2018) analysis of British newspaper coverage from 2013 revealed that the two most common modifiers of ‘immigrant’ were ‘illegal’ and ‘EU’/’European,’ consistent with analyses from the past decade, even though irregular migrants (those that lack legal status) only comprised 11 percent of the foreign-born population and there were more immigrants from outside the EU than within.\(^{24}\) However, racial framing is rare. Although national origin is frequently mentioned, the avoidance of racial terms may be attributed to efforts by anti-immigration elites to ‘detoxify’ immigration.\(^{25}\)

Blinder’s own experiment explicitly tested the effects of framing on perceptions, but it found that an ‘illegal’ frame did not influence perceptions on the size of this group. This was attributed to the fact that ‘illegal’ was already the most common modifier, and thus, participants were already overly exposed to the phrase for the experiment to add any further effect.\(^{26}\)
A 2016 paper by Blinder et al. moved beyond explaining quantitative perceptions to explore the effect of the ‘illegal’ frame on qualitative perceptions of immigrants. British media outlets’ consistent emphasis on illegal immigration causes illegality to be associated deeply with immigration, depicts immigrants as law-violators, and infuses immigrants with a taint of criminality, according to the conceptual metaphor theory. Language in news reports contributes to asylum seekers’ routine conflation with immigrants.27

Elite policymaking can also play a part in qualitative perceptions of immigrants. From the beginning of post-war immigration policy, there has been a preference for white immigrants. From the privilege of immigrants from white-majority former colonies28 to the lack of policy-maker interest at growing Western European immigration in the 2000s,29 elites have socialized the public to adopt diverging attitudes on different migrant groups. For example, Whitehall has been more willing to defend West Indian migration than South Asian, in reference to the former’s greater cultural proximity to ‘Britishness’.30

Ivarsflaten (2005) hypothesized that elites can influence public opinion by disseminating that reducing immigration is a solution to existing problems, thereby associating problems with immigration. Her experiment showed that while elites can successfully link some potentially-unrelated problems to immigration, other problems are already so strongly tied to immigration that elite visibility does not make a difference, for example a desire to preserve cultural unity.31

**Brexit Explained: Existing Research on the Leave Vote and Opposition to Immigration**

*Explaining the Leave vote*

In the lead up to the referendum, immigration was unequivocally the leading issue of the day, with 63 percent of voters selecting it as the most pressing issue at the end of 2015 in a YouGov poll, well ahead of healthcare (39 percent) and the economy (33 percent).32 Net migration reached
336,000, and 69 percent said that migration from the EU was too high.\textsuperscript{33} In their 2017 book \textit{Brexit: Why Britain Voted to Leave the European Union}, Clarke, Goodwin, and Whiteley’s regression analyses found that while evaluations of the economy, immigration and the National Health Service (NHS) each had statistically significant effects on support for EU membership, judgments about immigration actually exerted stronger effects than economic assessments.\textsuperscript{34} Their analyses also revealed that the benefit-cost variable of Britain’s ability to better control immigration in a post-Brexit world was one of the most significant predictors of the Leave vote.\textsuperscript{35}

Additionally, perceptions of potential post-Brexit immigration levels differed between Leave and Remain voters. Among Leave voters, 84 percent thought there would be less immigration if Britain left the EU, compared with 27 percent of Remainers, according to a YouGov poll.\textsuperscript{36} However, most polls conducted around the time of the referendum revealed that sovereignty was the number one stated reason that people voted Leave, and immigration was the second. YouGov’s poll showed 45 percent selecting sovereignty and 26 percent choosing immigration, echoed by Lord Ashcroft’s result of 49 percent selecting sovereignty and 33 percent selecting immigration. The British Election Study found around 30 percent for each.\textsuperscript{37}

The Leave campaign employed positive non-immigration messages including the renewal of sovereignty and democracy if the UK left the EU. This appealed to middle-class Eurosceptics who may have privately agreed with negatively rhetorical anti-immigration ideas but were conscious about potential tones of racism within the Leave campaign’s messaging. Another important message was the amount of money that could be redirected to the NHS if the UK left the EU. Although many Remainers contested the truth of this message, the idea that the NHS would be strengthened rather than weakened with a Brexit result resonated with Leavers.\textsuperscript{38} Additionally, many Leave voters from the left contested that the EU had become an ‘uber-capitalist’ club whose efforts threatened the possibility of nationalization of public services.\textsuperscript{39}
Overall, while there were other issues that contributed to Vote Leave, it is clear that immigration was the dominant concern. Even in surveys that placed sovereignty concerns over immigration, it is clear that the two are not mutually exclusive — in fact, to the contrary, the desire for sovereignty can be read as a desire to control immigration.

**Explaining Anti-Immigration Within the Leave Vote**

While this article investigates whether racism motivated anti-immigrant sentiment; existing literature has also generated differing explanations for this attitude, including the perceived increased job competition, burden on public resources, feelings of national identity, and threats to cultural unity.

Firstly, the perceived economic burden of immigration in terms of increased job competition has been a large contributor to anti-immigrant sentiment. A very timely and widely circulated 2015 Bank of England report documented that rising immigration could impact wages of UK-born people, especially in semi- or unskilled labor forces. This caught media attention immediately; a *Telegraph* headline from December 2015 read “Mass migration driving down wages offered to British jobseekers.” One study showed that the biggest impacts on wages of immigration occurred among low-wage workers, and that it was actually associated with an increase in wages of higher-paid workers, potentially creating widening inequality in the labor market. This could contribute to the differing views from different occupational groups about EU membership. Interestingly, the Migration Advisory Committee found that EU migrants actually appeared to have no impact on UK-born employment while non-EU migrants were associated with a reduction in UK-born employment.

Secondly, the perceived burden on the NHS is another economic concern that people had about immigration. The Vote Leave campaign strongly focused on the strain that immigrants exert
on welfare resources. At the beginning of 2016, healthcare professionals warned that the NHS was being ‘bled dry’ by ‘health tourists.’ In April, Vote Leave released statistics claiming to show increased waiting times from immigrants’ pressure on the system. This was effective as surveys showed that the percentage of people thinking that it was beneficial for the NHS to leave the EU was consistently higher than those who thought that remaining would help.45

National identity’s role in producing anti-immigrant sentiment has also been explored in existing literature. Clarke, Goodwin, and Whiteley (2017) found that national identity was a significant predictor for having positive or negative attitudes towards the benefits and costs of immigration. As is widely documented in anecdotal evidence, those who identified as English were more likely than those who identified as British to emphasize the benefits of leaving the EU.46 Britain is often viewed as an inclusive, outward-looking society in contrast to the English identity. European identity was a powerful predictor of the Remain vote,47 while English national identity is more associated with national sovereignty.48

Within the realm of identity, there has also been exploration on people’s perceptions on threats to cultural unity and minorities. There are widespread and entrenched negative attitudes towards minority groups among the whole British electorate, with survey data showing far more favorability towards the white majority than the four groups tested: Asians, Eastern Europeans, Blacks, Muslims.49 However, a strong majority of Leave voters and a solid half of Remain voters favored requiring EU migrants to go through the same application processes as non-EU migrants,50 indicating that the opposition to immigration in general may be based upon concerns with culture more than race. This fear of British cultural erosion is substantiated through investigations that find culture to be a bigger factor in anti-immigration than economic concerns. Sobolewska and Ford (2017) found that people who thought immigration posed a threat to culture did not always have the same misgivings about equal opportunity programs for minorities. In fact, only half of
those that were worried about culture said that equal opportunity programs had gone too far,\textsuperscript{51} indicating a deeper concern for cultural displacement than material. Some studies have also shown that concerns over national identity are more important than economic calculations in the context of European integration.\textsuperscript{52}

There is not an overwhelmingly dominant consensus in the literature about the relative importance of economic considerations and cultural, but the evidence for concerns about cultural erosion are strong. What follows is a deeper dive into how big of a role race played in determining aversion to immigration within the 2016 referendum.

**Data Analyses**

*Data and Methodology*

Analyses were conducted using Wave 8 of the UK Household Longitudinal Study (UKHLS), commonly known as Understanding Society. Wave 8 was conducted over two years from January 2016 to December 2017. Surveys were mostly conducted online, but households who had not participated were issued a face-to-face interviewer and there was also a small amount of telephone interviewing. The individual questionnaire portion of the survey was used, which surveyed 31,166 adults aged 16 and over from 16,015 households, who were mostly selected at Wave 1 in 2009.

*Analysis*

Wave 8 of the Understanding Society survey asks respondents how important “your ethnic or racial background” is to “your sense of who you are.” Using Wave 8 early release data, University of Bristol’s Paula Surridge, Siobhan McAndrew, and Neema Begum (2017) found that seeing race and ethnicity as important to one’s sense of identity was a statistically significant predictor of
voting Leave among White Britons. Among ethnic minorities, the importance of race to self-
identity was not statistically associated with either choice in the referendum.\textsuperscript{53} These results show
that, broadly, for the white population, concerns over race are likely to play into their opinions on
immigration.

But, to determine with more nuance the extent to which race plays a role in anti-immigrant
sentiment, a comparison of the strength of people’s concerns about race with their financial
corns was made through a series of different models and analyses. In order to compare racial
and economic attitudes, permutations of two variables were used. The first was the importance of
skin color to one’s identity, which was recoded into a binary variable (important/not important).
The second variable was satisfaction with income, which was also recoded into ‘satisfied’ and ‘not
satisfied.’ Subsets of people who satisfied different permutations of these two variables were tested
for their mean Leave-vote proportion.

The first question keeps income satisfaction constant: do people who are satisfied with
their income and think skin color is important vote in the same way as those who are also satisfied
with their income but do \textit{not} think race was important? The first is a key group as they are likely
to have less economic concerns, but worry about racial identity. The latter group is likely to be
also economically stable but more ‘color-blind.’ If race were not an important factor in the Leave
vote, we could expect the two groups to vote in the same way. From the first race-conscious group,
45.0 percent voted Leave, while 39.7 percent voted Leave from the non-race-concerned group.
This is a statistically significant difference (p-value \textless{} 10\textsuperscript{-10}), indicating that among people who
were satisfied with their income, those that thought their race was an important part of their identity
were more likely to vote to leave.
The second question keeps racial attitudes constant: is the first group (satisfied with income and race-conscious) less likely to vote Leave than those that are also race-conscious but dissatisfied with their income? If race were not an important factor in the leave vote, we could expect the first group to be less likely to vote Leave as they are more satisfied with their financial state. A difference-of-means test found that the two groups do in fact vote differently to a statistically significant degree (p-value = 1.66 * 10^{-5}). As stated above, 45.0 percent of the first group voted to leave, while 49.4 percent of the economically dissatisfied group did, when keeping racial attitudes constant. This indicates that among people who are racially-conscious, economic evaluations are still significant for the referendum vote.

A logistic regression was also modelled to substantiate these difference-of-means tests. The regression confirmed that when controlling for income satisfaction, the variable that measured the importance of race was a statistically significant predictor of the Leave vote (p-value = 0.005). The regression also found that income satisfaction was a powerful predictor of the Leave vote.

---

**Table 1**: Of those that are satisfied with their income, the Leave-vote proportion differs according to attitudes on race. p-value for difference of means for Leave vote: 6.66 * 10^{-15}

The next analysis tests the role of race within the Brexit debates in a different way. Under the null hypothesis, if race played no role, those that are dissatisfied with the economy should vote in the same way regardless of their concerns with skin color. Dissatisfaction with the economy is measured using three variables available in the Understanding Society dataset: income dissatisfaction, unemployment and lack of job security in the next 12 months. These variables were each further subset into two groups: those that thought race was important and those that did not. Since the race question is the only variable that differs, if the Leave-vote proportion differed between race-conscious and non-race-conscious groups of economically dissatisfied people, then opinions on race clearly played a role in determining the vote to leave.

The first test on people who were dissatisfied with their income found that there was a statistically significant difference between the mean Leave-vote proportion for those that thought race was important and those that did not. This indicates that for those dissatisfied with their income, attitudes on race had an effect on likelihood to vote Leave. The second measure of economic dissatisfaction — unemployment — found similarly significant results. Of those who were unemployed, racial attitudes also made a statistically significant difference, with 54.0 percent

<table>
<thead>
<tr>
<th>Voting Leave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance of race</td>
</tr>
<tr>
<td>(standard error)</td>
</tr>
<tr>
<td>Income dissatisfaction</td>
</tr>
<tr>
<td>(standard error)</td>
</tr>
</tbody>
</table>

Table 2: Results of logistic regression analysis.

Importance of race: 1 = important, 0 = not important
Income dissatisfaction: 1 = dissatisfied, 0 = satisfied
of the race-conscious unemployed voting Leave and 47.4 percent non-race-conscious unemployed voting Leave. The third measure of economic dissatisfaction was similar: for those that thought it was likely they would lose their job within the next 12 months, racial attitudes impacted decisions to vote Leave. It is evident that across three measures of economic dissatisfaction, attitudes on race still affected the vote to leave.

Discussion

Limitations

Racist attitudes are difficult to survey, and therefore survey questions do not always wholly represent true attitudes towards different racial groups. The variable in the Understanding Society dataset that was most closely related to a respondents’ potential views towards racial minorities was a question that asked how important “your ethnic or racial background” is to “your sense of who you are.” This question acts as an important proxy for attitudes on race. While this merely relates to one’s personal identity, placing greater importance on race generally translates into greater sensitivity to race. The assumption for the purposes of this discussion is that for white Britons, more value and importance in their own race, plus sensitivity to and concern for race, corresponds to a sense of superiority and/or more racist attitudes.

It is highly important to note the limitations in making such an assumption. It is possible to value one’s own white race or ethnicity in Britain and not view other racial groups negatively — since the question also specifies “ethnic background,” one probable example is a white Briton being proud of her Polish heritage. Additionally, for the non-white British population, placing importance on one’s ethnic identity does not have the same implication of superiority due to their minority status, which is a limitation not accounted for in the data. However, as mentioned,
Surridge, McAndrew and Begum (2017) found a statistically significant correlation between placing importance on race and voting Leave for white Britons, which intuitively adds conviction to the assumption for discussion purposes. Furthermore, ignoring the use of this question as a proxy for racism, simply taking the variable at face value (the importance people place on racial identity) yields valuable findings in itself.

Related limitations apply to the questions measuring economic concern. Similarly, these questions probe satisfaction with one’s own financial circumstance. As there were no questions evaluating the national economy, the questions used were: whether a respondent was satisfied with her income, whether she was unemployed, and a self-evaluation of likelihood to lose her job in the next 12 months. These three different variables were used to build a more robust model of national economic dissatisfaction. It is important to note that personal satisfaction does not always translate into national, and vice-versa. Nevertheless, how attitudes towards race vary by economic status in itself is revealing and worthwhile to discuss. With these limitations addressed, the discussion of results should be taken in a broad sense, indicative of avenues of future research.

**Economic Concerns Versus Racism**

Regression analysis demonstrated that when income was kept constant, racism was a statistically significant predictor of the Leave vote. This is consistent with intuition and past literature on attitudes towards racialized immigration in Britain.\(^5\) When respondents are subset into those that are satisfied with their income and those who are not, both groups demonstrated that evaluations of race contributed to the Leave vote.

A difference-of-means test confirmed that for respondents who were economically satisfied, racism mattered. Their attitudes toward race factor into their vote choice: if they are
simultaneously racist, they are more likely to vote Leave. As the survey question was about personal financial situation, it is unknown if this group has national economic concerns. But, even so, race did play a statistically significant role in pushing respondents one way or another. The vote was not simply about economics, contrary to some elite narratives, even for those that were financially dissatisfied. They did not all follow the same voting patterns; their racial attitudes mattered in determining their vote to leave.

Consistent with past literature but inconsistent with contemporary elite rhetoric in the Brexit campaigns, people responded to their sociological concerns, despite efforts from some in the Leave camp to shy away from race-baiting. A review of the Leave campaign’s strategy and media coverage suggests that economics should have been more important. Granted, the analyses show that for more racist respondents, economics mattered. Those that are racist were more likely to vote Leave if they were also dissatisfied economically, and less if they were not. Keeping immigration at the fore of their campaign, Vote Leave focused on economic and cultural concerns with mass migration, sometimes twisting economic facts. The previously mentioned Bank of England report that garnered significant media attention was repeatedly used and misused. A Telegraph headline from December 2015 read “Mass migration driving down wages offered to British jobseekers.” Sir Stephen Nickell, one of the authors of the report claimed that his findings had been misrepresented and that the impact was “infinitesimally small.” Conservative Member of Parliament Iain Duncan Smith said during the campaign that British wages are on average 10 percent lower because of EU immigration within the past decade. This claim was repeated by two senior figures in the campaign, including Boris Johnson. Nigel Farage, outspoken former leader of UKIP, had been consistently espousing for years prior to the referendum that EU immigrants
are a detriment to British jobs, saying in 2014 that there was a “massive oversupply” of foreign, unskilled labor, pushing wages of the British down.58

Additionally, immigrants’ burden on the NHS was a Leave camp favorite. Conservative MP Priti Patel said “Current levels of migration are causing unsustainable pressure on our public services and we can see that the NHS is creaking under the strain.”59 This is likely to fit with people’s intuitive logic that immigrants use NHS resources without contributing to its operation. The observed media and elite influence highlight efforts from the top to rid the debates of racism, instead using economic framing as much as possible, potentially as a façade of latent racism within sects of Leave voters. The data analyses demonstrate that latent racism was in fact still a major contributing force to the Leave vote, which the campaign did not dare to publicly admit. These results bolster the argument that economic framing was effective in appeasing social desirability effects of public rhetoric, while still allowing the Leave vote to flourish through private racist motivations.

National Identity

Using national identity survey questions from Wave 6 and combining it with Wave 8 of the Understanding Society data, Surridge, McAndrew and Begum (2017) found that having a strong British identity — responding to the question “How important is being British to you?” with a 6 or more out of 10 — was a significant predictor of support for leaving the EU. Notably, this identity had a statistically significant association with the Leave vote for both white Britons and non-White Britons.60 The fact that a strong national identity is associated with a desire to leave a supranational institution is not surprising, but the consistency among the white and non-white population is pertinent. Wanting to protect British national identity from the EU is therefore not
associated with ‘whiteness.’ For all races, feeling British led to a desire to reject the EU, suggesting that national identity concerns within the referendum debates were not framed around race.

**Culture**

People’s evaluations of race do not necessarily suggest the same about people’s attitudes towards cultural unity. Unfortunately, Wave 8 of the Understanding Society dataset’s questions on attitudes towards culture were exclusively asked in the ethnic minority boost sample. Thus, there was no way to analyze this data in conjunction with the referendum vote choice to yield an accurate picture of how concerns about cultural unity affected immigration debates in Brexit.

Despite these difficulties, both past literature and the current narrative point to the fact that people are protective of their culture. Section two of this article reviewed literature that supported the importance of cultural identity. The 1997 British Election Study found that 71 percent of white British prefer immigrants to adapt rather than keep their customs and traditions.\(^{61}\) It is clear that many white Britons prefer cultural unity to pluralism. This is not necessarily based on race, but since the white population is indigenous to Britain, perhaps there is a preference for sustenance of the native culture. Therefore, there could be resistance to immigration which would upset this culture. Since cultural concerns are not based on race, this hypothesis is consistent with wanting to limit immigration from the EU, even if the immigrants are white, because immigrants from within Europe also pose a cultural ‘threat.’ This article demonstrated that race was an important consideration in the Leave vote, but past literature has been able to place this within debates on culture. For example, Ivarsflaten’s (2005) model showed that concerns about cultural unity had a larger impact on anti-immigration than racism.

An analysis of UKIP strategy and support affirms the importance of cultural concerns. Public support for UKIP was partly motivated by feelings of anxiety over immigration, especially
when this could threaten people’s identity and culture. At a UKIP conference in 2014, leader Nigel Farage said that immigration was making parts of the country “unrecognizable” and like “a foreign land.” He also later expressed during a press conference his discomfort at hearing foreign languages being spoken on a train. In the final month of the Brexit campaign, Michael Gove and Boris Johnson, key Leave figures, promised a points-based system that would require migrants to speak English.

When elite influence is taken in conjunction with past British public opinion showing aversion to cultural pluralism, it is apparent that people’s evaluations of culture may have played a part within anti-immigration debates in Brexit, separate from race’s impact. This article demonstrated that racism was a predictor of the Leave vote, but did not solve the puzzle of why there would be aversion to EU immigrants based on racism, given that the majority of these immigrants are white. A very likely explanation is that those who are sensitive to race are also sensitive to cultural unity. This article’s claim that racism is a powerful predictor does not counter the potential predictive power of a preference for cultural unity. Further research is necessary to disentangle the two.

Conclusion

After June 2016, there has been a multitude of analyses that strived to dissect why Britain voted to leave the EU. Immigration was unanimously recognized to be the primary issue not only for the Leave campaign’s strategies, but also to voters. What caused anti-immigrant sentiment was a more complex question. The mobilization of concern over the perceived economic burden of mass migration was key, from increased job competition to a strain on welfare resources, namely the NHS. Analyses also showed strategies of cultural scaremongering to stoke public fears of cultural pluralism. There was also plenty of robust scholarship on public opinion towards immigration more generally, which also strived to compare economic concerns with sociological.
This article sought to locate race within the anti-immigration debate. It asked the question of whether anti-immigrant sentiment was motivated more by racism or concerns with the economy in the 2016 referendum. Using new Understanding Society data from 2018, this article found that race mattered even when holding economic evaluations constant in determining individual referendum choices.

The amalgam of past literature and this article’s analyses suggest that, although the British public and elites who advocated to exit the EU insisted that their campaign was not based on racism, people inherently respond to their sociological concerns when making a political determination about immigration. It is possible that many white Britons voted to curb immigration from the EU because they preferred that the UK retain its Anglo-Saxon culture — which would extend antipathy to include white EU immigrants. However, regardless of cultural attitudes, people’s racism was translated through the ballot box. Brexit was not simply about perceived declining economic conditions of the country — in 2016, the United Kingdom played the race card.
Notes:

3 Ford 2011, 1020.
4 Ford 2011, 1022.
5 Hansen, cited in Ford 2011, 1022.
7 Ford 2011, 1026.
10 Kaufmann 2014, 272.
11 Kaufmann 2014, 274-75.
13 Sniderman et al., cited in Ford 2011, 1019.
16 Husbands, cited in Ford 2011, 1020.
18 Ford 2011.
19 Ivarsflaten 2005.
21 Saggar 2003, 185.
23 Baker et al., cited in Blinder 2018, 1448.
26 Blinder 2018, 1455.
29 Ford 2011, 1022.
33 Clarke, Goodwin, and Whiteley 2017, 11-12.
34 ibid., 84.
35 ibid., 161.
38 Clarke, Goodwin, and Whiteley 2017, 59-60.
39 ibid., 15.
40 Nickell and Saleheen 2015, “The impact of immigration on occupational wages: evidence from Britain.”
41 Dominiczak and Spence 2015, “Mass migration driving down wages offered to British jobseekers.”
43 Clarke, Goodwin, and Whiteley 2017, 113.
44 Migration Advisory Committee, cited in Clarke, Goodwin, and Whiteley 2017, 113.
45 Clarke, Goodwin, and Whiteley 2017, 48.
ibid., 168.
47 Hobolt 2016, 1269.
49 Clarke, Goodwin, and Whiteley 2017, 102-3.
52 Hooghe and Marks, cited in Clarke, Goodwin, and Whiteley 2017, 64.
55 Dominiczak and Spence 2015.
56 Chu 2017, “Impact of immigration on native wages ‘infinitesimally small’ says author of study cited by leading Brexiteers.”
57 Chu 2017.
58 Graham 2014, “Nigel Farage: 'Massive oversupply' of foreign labour is forcing British wages down.”
59 Clarke, Goodwin, and Whiteley 2017, 48.
60 Surridge, McAndrew and Begum 2017, 19.
62 Goodwin and Milazzo 2015, cited in Clarke, Goodwin, and Whiteley 2017, 64.
63 Sparrow 2014, “Nigel Farage: parts of Britain are 'like a foreign land.'”
64 Clarke, Goodwin, and Whiteley 2017, 54.
Ecotourism and the Middle East

Andrew Ver Steeg
Abstract

Tourism provides a major source of industry in various nations around the world. Although historically, tourism has proven ecologically harmful to its host nations, a new approach, called ecotourism, takes a more conservationist and sustainable approach. This article discusses the agents, aims, pitfalls, and potential of ecotourism in the contemporary global sphere, particularly in the Middle East. It then examines the successes and challenges experienced by two case study nations: Jordan and Oman. Following an analysis of ecotourism in these places, the piece recommends a number of elements for consideration for nations aiming to boost the environmental considerations of their tourism sector, including but not limited to education and environmental regulations set by policy.
Table of Contents

Introduction 58
Ecotourism Overview & General Discussion 59
Actors 60
Aims 65
Pitfalls 65
Potential 68
The Twin Case Studies: Jordan & Oman 70
Conclusion & Recommendations 80
Notes 84
Introduction

According to the World Travel and Tourism Council, a tenth of global jobs and GDP are rooted in the tourism industry.\(^1\) Ecotourism specifically is one of the tourism industry’s fastest growing sectors and has seen growth rates as high as 10-30%.\(^2\) Though for some, tourism may conjure up pleasant thoughts about sipping margaritas in Cancun, millions of people interact with and indeed depend upon the sector for a livelihood. As a champion of these millions and their communities, as well as of the environment, education, and economies of millions more, ecotourism and its growing popularity stand ready to improve the future of the world at large.

Ecotourism is a recent concept, the origins of its contemporary usage beginning in the 1970s.\(^3\) In keeping with social movements at the time, the notion of ecotourism paid attention to how to alleviate negative impacts on the environment while advocating for the empowerment of the marginalized people involved, i.e. those populating the local communities wherein tourism took place. Though the meaning of ‘ecotourism’ has and continues to change, many attempts have been made to define it and a rough consensus on its core principles has emerged, evidenced by The International Ecotourism Society’s terse definition: “Ecotourism is responsible travel to natural areas that conserves the environment, sustains the well-being of the local people, and involves interpretation and education.”\(^4\)

This paper hopes to introduce the reader to the aims of the ecotourism industry, its different actors, and some of its pitfalls and potential. In doing so, it will orient the reader to ecotourism and its theoretical contours by employing examples from across the globe, which will provide a framework for understanding ecotourism generally. The scope will then narrow to two nations that have already done much to lead the field in the Middle East, Jordan and Oman. Along with the author’s personal experience with both countries, they are at the forefront of ecotourism in their
region, there is a great corpus of research in English documenting their progress, they are both stable politically, and the differences between the two make them an excellent pair side-by-side study. Though the framework established earlier will not map cleanly on to these two case studies, it will inform the discussion as it moves on next to the unique successes and challenges of developing the ecotourism sector in these countries.

The paper concludes with larger lessons and recommendations that have relevance for the Middle East as a whole. It is undeniable that the region is both dependent upon and ripe for tourism development, evidenced by the fact that it witnessed international tourist arrivals grow from 24.1 million in 2000 to 58.2 million in 2010.\textsuperscript{5} This growth coincides with a broader, more international trend, as demonstrated by the U.N. General Assembly’s declaration of 2017 as the “International Year of Sustainable Tourism for Development.” It is hence the author’s stance that, while it faces some challenges, ecotourism has the potential to sustainably and sensitively aid the development of Middle Eastern economies. Indeed, the Middle East is not only well-positioned to participate in the global trend towards more sustainable tourism development, but can lead the drive through ecotourism efforts.

**Ecotourism Overview & General Discussion**

Though it may seem pedantic, defining ecotourism and its dimensions is necessary to understand how to judge the successes or failures of any ecotourist operation. Unfortunately, ecotourism as a term encompasses many different interpretations, and it thus resists easy definition. Some have simply used it to describe any form of outdoor recreation and nature-based trips, although even in its earlier usage many academics stressed its educational components.\textsuperscript{6} Others prefer to emphasize sustainability and claim that ecotourism’s primary focus should be healthy
practices that diminish the economic and environmental impacts of normal tourism on surrounding community. Building off this last usage, others yet foreground the local community at the heart of ecotourism, requiring inclusion of local inhabitants in the profits and administration of the operation. Clearly, ecotourism is a convoluted term and unraveling its many meanings merits its own paper. Accordingly, this paper will continue to roughly abide by the aforementioned definition supplied by The International Ecotourism Society, chosen for its concision and wide recognition. This definition differentiates ecotourism from conventional tourism, which may be understood as tourism without any of the normative principles, objectives, or constraints that accompany ecotourism. The sole purpose of conventional tourism is economic profit, often expressed as high-volume mass tourism which can lead to the degradation of a place’s attractiveness. Conversely, ecotourism strives to achieve greater conservation, responsibility, sustainability, and education.

**Actors**

In this definition, then, an ideal ecotourism model incorporates an attentiveness to the needs of the local community, a combination of cultural and environmental attractions, education about said attractions, and a minimization of ecological impact. However, even these parameters leave room for variation and many different ecotourism operations across the globe take different forms. Some are state-owned and operated while others are administered by NGOs, and some are branches of transnational companies while others are local industries. Every version of ecotourism involves, albeit to varying degrees, four main actors. As laid out by Dr. Peter Bjork, “The following groups of actors have to co-operate: the tourists, the tourism companies (the tourism industry), the authorities, and the local people.” For clarity, this paper will retain the terms used by Dr. Bjork
with the sole alteration of naming the ‘tourism companies (tourism industry)’ the ‘private tourism industry’ to distinguish it from the ecotourism industry as a whole.

Local Communities

Local people tend to be the actors with the least agency, and thus demand prioritization in any sustainable ecotourism model. In fact, to refer back to the definition of ecotourism, the well-being of the local people is a requisite characteristic. Local communities differ from place to place, including nomadic peoples, small villages, or large townships. Some communities are more family-based and others more civic-based, and thus approaches to incorporating them in the decision-making process varies too, but it is important to meet them on their own terms rather than discounting their modes of life as unfit for autonomy. There are a number of ways in which this discounting can occur, sometimes even by infringing on the prerogatives for self-determination and resource access provided them by the U.N. Declaration of Human Rights, which often targets the most vulnerable demographics of “women, indigenous peoples, people living in disputed areas and children.”11 In addition to having a say in the initial decision-making, local communities deserve inclusion in the running of ecotourism projects in recognition that their opinions and values are not static and do change with the progress of a project. One way to accomplish this is to include them through employment, intentional purchasing, and other monetary chains. Panama, where 56% of tourism income returns to local households, has adopted this principle and sought to locally supply “products, labor, tourism services, and increasingly ‘green services’ in energy and water efficiency and waste management.”12
Authorities

The authorities’ role is more fluid in that it can serve to reinforce the local communities’ wishes or sometimes to oppose them directly, often when taking the side of the private tourism industry. Enticed by the prospect of profit, development, and foreign investment some governments are too eager to develop a site for ecotourism before obtaining the consent of and considering the impacts on the local people. This can occur especially when the local people live more traditional lifestyles and thus remain on the fringes of conventional citizenship; for example, the high illiteracy rates among Maasai tribespeople led to the local government overseeing the conservation and management of their traditional pastoral lands on their behalf. This poses the danger of excluding another key actor, the local community, in the ecotourism process. Without a voice in the process, these parties lose access to the financial benefits of tourism development while their proximity entails enduring many of the effects, whether they be environmental degradation, elevated cost of living, or cultural intrusion.

However, authorities also provide the institutional support and assets in backing local communities’ interests. Whether at a regional or national level, legislative, executive, and judicial powers all can play a critical role in advocating for local communities. The national government of Bhutan has sought a “high value, low volume” model of ecotourism that has minimized the negative impacts of tourism on the local populace, conserving their environment, culture, and livelihoods. Furthermore, in Cambodia, the endorsement of community-based organizations by the national government, particularly the Ministry of the Environment and the Ministry of Tourism, has empowered those organizations to make their own decisions in determining the role of outside operators in local ecotourism projects. The role of governmental authorities is vital to the success of sustainable ecotourism and must be used wisely to further amplify the benefits.
accrued to local people, rather than losing their input in an exploitative quest to develop the economy.

__Tourists__

One often overlooked group is the tourists themselves. Though it is tempting to categorize this actor as reactive—a consistent demand-side force that responds to changes in the other parties—tourists themselves can influence what ecotourism looks like. Bjork has astutely considered the matter:

>A central question still unresolved is how much are tourists willing to pay for trips that are more sustainable than mass tourism? Although the environment is in focus, we have to keep in mind that even this kind of tourism must be based on profitmaking activities, and tourists must have the opportunity to enjoy their stay (D’Ayala, 1995; Giannechini, 1993). Another important unresolved question is how much comfort do ecotourists need? Are they willing to pay a higher price for a lower standard, i.e. usually a more authentic reality?

No matter how well-designed an ecotourism system is, it only succeeds if it attracts tourists. While some, more environmentally-conscious tourists are willing to ‘vote with their dollar’ and pay higher prices for a more sustainable experience, far greater potential lies in attracting wider tourist audiences. Education about the cultural and ecological significance of a site, a tenet of good ecotourism, can also lose its impact if ‘preaching to the choir,’ or just reaching already environmentally-conscious tourists, who currently constitute the majority of ecotourists.16 With these demographics in mind, we can understand how ecotourists play a part in the process today and how to envision ways in which to broaden their role in the future, by incentivizing more diverse participants.
Private Tourism Industry

As for the private tourism industry itself, it provides the foundation for the interaction of all the other actors. At a minimum, the industry must distribute benefits to the local community, abide by governmental guidelines, and attract paying tourists, coordinating between various demands and expectations. The ecotourism businesses’ profits underwrite this confluence of actors, and revenues must be sufficient to both sustain the businesses themselves and to provide benefits to the other parties. To strike this balance, when the ecotourism operations are run by private-sector, international, and corporate tourism companies, it is crucial that these companies avoid extracting the tourism money spent in local economies by hiring much of their staff from outlying communities. For instance, in the Peruvian Tambopata Candamo Reserved Zone, the employment of the indigenous tribe leaves them some 60% of the profits.\textsuperscript{17} Some prefer to directly distribute a portion of the profits back to the community as in the 19% of all revenues in the Maasai Mara Reserve in Kenya that go straight to the community there.\textsuperscript{18}

Ecotourism businesses that are local, smaller-scale, or community-operated, will often bring in NGOs for additional funding and support. NGOs can assist in guide training, research, conservation area management, investment, and community planning and stakeholder meetings, all small burdens that can cripple a nascent ecotourism operation before it gets off the ground if not addressed.\textsuperscript{19} However, the presence of several NGOs working towards the same end can result in a “duplication of roles” leading to competition or rivalry within the private tourism industry which disrupts the collaboration with other actors.\textsuperscript{20} One study even found, using meta-analysis of over 200 cases, that “the proportion of successful projects is greater for NGOs working individually than for those collaborating
with other NGOs.” At least when partnering individually then, NGOs often complement or even fill the role of the private tourism industry in managing operations, earning profits, strengthening community, and promoting conservation, especially in developing countries where governmental oversight tends to be weaker.

Aims

Though each actor seeks different aims, it is easy to note shared objectives that have emerged in the prior discussion. Among these are emphases on economic development, community development and involvement, cultural sensitivity and preservation, sustainability, environmental stewardship, and education. These constitute the essence of ecotourism’s ultimate telos: reduce the negative impacts of tourism and increase the positive ones. Though simple in theory, it requires a substantial amount of thought and foresight to accomplish, as well as a case-specific method of harmonizing competing goals. For example, stakeholders focused on conservation might advocate for no-use resource preservation as the priority while those concerned with commercial development might be willing to go further in exploiting natural resource use in order to bring economic prosperity to the local communities. No standard equilibrium exists, but must be negotiated in every given instance.

Pitfalls

Unfortunately, despite these lofty intentions, many attempts at ecotourism fall short and can suffer from a host of common sociocultural, environmental, and economic pitfalls.
Sociocultural Risks

One damning sociocultural fault strikes at the core of ecotourism, as it questions the authenticity of the ecotourism movement. While ecotourism’s public intentions are desirable and even noble, the fact remains that by encouraging and enabling visitations from afar, the ecotourism industry can inadvertently exoticize the sites, the local communities, and the cultures inhabiting them. While this process elide those environmental aspects that have ecological but not commercial value, like the conservation of species unattractive to tourists, it also obfuscates the voices of local communities who may not share the same marketable goals. Much of this tendency originates in the prevailing neoliberal logic of the world which prioritizes profit and the ability to attract tourists and their money.22

Beyond commercialization, ecotourism at its worst evokes ‘poverty tourism,’ that moniker given to projects that found their appeal on the exposure to abject conditions or alien lifestyles of other people. One Kenyan woman living in a ghetto noted “that it is morally unfair that tourists keep on coming to the place she calls home.”23 In the same fashion, those visiting more traditional, pastoral communities to marvel at their ‘backward’ practices or ‘inferior’ quality of life may bring economic wealth to the region but at the cost of the humiliation, alienation, and degradation of the local community. Even at the Huaorani Ecolodge in Ecuador, an operation that seems to be doing many things right, ecotourists pose with and photograph the tribespeople stricken with polio or physical disfigurements.24 In addition to this zoo-like exoticization, the Huaorani have complained that the ecotourists seldom leave tips for photos or pay enough for the traditional crafts they painstakingly make.25 A proper application of ecotourism would thus be one that monitors and limits this kind of belittling visitation while also ensuring fair prices or the expectation of reasonable compensation.
Environmental Dangers

Though the name ‘eco-tourism’ and its professed attentiveness to conservation intuitively absolve the industry of environmental harm, no form of tourism is without its environmental impacts. Habitat disruption, litter, trail deterioration where applicable, and light, noise, and physical pollution all contribute to environmental degradation, even when confined to designated ecotourism areas that are often just as vulnerable as the more strictly-protected sections. In part, ecotourism’s allure promises the access to pristine, ‘hidden’ environments, but developing these untouched refuges exploits them anyway, undermining the very thing many ecotourists seek. Furthermore, much of ecotourism’s potential takes place in developing countries, and the focus of the government on more pressing issues like providing adequate nutrition, security, healthcare, education, and employment precludes them from addressing the environmental impacts brought by ecotourism.

Economic Pitfalls

The economy can foil ecotourism’s aims in other manners too. Tourism is a rather variable and vulnerable industry that depends on tourists’ willingness to participate, which fluctuates based on everything from seasonality to the global economy to international politics. Investment in the sector does not allow stakeholders to recoup much if operations go under, and this imperils the livelihoods of the local communities and the ecotourism businesses. Even when ecotourism is thriving its success can negatively affect local communities. Through the importation of foreign goods, the involvement of large international tourism companies, and other factors, “The World Bank estimated in 1988 that 55 percent of gross revenues for all tourism in the developing world
eventually return to industrialized countries.” This is a distressingly high percent of profits dissipating from where it is spent and where it is supposed to bring prosperity. The presence of foreign investment, the concomitant inflation, and the influx of migrant workers drawn to the promise of employment in a lucrative industry all further detract from the economic benefits that would otherwise land squarely in the local communities. These drawbacks go to show how crucial it is to require participation from all actors, particularly local ones, to attain ecotourism’s desired sustainability and other positive attributes.

Potential

While these issues complicate the application of ecotourism, they do not preclude it. Ecotourism still brings several benefits to the table and, if executed properly, can meet its grand aspirations. In the same three realms, sociocultural, environmental, and economic, examples of successful ecotourism operations abound.

Sociocultural Promise

In Nepal, women often face lower literacy rates, access to resources, and decision-making powers than men, but the Langtang Ecotourism Project has sought to even this imbalance by bringing Nepalese women into the planning and management of the operation while providing them with opportunities to supplement their income via cultural performances and the sale of crafts. Although the Australian government originally administered Mutawinjiti National Park, the local aboriginal populations protested until they were included in the process and they now co-manage the park in conjunction with the National Parks and Wildlife Service, all while running every tour in the area through a company that provides their guides with professional training to
build on their traditional knowledge of the area. Similarly, Fiji Rivers, an eponymous ecotourism company, directs a large portion of its funds to education and training the local population, capitalizing on profits and tourist philanthropy to offer classes on conservation, swift water rescue, nature guide interpretation, first-aid and CPR, and other marketable skills.

**Environmental Capacity**

Environmentally, achievements can be as simple as sustainable construction, like with the Crosswaters Ecolodge in China whose construction materials are almost entirely recycled, local, and renewable materials like bamboo. The designers of the lodges used a mixture of GIS mapping, feng shui principles, and input from academic and industrial ecologists to minimize its impact on the habitat. More ambitious achievements include that of Frégate Island Private in the Seychelles where plantation agriculture had depleted 90% of native flora and fauna: the eco-resort has completely rid the island of pesticides and rats to protect its endemic bird species, established a nursery reviving many near-extinct native plants, and funded a nursery that witnessed a 1300% increase in the giant tortoise population in just two decades. Communities in Cambodia were able to conserve mangrove forests despite the high start-up costs and other impediments to initiating REDD+ programs by instead harnessing their collective might in Community Protected Areas to set up ecotourism projects that incentivized preservation while compensating the families dependent upon tree-cutting.

**Economic Potential**

This compensation is an integral part of the economic triumph of ecotourism. The Hidden Valley Inn Reserve in Belize hires the majority of its unskilled labor from local villages, buys fresh
produce from outlying markets, and even outside of its tours runs a coffee cultivating and roasting business that supplies additional income for the company and the local economy. The Tambopata region of the Peruvian Amazon, which, as noted earlier, leaves locals with the lion’s share of ecotourism profits, also poses a unique advantage. One study considered competing land uses in the area, namely ecotourism, logging, cattle-grazing, and gold mining industries, and weighed the various potential economic benefits of all. They found that, without exception, ecotourism provided the greatest net economic benefit of any alternative, without even taking into account its social benefits or positive externalities.36

The Twin Case Studies: Jordan & Oman

Ecotourism is clearly an expansive term, one that encompasses a diverse group of actors, various countries and cultures, and a whole host of success and failures. Keeping in mind the theoretical understanding and insight thus far gained in this paper, undergirded as it was by numerous, real-world examples, the discussion next focuses on two Middle Eastern countries, Jordan and Oman. With the previous section situating the reader in the ecotourism field, turning the attention to Jordan and Oman will shine more light on the specific challenges and triumphs of both. From this analysis will come lessons about both that bear relevance for the ecotourism endeavors of other countries, notably those also in the Middle East.

Jordan

Jordan enjoys a variety of different biomes and habitats, ranging from the forested north to the wetlands in the east to the rocky desert in the south. This heterogeneity has not gone unnoticed, and the monarchy founded and funded the Royal Society for the Conservation of Nature (RSCN)
to harness and conserve Jordan’s great natural resources. Though established by the state, RSCN is an independent NGO tasked with a public mandate to protect Jordan’s environment, and it also spearheads many of the major ecotourism operations in the country. Other smaller businesses have filled the lacunae RSCN left, creating a patchwork of ecotourism actors that blurs the lines between the four categories outlined earlier. Nonetheless, these actors have earned Jordan the reputation of being one of the more innovative, environmentally-conscious, and committed ecotourism leaders in the region. Altogether, while Jordan faces some difficulties in its quest to build a sustainable tourist sector, it has realized many more accomplishments, and both categories mark it out as valuable case study from which to learn ecotourism.

Challenges

As a pioneer in the ecotourism game in the Middle East, Jordan started tapping its potential relatively early. Unfortunately, its eagerness in rapid change led to some hasty, haphazard decisions that have since negatively impacted its ecotourism prospects. Notably, in and around Petra—one of the seven wonders of the ancient world and a massive tourist destination for Jordan—this development has resulted in a few longstanding issues. The construction of more permanent trails, paved roads, and other accessibility-oriented alterations ate away at much of the habitat for the local juniper trees and other flora, and the streamlined corridors now intensify the effects of the seasonal flashfloods that visit the site. Ironically, the government constructed a brand-new Archaeology and Tourism Faculty Building on land that not only had been considered protected, but had retained the last vestiges of greenery in the area. These kinds of mistakes have lasting effects but are quite preventable and require simple foresight to avoid. Although it is too late to apply this lesson in the already established ecotourism sites in Jordan, it proves valuable
when approaching the creation of new projects across the country so long as thoroughness in planning is prized over haste in development.

Another area in which Jordan has room for improvement lies in education. Particularly, two separate realms of education could be improved upon, and they are in the education the general populace and the education and training of a skilled workforce. First is educating the public and Jordanians about the benefits of ecotourism to raise awareness, instill a conservational ethos, and encourage further participation in and support for the ecotourism industry. The author noted during time spent in the country volunteering with RSCN in 2017 that there seemed to be a disconnect between the work the NGO was performing and the interest of average citizens. Though proud of the environmental and cultural heritage of the nature reserves RSCN administered, many of the Jordanians the author spoke with were unclear about what exactly the organization did, why habitat stewardship was crucial, or what benefits ecotourism could bring. Others have remarked on this lack of awareness, saying that “…there is a low level of awareness, both within and outside the tourism industry, of the environmental impacts of tourism and of alternative approaches like eco-tourism.”

Thus, through school curricula, social media campaigns, informational materials, and events, both state and non-state entities in Jordan could further educate the populace on the role of ecotourism in their country.

Though strides are being made in this direction, Jordan still suffers from a dearth of professional, skilled, and well-trained tourist staff. This arises in part from hiring staff members for ecotourist projects from local communities, which, although it benefits the economic development and inclusion of the local population, can result in an unskilled workforce. This is not inevitable however, and it would be wise for ecotourism companies hiring in this fashion to invest in their staff by training them and even certifying them to be guides. RSCN funded several
Jordanians to acquire “nature guidance qualifications and accreditation from the Field Guide Association of Southern Africa in 2014. These individuals, however, were only four in number. Jordan stands to gain from increasing this number and adding to their ranks of professional staff, since the increasing popularity of the nation’s ecotourism scene will only merit more world-class guides.

Another challenge to ecotourism in Jordan is, paradoxically, the success of the industry itself. In Wadi Rum, a desert nature reserve in the south of the country, the local Bedouin tribes live in and run ecotourism camps. Using interviews and surveys, one researcher investigated the impacts of ecotourism not on the physical environment, but on the sociocultural milieu of the Bedouin communities. He found that the allure of easy profits to be made in operating tours and the draw to provide for one’s family led to “the creation of an incentive to not finish schooling and begin working in the industry as soon as is possible.” This phenomenon had taken such a hold that 100% of the study’s participants said they had dropped out of high school or had not progressed beyond high school. Furthermore, the capitalist structure of the reserve and the multitude of camps from which tourists can choose has created unhealthy competition between Bedouin camps and their families, leading to the deterioration of friendly ties, kinship, reciprocity, and hospitality. While these findings are confined to Wadi Rum, it is not difficult to think of parallels across the country, like in the Dana Biosphere Reserve where Bedouin communities are similarly employed.

**Successes**

Yet, portraying Jordan’s ecotourism scene as a network of difficulties and missed opportunities would be both unjust and inaccurate. There are numerous projects that have earned
recognition for their sustainability and positive impact, and the industry has been smart in building up its potential. In 2012, 18% of Jordan’s GDP came from tourism alone.47 More interestingly, while the largest portion of these tourists, 27.6%, came from foreign countries, the next largest portion, 24.5%, came from other Arab countries.48 This distribution of tourists is an asset for Jordan, as demonstrated during the plummet in all tourism revenues for Middle Eastern countries during the Arab Spring due to concerns about stability.49 The tourists from other Arab countries, especially the Gulf, prefer Jordan for its “extremely favourable price-performance ratio and its comparative liberality.”50 The diversity in tourist demographics allows Jordanian ecotourist operators to cater to different markets accordingly: should issues of political instability dissuade European and other Western tourists, Jordan can still rely on intra-regional visits, and should oil prices plummet and jeopardize the ability of these Gulf citizens to visit, then Jordan can turn to more global markets that conversely are more willing to pay for travel.

Another way in which Jordan can capitalize on this interplay of divergent elements is in stressing its attractive climate. Jordan shares with its neighbors that idyllic Mediterranean climate thought by some to constitute the world’s most popular tourist destination, with more than 120 million visitors each year since 2005.51 As a result, Jordan does a good job of depicting itself as a winter getaway for colder, northern climes. Simultaneously, its more northern location in regard to the Gulf promises cooler climes and even the presence of snow in some parts, a rarity for Saudis, Emiratis, and others. Though, as noted above, these Gulf Arab tourists appreciate the relative liberality of Jordan, its status as a Muslim-dominant, Arab nation make it the perfect balance between an exciting getaway and a familiar friend, reinforcing its appeal. Between the differently-inclined markets to its south and north, Jordan can diversify its reliance on any one tourist demographic and thus enjoy a more robust ecotourism industry.
More specifically, Jordan has found success in individual operations, like that of RSCN’s Feynan Ecolodge in the Dana Biosphere Reserve. This lodge is entirely solar-powered and is lit at night with candles made on-site, helping reduce light pollution in the process, while it further reduces waste via the composting of biodegradable trash and the burning of charcoal made from discarded olive pits for warmth.\textsuperscript{52} The entire staff of the lodge consists of people from the local villages, many of whom are women who otherwise seldom find work outside the home, and the camp drivers are all Bedouins who supplement their income with this part-time position.\textsuperscript{53} Furthermore, as is repeated in some of the camps in Wadi Rum and elsewhere in Jordan, RSCN offers chances for tourists to work alongside the Bedouins so they can do more than observe. This leads to the retention in Bedouin life and the appreciation of foreigners for activities like herding and the making of bread, candles, hair tents, and traditional eyeliner.\textsuperscript{54}

Additionally, while locals make and sell these crafts on site, RSCN has tapped into its nation-wide structure to move and sell these products throughout the country. Thus, whether made by women in the Dana Biosphere Reserve, the Ajloun Forest Reserve or anywhere in between, tourists can purchase these fruit rolls, olive-oil soaps, teas, calligraphy items, silver jewelry, organic herbs, ceramics, biscuits, jams, at any of RSCN’s locations, even its Wild Jordan branch in the capital city, Amman.\textsuperscript{55} In this way, RSCN is able to preserve the traditional methods and craftwork behind these products while opening up new markets and increasing profits for the communities creating them. While as a singular example the Feynan Ecolodge is exceptional in many regards, RSCN has replicated a similar model in its other reserves, demonstrated by its craft production network, and the organization’s larger efforts have only boosted ecotourism’s growth in Jordan.
Oman

Though both Jordan and Oman are Arab, predominantly-Muslim, and monarchical countries in the Middle East, they vary markedly. For one, Oman, unlike Jordan, is oil-rich, and the extraction of fossil fuels is the backbone of the economy. Nonetheless, tourism plays an increasingly important role, especially as a lauded alternative to diversify Oman’s economy and reduce its dependency on oil. Oman gravitates to ecotourism in particular as a means of sustainable development, and the country boasts a strong record for environmental stewardship, having established “more than 15 natural reserves and protected areas stretching over an approximate area of 30,000 square kilometres” while also signing and abiding by the Convention on Biological Diversity and the Convention on the International Trade of Endangered Species. Both in name and actuality then Oman has taken great strides in providing the institutional and legal framework necessary for ecotourism to flourish. The Omani Ministry of Tourism and Ministry of Environment and Climate Affairs, which together combine the roles of two ecotourism actors, the authorities and the private tourism industry, have supported much of this progress. Their collaboration and engagement with ecotourism is in recognition of the country’s great array of natural resources that include world-class diving locations, beaches, oases, mountains, and unique wildlife like the Arabian Oryx and several endangered species of sea turtles. While Oman sports this great environmental diversity, the literature has not kept pace, and few studies have thoroughly covered the state of ecotourism in the country. Even so, we can witness the great success Oman has found with sea turtle conservation, its potential for further ecotourism progress with oases, and some of the impediments to this grand venture.
Challenges

Unfortunately, Oman has encountered a few challenges to ecotourism development that has prevented the industry from living up to all these aspirations. Buerkert et al. discuss the case of the oasis village of As Sawjarah, which is somewhat similar to Misfat al-Abryeen in its continued production of traditional farming and livestock practices, though it lies in the mountain area of Jabal al Akhdar and thus utilizes a unique type of terrace-based agriculture. Noting the increasingly feeble returns on agricultural sales and residents’ expressed desire to abandon the oasis, Buerkert et al. proposed a radical transformation of the village into an ecotourism project wherein residents would be trained to work as guides, leading tours and engaging visitors in traditional activities like fruit-harvesting, grain-grinding, livestock-tending, bread-baking, and making of handicrafts like baskets or walking sticks. They called for this culturally and ecologically sensitive development as an alternative to the country’s otherwise large-scale, industrial, and non-educational tourist industry. Unfortunately, when the author here visited As Sawjarah in the summer of 2018, no major steps had been taken in this direction and, conversely, the exact large-scale, industrial style of tourism Buerkert et al. advised against had taken root just across the canyon. It did so in the form of Anantara Jabal al Akhdar Resort, a “fortress-style” compound sporting 115 rooms, tennis courts, a spa, biking, an Italian restaurant, and even an archery range. A far cry from the sustainable, culturally-attuned model of ecotourism, this property goes to show that the fight for better tourism in Oman is not yet won.

This occasionally anemic commitment to ecotourism has manifested itself in another lamentable fashion in Oman. The government established the Arabian Oryx Sanctuary in the 1980s, creating a refuge for “Nubian ibex, caracal, Arabian gazelle, sand gazelle, sand cat, the last wild breeding population of the Houbara bustard in Arabia and a great number of other animals,”
as well as sites of great geological and archaeological interest.\(^{64}\) The sanctuary even included an information center, as well as tour guides run by the local pastoral tribe, providing economic benefits and environmental education both.\(^{65}\) Alas, the government shrank the sanctuary to one tenth its original size in 2007 after the discovery of oil in the region, spawning a decline in Oryx numbers from 450 in 1996 to 65 in 2007.\(^{66}\) In this individual case, competing priorities led the Omani government to undermine the integrity of the ecotourism venture for the economy’s sake. Clearly, steadfast dedication to the guiding principles behind ecotourism, behind “responsible travel to natural areas that conserves the environment, sustains the well-being of the local people, and involves interpretation and education,” is missing from some of these projects.\(^{67}\)

One other stumbling block Oman faces in its pursuit of ecotourism is the composition of its labor force. Migrant labor issues are common in the Gulf, and Oman is no exception. In the tourism industry alone, “expatriates represented 88.40% (112,144 out of 126,857) of the workforce in this sector, which is 7.6 times higher than the local employment.”\(^{68}\) It is difficult to engage with, deliver benefits for, and include local communities when the overwhelming majority of the staff in an ecotourism operation come from other countries. In this case, migrants predominantly from South Asia come to fill unskilled labor roles and typically lack the interpretive training and lingual skills to be adequate ecotourism guides. Indeed, many ecotourism sites, especially those run without support from or coordination with the public sector, suffer from a paucity of qualified guides trained in general tourism issues, conservation principles, and context-specific sociocultural understanding.\(^{69}\) It is unlikely Oman will witness a sudden surge of Omanis desiring to work in the ecotourism sector unless action is taken to either decrease the flow of migrant labor or to incentivize this work for citizens. Efforts made towards the latter also help combat the attitudes of farmers and villagers like those in As Sawjarah, where incomes come mostly from non-agricultural
activities, familial remittances, and government subsidies, encouraging movement to urban centers in lieu of ecotourism development.\(^{70}\)

**Successes**

Despite these setbacks, Oman has done a great many things right in establishing its ecotourism sector. The country is home to substantial populations of multiple sea turtle species, including loggerheads, green turtles, hawksbills, and olive ridleys.\(^{71}\) By both establishing the Ras al Jinz Turtle Reserve around the coasts where the animals nest and building a resort there that sources tours to view the laying and hatching of eggs, the Omani government has promoted a popular and profitable activity; even in its early stages in 1997, the reserve grossed over $100,000 USD in revenue and visitors have continued to increase since, with an average annual growth rate of 21.53\% between 2010 and 2015.\(^{72}\) During the author’s stay at the reserve in the summer of 2018, tourists came from everywhere from China to Scotland to the U.S. to view the hatching of the eggs, although many too were Arabs from other Gulf states. Moreover, the government has incorporated education into the resort, supplying an information center as well as knowledgeable guides.\(^{73}\) The author’s own experience attests to the efficacy of these programs, observing the great detail and accessibility of the multimedia information in the resort’s museum-esque sea turtle center.\(^{74}\) The guides that conducted the tours to see the turtles laying eggs at dawn and dusk were incredibly proficient, alternating instructions on how to avoid disturbing the turtles with interesting facts about their reproduction habits, lifecycles, migration patterns, and more.

Even beyond this realized success, Oman offers much more potential for ecotourism growth. Mountain oases, hotspots of agriculture, biodiversity, culture, and beauty, are a promising avenue. One such oasis, Misfat al-Abryeen, is devoted to agricultural production on a small scale,
using the traditional Omani aflaj irrigation system. Although no ecotourism operation currently exists, many tourists, the author included, have visited the oasis to observe and enjoy this intersection of picturesque surroundings and living history. In fact, Zekri et al. argue that the positive externalities generated by the farms exceed $360,000 USD annually, and that if a simple, minimal entrance fee were charged, that farmers could achieve between 6-21% higher profits, even with the accompanying loss of some tourists unwilling to pay. In this case the local communities, i.e. farmers, are already doing what an ecotourism project would anyway by maintaining a cultural and ecological heritage, and all that remains is to capture the economic benefits of this activity. In tapping this ecotourism potential and others across the country, Oman also has the luxury of providing strong, state-backed investment. The oil wealth of the country will not last forever, but so long as it does, it can be directed to fund longer-term, more sustainable ventures precisely like those in the sector of ecotourism.

**Conclusion & Recommendations**

There is no doubt that Oman and Jordan both face their own set of ecotourism challenges and successes, and each is uniquely defined by its physical, economic, and cultural context. While issues remain with labor, luxury development, and the balance of competing priorities in Oman, the realized achievements and further prospects for environmental protection and diverse and sustainable economic growth are staggering. For Jordan, though progress could be made in educating the public, investing in guide-training and other ecotourism infrastructure, and cautioning the easy growth of capitalist ecotourism, the country’s marketing to a diversity of tourists, capitalization on its natural climate, comprehensive ecotourism model, and innovative commercialization of crafts are accomplishments worthy of both praise and replication. As nations
that have devoted considerable resources to developing an ecotourism industry, both can impart lessons to other countries in the Middle East.

Nascent ecotourism sectors have begun to arise in Tunisia, Lebanon, Israel, Egypt, the U.A.E., Saudi Arabia, Qatar, Iran, and Turkey. Few, however, can match the ecotourism efforts, environmental diversity, and political stability of Oman and Jordan, all contributing factors in their preeminence in the region. Oman’s tourism industry remained relatively unaffected after the events of the Arab Spring, with arrivals growing 8% in 2013, and although Jordan saw a decrease of 5% in the same year likely due to its proximity to Syria, it still outpaced Egypt, Lebanon, and the averages for the Middle East and North Africa regions. Meanwhile, critics have found fault with Iran, Turkey, and the UAE for over-exploitation and unsustainable development of their ecotourism industries. More specifically, Iran remains a small player in the global scene while some of Turkey’s forest are quite mismanaged.

So, what exact lessons can be learned from Jordan and Oman? Both enjoy a strict set of regulations that provide a legislative and institutional backbone to conservation efforts and ecotourism development. The Omani government ratified the Convention of Biological Diversity in 1994, while in 1995 Jordan passed The Law of Environment Protection no. 12 to fortify possible weakness in other legislation pertaining to environmental conduct. In the same vein, both countries have a string of nature reserves that sport unique tourist opportunities, whether its hiking in Jordan’s Ajloun Forest Reserve or watching loggerheads lay eggs in Oman’s Ras al Jinz Turtle Preserve. These reserves ensure the conservation of the physical constituents of ecotourism, allotting greater resources and attention to the involvement of local communities, the perfection of sustainable operations, and the maximization of healthy economic growth. While Jordan has been on occasion too eager to develop quickly, sometimes not giving sufficient thought to future side-
effects as mentioned earlier in the case of Petra, it appears to have kept a sustainable pace of
development overall. So has Oman, despite its relatively late entry into the world of ecotourism.\textsuperscript{80} This would be a good standard by which other Middle Eastern countries in an incipient stage of
ecotourism can abide so they do not develop too quickly and find themselves confronted by issues
that could have been avoided with prior planning or moderation.

Though unfortunate, it is likely that future political unrest of some sort or another will erupt
again in the Middle East and that it will inevitably impact other nations in the region, if only by
virtue of their proximity, like during the Arab Spring. Hence, it is wise for countries to seek, as
Jordan and Oman have done, to diversify their markets and attract tourists from within the region
or other areas of the globe. This way, if tourist arrivals from one collection of nations, like the
‘West,’ suddenly plummet, the industry is not crippled. Again using Jordan as a template,
incentivizing traditional craftwork and arts among local communities by selling their goods on a
national level without impacting the production processes would work well anywhere such
craftwork, a staple of Middle Eastern merchandise, is prevalent. Both Oman and Jordan also
emphasized the importance of education in their ecotourism projects. They did this in structured
ways like through Oman’s information centers at its sea turtle resort and elsewhere, but also more
experientially like through Jordan’s chances to work alongside Bedouins in performing their daily
tasks. Both could still stand to further professionalize and train their ecotourism guides and staff,
and this poses an exciting realm of unexplored opportunity for other Middle Eastern nations
looking to improve on their models as well.

More recommendations can be drawn from additional study of the examples provided here,
and it is the author’s hope that this paper inspires some to think critically about approaching new
ecotourism developments in the Middle East. Ideally, the great number of parameters by which to
define ecotourism, its actors, aims, pitfalls, and potential adequately prefaced the nuances of this analysis for a lay reader. Given these parameters, the difference of experience between examples, even between two good examples like Jordan and Oman, goes to show that there is no one right way to approach ecotourism. Indeed, it goes to show that the field remains heterogenous and ripe for innovation. Countless ways to improve upon conventional tourism in the Middle East exist. Hopefully, the above discussion has given some idea as to just what they may look like.

Though at times it seems like attempts at ecotourism face similar criticisms as those leveled at conventional tourism, they are few and far between, and the successes of said attempts, however problematized, still helps push the field in the right direction. Ecotourism presents a far more sustainable model of tourism development by paying mind to its community, environmental, economic, and social impacts. These are all imperative, as the U.N.’s advocates, in the larger quest for tourism to promote “1. Inclusive and sustainable economic growth 2. Social inclusiveness, employment and poverty reduction 3. Resource efficiency, environmental protection and the fight against climate change 4. Cultural values, diversity and heritage 5. Mutual understanding, peace and security.” These aims, linked to the U.N.’s 2030 Agenda for Sustainable Development, span beyond ecotourism and possess great importance for both the world at large and for future generations. They are, at their core, goals for humanity. Despite this grandeur, ecotourism, and particularly ecotourism in the Middle East, has the capacity to help realize these lofty aspirations.
Notes:


13 Chepkwony, Irene Chebet. AN ACTOR CENTRED APPROACH TOWARDS AN UNDERSTANDING OF ECOTOURISM – THE CASE OF MAASAI MARA. p. 146.


15 Mom, Sary. INTERACTIONS BETWEEN STATE AND NON-STATE ACTORS IN RESOURCE GOVERNANCE: A CASE OF COMMUNITY PROTECTED AREAS (CPAs) IN PEAM KRASAOP WILDLIFE SANCTUARY, KOH KONG, CAMBODIA. Chiang Mai University, Sept. 2016.


18 Chepkwony, Irene Chebet. AN ACTOR CENTRED APPROACH TOWARDS AN UNDERSTANDING OF ECOTOURISM – THE CASE OF MAASAI MARA. p. 146.


20 Chepkwony, Irene Chebet. AN ACTOR CENTRED APPROACH TOWARDS AN UNDERSTANDING OF ECOTOURISM – THE CASE OF MAASAI MARA. p. 146.


25 Ibid.


31 Ibid.


33 Ibid.

34 Ibid.


39 Ibid.


45 Ibid.

46 Ibid.


48 Ibid.


54 Ibid.
60 Ibid.
61 Ibid.
65 Ibid.
70 Ibid.
71 Ibid.
73 Ver Steeg, Andrew. Field Notes. 28 July. 2018.

[https://digitalcommons.dartmouth.edu/dujpew/vol1/iss2/3](https://digitalcommons.dartmouth.edu/dujpew/vol1/iss2/3)

Ibid.


Faith in Democracy:

Korean Churches as Engines of Pro-Democracy Protest

Will Matheson
Abstract

This paper delves into the characteristics and political actions of Korean Christian churches in the 1970s and 1980s to explore their influence on the societal transition to democracy. Much of the scholarship on this period has pointed to the various actors who took part in the pro-democracy movement. However, it is equally important to explore the ties between these organizations, facilitated by the church as a social institution. While scholarship has examined official joint-declarations and coalitions made by religious institutions, this paper seeks to explore how aspects of the Protestant and Catholic Church themselves created interpersonal networks among protesters and shaped their political action, not just as Christians but as members of the other major groups identified with the pro-democracy protests as well.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>91</td>
</tr>
<tr>
<td>Authoritarianism in South Korea</td>
<td>92</td>
</tr>
<tr>
<td>Christianity in South Korea</td>
<td>94</td>
</tr>
<tr>
<td>Social Capital and the Churches’ Organizational Impact</td>
<td>98</td>
</tr>
<tr>
<td>Church Relationships with Students and Youth</td>
<td>101</td>
</tr>
<tr>
<td>Church Relationships with Labor</td>
<td>103</td>
</tr>
<tr>
<td>Discussion</td>
<td>107</td>
</tr>
<tr>
<td>Notes</td>
<td>109</td>
</tr>
</tbody>
</table>
Introduction

The set of movements in South Korea from the late 1960s through the 1980s that called for establishment of a democratic government encompassed a coalition of societal groups, including religious groups. Both Protestant and Catholic Christians joined the movement in protest of the authoritarian regime. These churches briefly served as the face of the movement, publishing admonitions of the state of Korean society and occasionally organizing protests in order to criticize the regime.

However, the churches’ role extended beyond declarations. As social institutions, the churches had a broad influence on the movement and the various groups that constituted it. Through both formal and informal connections, the churches had influence over, and were influenced by, other activist groups. Some of these influences were intentional, as groups within the churches formed community engagement organizations that directly connected the churches with Korean society’s disadvantaged actors. Others were less obvious. The churches’ influence on people as members of denominations, as youth within their ranks, and as civically engaged members of society helped create the conditions for anti-governmental action by different societal groups.

The quality underlying the churches’ influence was social capital. As institutions that fostered unique social interactions, the churches influenced how a number of Koreans related to one another. Specifically, this organizational influence imbibed Korean citizens with certain community organizing skills, as well as networks that helped initially foster pro-democracy activist groups, protect those groups from government repression later on, and develop linkages between those groups.
Authoritarianism in South Korea

With Korea’s liberation came its division. Due to political vacuum, competition, and the influence of strong global powers (the US and USSR) partitioning Korea, the division between North and South became clear and extreme. These tensions spawned a brutal civil war (1950-53) that decimated virtually the entirety of the country. The violence used to establish control at this time framed leadership in the following decades.

In South Korea, the US desired one quality in a leader: vehement opposition to communism. As a result, it placed Syngman Rhee into power, partly due to the fact that Rhee had spent time in the US and established trust with Washington decisionmakers who understood his commitment to fighting communists. Rhee’s government was decidedly Christian: he was a church elder, his vice president and acting prime minister were church ministers, a quarter of the National Assembly was Christian, most of them church leaders, and an estimated 40% of all political positions were filled by Christians, despite the fact that Christians made up around 10% of the population. After coming to power, Rhee steadily chipped away at democratic institutions to ensure his continued reign. In 1959, Rhee forced the legislature to allow him to run despite having reached his two-term limit. Moreover, he executed his primary political opponent. After winning his 1960 election, students protested en masse in the 4.19 student uprising, which, after a severe government crackdown, sparked nationwide insurrection. With the military losing faith in his leadership and with pressure from the US, Rhee abdicated in 1960.

Rhee’s abdication led to chaos within South Korea. As a result, in 1961 General Park Chung Hee took power via a quiet military coup. In 1963, he was officially elected president, and then re-elected in 1967. While some initially hoped he would liberalize the country after Rhee’s rule, he instead continued Rhee’s example and seized power for himself. In 1969, he forced a
Constitutional amendment that allowed him to run for a third term. In 1972, he extended this reach, eliminating the democratic constitution in favor of the Yusin Constitution, effectively institutionalizing authoritarianism.² Park’s iron grip extended beyond politics to Korean society itself. His government aggressively regulated its people’s lives, implementing a curfew and regulating hair and clothing style.

Park enjoyed some support for his actions that the public would not have tolerated from Rhee due to Park’s economic policy. With state control of banks, the state effectively dictated the entire country’s economic policy. Park focused on industrializing the nation, exploiting cheap labor and using the state to help grow a small handful of companies at the expense of free market competition. This industrialization process led the government to recruit thousands of workers, including many young women, away from their agricultural jobs in the country to industrial jobs in the city, changing the makeup of the nation. Part of this economic control was state control of labor, politically demobilizing unions and allowing exploitative working conditions. Despite this harsh rule, many people looked up to Park for creating this ‘economic miracle’ that reversed the poverty and ruin the country suffered from the war.

Segments of society did, however, protest his rule. After he amended the Constitution to allow for him to run for a third term in 1969, students began to speak out. In the 1970s, the protest movement diversified, with intellectuals and professors, journalists, lawyers, workers and those associated with the labor movement, and Christian political activists joining the fight. In particular, Christians played a large role in the mid-1970s, before the movement truly diversified and society more broadly began to vocally protest Park’s rule. This diversity of protestors created a diversity of issues about which the movement was concerned, a diversity of demands upon the state, and a diversity of protest tactics used.
In 1979, the head of the KCIA, Kim Jae-Kyu, assassinated Park; in 1980, General Chun Doo Hwan took power and instituted martial law. In the same year Chun took power, students protesting his rule created the Kwangju Uprising (or the 5.18 Uprising), in which the army fired into a crowd of students and attacked the citizens of Kwangju. As Chun’s government became increasingly repressive, the democracy movement became more radical and more covert. In the late 1980s the middle class also became more vocal, and international attention due to the 1988 Olympics increased pressure on the regime to liberalize. As a result, a reform in 1987 allowed for direct presidential elections. Chun was succeeded by Roh Tae Woo, another General, after two progressive candidates split the progressive vote. However, the democracy movement became better established and institutionalized, and by 1993 Korea elected its first non-military leader.

**Christianity in South Korea**

Unlike its introduction to most other East Asian areas at the time, Christianity started on its own internally rather than being introduced by outsiders. Reform-minded Shilha Confucian Scholars encountered Western religious texts and explored them to better understand the world. In 1631, Chong Tu-won sent an envoy to China which brought back the first Catholic missionary in Korea, Matteo Ricci. Again, the interest in Christianity was mainly academic at the time. In the 1780s, Yi Song-hun traveled to China, was baptized, and then returned to Korea with religious texts. The beginnings of Christianity in Korea sprouted from these texts, with Koreans performing the practices they found in these texts, incorporating them into traditional Confucian practices, all without the presence of missionaries. As a result, a Korean indigenization process shaped how Christianity emerged in Korea.
Soon thereafter, a slow trickle of representatives of the church entered Korea. Chinese priest Zhou Wen-mo arrived in 1795. A half-century later, in 1844, the first two clergymen arrived in Korea. Over this period of time, the church experienced rapid growth, with an estimated 4,000 Korean Catholics in 1795 and 10,000 in 1801; moreover, by 1910, Korea had 69 churches and 71 priests, including 15 Koreans. At this time, the religion was primarily practiced by the yangban class.

In the mid 1800s, a religious thinking which incorporated aspects of Catholic thinking arose, called Tonghuk. In 1894, this thinking led to the Tonghuk Uprising, in which Korean peasants rebelled against the Choson dynasty. Due to its relation to Tonghuk, its emphasis of loyalty to God over loyalty to king, and its tenets that challenged traditional Confucian customs such as ancestor worship, Christianity the Choson dynasty began persecuting Christians.

Protestantism arrived later than Catholicism, with John Ross visiting the Korean border from China in 1874 and 1876, then translating the New Testament into Korean in 1882. In the same year, the Treaty of Amity and Commerce with the US opened Korea so missionaries could enter as representatives of diplomatic legations or charitable organizations. At the end of the 19th century and the beginning of the 20th century, the Protestant Church took on a global focus of missionary work, which helped to spread Protestantism in Korea. Missionaries entered Korea, and as part of their work, they began the tradition of establishing primary and secondary schools, universities, religious schools, and hospitals. By 1905, the Protestant Church had baptized an estimated 9,000 Koreans. Part of this growth was due to political cover; the Choson dynasty supported limited Protestant church action as a result of Horace Allen saving the life of Queen Min’s nephew after the Kapsin Coup in 1884. Moreover, it appealed to the lower classes in Korea’s rigid social structure, offering them religion, education, and medicine. One important factor of its
growth was its material focus: due to the indigenization process combining it with shamanism and Confucian ideas of prayer, Christianity in Korea had a strong this-worldly focus. This characteristic of the church exists to this day, with the church owning millions of dollars’ worth of assets and many still praying for material prosperity.

Japan’s occupation of Korea posed a new challenge to Korean Christianity. In 1910, the Japanese empire began persecuting Christians in Korea, signified by the Conspiracy Case of 105. Specifically, the occupation Japanese government fabricated allegations that Korean citizens, many of whom were Christians, had been plotting assassinations attempts, and sentenced 105 Koreans in an unjust trial. This event stoked anti-Japanese sentiment in the Christian community in Korea, and Christians played a disproportionately large role in the March 1, 1919 independence movement, both in terms of leadership and proportion of the population arrested. In the war mobilization years of occupation, the Japanese government mandated that all Koreans perform Shinto shrine worship to the Japanese Emperor, sparking a fierce debate among denominations over whether or not to refuse and risk persecution. Moreover, the ‘colonial regime created a state-controlled church, or Kyodan,’ further stoking anti-Japanese sentiment. In short, the Colonial period introduced the Christian church to the role of combatting an overbearing government.

The Protestant church saw a second explosion of growth after the end of World War II. In South Korea, the new US-controlled government reversed the suppression of religion and created a climate for Western religions to flourish. At the same time, in the north, the Soviet-influenced government increased religious persecution, spurring an exodus of Christians to the south. As a result of the war, many refugees joined the well-organized church, which provided material and social support. In fact, due to its organizational strength and relation to the West, the church became a primary network for distribution of post-war relief aid. This growth culminated in
‘between seventy and one hundred denominations and other groupings within the Korean Protestant church in the 1970s.’

The Protestant church in this period was strongly tied to the authoritarian governments in charge. However, a number of liberal Protestant denominations, as well as many Catholics, took part in anti-government protests for democracy. Specifically, they entered the movement in the 1970s, with involvement peaking in the mid 1970s and continuing to a lesser degree into the 1980s, as measured by the number of protest events by Christians over that time period. Many believe that General Park Chung Hee’s successful attempt to amend the constitution in 1968 to allow for him to serve for a third term sparked the beginning of this growth of church involvement.

In the political arena after liberation, a disproportionately large number of Korean leaders were Christian. This dynamic helps explain why the Protestant church was so close to the authoritarian government. After the war, the US government supported the church as a social organization because it had become a critical conduit for funneling aid to rebuild the nation. As a result, those with economic and social capital assumed both church and political leadership. Thus, many of the leaders of the church were invested in the government starting under the rule of Syngman Rhee. This also created a dynamic where leaders of denominations competed with one another, which occasionally engendered factionalism in denominations. This dynamic explains why ‘only 542 pastors and 831 churches out of 11,582 clergy and 17,793 churches in the Protestant community were involved in this activism against the regime.’ In the context of the pro-democracy movement, then, the Protestant church as a whole took a passive stance. However, a subsection of the church split with the majority and played a vital role in the movement.
Social Capital and the Churches’ Organizational Impact

Today, relational trust in Korea is quite weak. In a measure of generalized trust based on survey responses to the question ‘Most people can be trusted,’ only 26.5% of the population exhibited positive responses. Of nine sample countries, this rate was second-lowest; for comparison, China’s generalized trust rate was 60.3%. This demonstrates an important aspect about Korean society: most people regard others with suspicion, which both demotivates creating coalitions and broad networks and solidifies closed networks. This lack of generalized trust was even greater during the post-war era and under Park’s leadership: ‘Traditional rural communities were dismantled during the rapid industrialization process, yet the drastic social change was not accompanied by a corresponding development of organizational principles suitable for urban and modern life.’ Traditionally, many groups such as local villages outside of the capital provided such organizational platforms; however, with industrialization, they could no longer play this role.

The church filled this vacuum of organizational influence and trust-building institutions. In Korea, both Catholicism and Protestantism had a marked impact on social capital. First, Catholicism emphasizes charity and philanthropy, thus creating involvement in local communities. Protestantism traditionally does as well, though due to its material focus in Korea many denominations do not emphasize it to as great of an extent (nevertheless, a subset of denominations does emphasize it heavily). Second, both religions have high levels of institutionalization. In addition to the incredibly high number of denominations, churches, and members in Korea in the 1950s-1970s, the churches possessed vast organizational networks. The Protestant church controlled around 75% of all colleges and universities, a quarter of high schools, around 100 hospitals, and radio broadcasting, newspapers, and television stations; the Catholic church had a similarly extensive reach. Finally, the Catholic church had significant autonomy from the state,
providing the organizational platforms for activism (the Protestant church did not enjoy similar autonomy due to its strong ties to the state).¹³

General sociological reasoning supports this conclusion that religious institutions influence social capital. A meta-analysis of studies of religions in America found that religions create organizational environments that teach and put into practice community organizing and leadership skills. In addition, as social institutions they foster trusting networks both among youth, and between youth and other members of society. Finally, they provide religious organizations that focus on specific causes and/or issue groups, creating networks between church members (especially youth) and social causes.¹⁴ In Korea’s case, these organizations were directly linked to the other major groups of protesters: students and labor. Significantly, the historical factionalism within the Protestant church created such a large ‘constellation of civil society organizations’ that the government struggled to exert total control over pro-democracy activity.¹⁵

This form of autonomy played a particularly important role in shaping pro-democracy demonstrations. Due to heavy state repression, many other groups (students, labor, etc.) could not access institutions free from the repressive state apparatus. While the Catholic church possessed the most autonomy, the Protestant church’s organizational structure also lent it to pro-democracy political activism. Historically, church leadership divided the country by region and assigned missions to each region, creating divisions between different mission societies as they developed identities that in part reflected their local areas as part of the broader indigenization of Protestantism in Korea. They were influenced by the broader historical factions such as north versus south and political center versus the periphery. The Shinto shrine worship controversy further stoked factionalism. The divisions were accentuated after the war because different groups – stoked by tensions from historical regional divisions, power changes among church leadership,
and the Shinto shrine worship controversy – established their own separate denominations, including divisions among educational institutions. As a result, despite the Protestant church’s closeness to the government writ large, smaller units of the church had their own forms of autonomy from the larger church structure, and thus the state.16

Other historical events also helped create the dynamic in Korea where the church was unique in its role in facilitating social capital necessary for the pro-democracy movement during the crackdown in the 1970s. Despite the lack of generalized trust within Korea, church denominations created social spaces to develop trust. Moreover, after the war, most international relief was channeled through the church, imbuing it with legitimacy in local communities and creating networks of people.

The church was a uniquely important political organization. During the period of intense social and political regulation, the church not only provided the space, but also the training, for community action: ‘Koreans joined the Protestant church and practiced “Korean politics” … [the] church organization attracted politically active Koreans more than other forms did because it encouraged the participation of lay members in its pastor-elder sessions, presbyteries, and general assemblies.’17 Because the church actively involved the working class, it thus disseminated community organizing skills and created networks between members of the working class and members of other sectors of society.

One illustration of the church’s ability to foster trust was the Rhee regime itself. Rhee’s government was solidly Christian. The US was only concerned with getting political leadership that opposed communism, and not specifically getting Christian leaders in charge. The fact that they favored Christians was merely corollary, as many Korean Christians’ were vehemently opposed to communism.18 Moreover, the US government simply did not have the capacity to instill
Christian leadership to the extent that it existed within the Rhee government. Rather, Rhee’s regime favored Christians due to its leadership’s foundation in the church. Because the authoritarian rulers could trust the people they knew from the church, they were able to put them in power and create an authoritarian regime that would not splinter internally. Again, the church as an institution created social networks and trust which spilled over to political organization.

**Church Relationships with Students and Youth**

Students made up practically the entirety of the pro-democracy movement in its beginnings. The church’s relationship with students and youth influenced how the democracy movement grew, including how the church came to take up the cause in the mid-1970s. As mentioned earlier, the church provided unique organizational contexts that educated youth, trained them in skills like leadership, public speaking, and community organizing, and created social networks imbued with trust, thus laying important groundwork for the movement to later arise. Moreover, during the period of demonstration the church provided autonomy necessary to continue anti-government activity during a period of severe repression.

Beyond creating the organizational conditions within church activities, the church also created broader social conditions that helped give rise to youth social organization. It can certainly be said that without the great contributions of the Protestant and Catholic churches in Korea, higher learning institutions in the country would not have been as numerous or as autonomous in the 1960s and 1970s. These ‘schools...and their alumni associations are civil society institutions…institutionalizing loyalties and obligations that are far beyond the control of the state.’

Many of the church’s programs helped foster such social capital, preceding and during the movement. For example, the Ecumenical Youth Council (EYC) demonstrates how church-
sponsored organizations fostered inter-religious coalitions among youth groups. The EYC is a
great example of how the church's status made it subversive due to its autonomy, existing outside
of traditional spaces of protest and legal status. Its creation was distinctly political; its mission was
to ‘draw upon the potential power of students and youths while moving their activities away from
the campuses to avoid restrictions set by ED number 4.’ Other examples of church organizations
developing ties among youth and between youth and the church are the Korean Student Christian
Federation, the YMCA, the Seoul District Youth Association of the Presbyterian Church, the
Christian Church Youth Association, and the Intervarsity Christian Fellowship. Additionally, the
church fostered social capital among young women, specifically with organizations such as the
YWCA, Korean Coalition of Christian Women, and Korean Church Women United. This
development helped broaden relational ties among politically active church members and increase
political participation.

The youth social organization the churches fostered in turn spurred their direct involvement
in the movement because they defended the students from government crackdown. For example,
in 1975 the government demanded that Dr. Park Tae Sun step down from his post as the head of
Yonsei University. This action caused church members to join in and support protests by students,
showing the link between the church and student groups in society. Indeed, Dr. Park expressed his
gratitude toward the United Board for Christian Higher Education in Korea in New York ‘for all
the spiritual and material support they have given us.’
Church Relationships with Labor

Many of the church’s labor-associated groups expressed political purposes. For example, the Urban Industrial Mission (UIM), a unit of the Protestant National Christian Council, was expressly ‘aimed at protecting workers from abuse.’ These alliances helped the church funnel resources - particularly human capital - to the more political groups in the democracy movement. As with its relationship with student groups, the churches’ relationship with these organizations in turn caused them to become more politically involved in the protest movement due to the relational networks they created: ‘Urban Industrial Mission (UIM) ministers and church leaders first indirectly learned of the situation from the laborers themselves who attended their churches. They then obtained manual labor jobs in order to work alongside laborers and help them organize.’

This example shows how the networks formed within the church led to a broadening of the pro-democracy movement in the form of information dissemination and the church and its members getting more involved. Other examples of church organizations developing ties between labor and the church are the Seoul Metropolitan Community Organization, the ‘Street Learning Place,’ the Social Development Corps of Christian Students, and the Young Catholic Workers.

The Seoul Metropolitan Community Organization had close ties to the church and protested crackdowns on it, again demonstrating the societal web between the church and various social groups. This organization, designed to involve church members with lower class issues, fostered the human capital necessary for later pro-democracy action because members ‘helped to organize the community in order to protect the rights of the poor.’ This organization fostered ties with student groups as well because many Christian students joined the group. The Seoul Metropolitan Community Organization exemplifies the ways in which the Church’s organizational
structure fostered inter-group solidarity and the human capital necessary for pro-democracy protest.25

Many of these associations with labor also created intersections with youth activists. Night schools exemplify this inter-group alliance building. For example, Hyunje Church ran a night school, called ‘Street Learning Place.’ With night schools, the church created organizational connections not only among students, but also laborers for whom night school was the only option for group organization, discussion, deliberation, and interaction. Moreover, this focus on education meant the church developed the labor movement’s human capital.

The Christian Academy is another example of how the church created organizational networks and invested in human capital. It was initially created to assist with reconstruction after the Korean War (with a specific focus on establishing educational and social programs), again demonstrating how the church developed social capital among Korean citizens who would later engage in the democracy movement. Moreover, as participation in the democracy movement began to include more labor activists, it focused on assisting citizens negatively affected by industrialization. Thus, the organization played the role of a support network during the protest movement, assisting other groups in their work through both boosting morale and organizational support. The director of the organization, Pak Kyongso, explained that the group ‘trained laborers, farmers, youth, women, and ministers in the church… [to] raise awareness about democracy for all groups in society.’26 This explicit political focus on helping the disadvantaged groups calling for democracy shows how the church served as a training ground for grassroots organizing, instilling in and between various social groups not only skills but also contacts and networks.
Societal Subversion: Public Awareness, News Media, and Peaceful Protest

Beyond labor and youth, the church influenced grassroots organization via its role as a distinguished social institution. In the fight against the government, the church played a vital role in countering its propaganda war. Due to both churches’ strong association with anti-communism, the government lost its best weapon to use against its enemies: labeling them communists and arresting them on the grounds of national security. In addition to imbuing the movements with which it was associated with credibility, the church used this reputation to criticize the government and perform political activism in such a way that would gradually turn broader segments of society against the regime. Many of the subversive tactics the churches employed parallel those taken they took during the colonial period, when they provided a platform for nationalists to organize and criticize the colonial government.

The churches, as institutions, constituted vast informational networks that could be used to sway public opinion. As groups such as labor associated with and exerted influence on the policies of the churches, they could better politically mobilize. For example, the Christian Broadcasting system helped spread stories about repressive labor policies in order to build support for, and demystify the fear of, the labor movement. Moreover, such informational networks established a human rights discourse that framed how people thought about the movement and expressed themselves against the regime.

The churches’ social capital formation, creating social relations and trust, in turn created informal informational networks that played this same role as the more established information distribution networks. The Human Rights Mission ‘quickly became the central node in the Christian network that linked together a variety of other groups, thereby expanding the organizational infrastructure of Christian activism.’ Moreover, it was due to ‘the Thursday Prayer
Meeting, the Human Rights Mission, and other Christian organizations, such as the Association of Family Members of Prisoners of Conscience, that Christians were able to publicize the stories of political prisoners and state torture to the larger public,’ demystifying activists and eliminating fears propagated by the regime.29

Another tactic used by the church to change public opinion and give credibility to labor and youth activists was allowing the arrests of its members. Triggered by non-violent modes of protest by leaders in both churches, these arrests ‘led to solidarity among many protesting groups (who all coalesced around demanding the release of the church leaders) and broadening the base of the protest movement by winning the favor of many indifferent members of society who were horrified to see the arrests of these peaceful church leaders.’30 Equally important, these protests helped to further develop social capital because, as exemplified in the case of the arrest of Bishop Daniel Chi, ‘all Catholics automatically became involved… we got much educating done at the almost biweekly prayer meetings. In church is where most Koreans learned… about their government.’31 These prayer meetings involved testimonies from political prisoners and the distribution of the Human Rights Committee’s Human Rights Newsletter, helping educate members about various groups involved in the movement. With each arrest, the church provided a base of organizational networks to coordinate a public pressure campaign to protest the arrests. Moreover, due to the degree of inter-denominational autonomy within the church, the government struggled to prevent local communities which had organized in their churches from then mobilizing after these arrests in the form of prayer meetings and taking part in the vast organizational networks connecting the church to labor and youth.

An important aspect of this subversion is the physical capital the churches provided. The government could not attack worshipers in the same way it attacked student or working-class
protesters in the streets. Under the guise of, or in tandem with, religious worship, church members could attack the government despite its severe repression. Due to their being powerful institutions, the churches provided physical space for other activist groups to meet and organize, as well as develop more substantive alliances with other anti-government groups. Moreover, many churches and cathedrals, most notably the Myǒngdong Cathedral, served as physical sanctuaries for protestors fleeing police violence. Beyond direct contact with students and workers, then, the churches helped foster anti-government political action due to their institutional structures and resulting influence on social relations and information.

**Discussion**

The Christian churches, both Protestant and Catholic, played pivotal, changing roles throughout the movement. In the mid-1970s, the churches took up the pro-democracy fight when the state began to effectively repress the student movement; by the late 1970s, the churches then stepped back from this role as the visible face of the movement as it diversified.

However, in addition to this more visible role making declarations during the 1970s, the churches played a substantive role as a support network that helped lay the conditions for, as well as greatly assist other groups in, protest. The churches were never the most powerful anti-governmental force; after all, only a small portion of the Protestant church participated, and the Catholic church alone did not have the capacity to exert enough pressure on the regime. In terms of anti-government action, the churches alone were extremely limited as one large unified institution of Korean society, but their organizational structure helped to facilitate a much broader, if weaker, effect promoting the democracy movement at a time when the government was succeeding in stamping out the other manifestations of protest.
Despite this broader failure to engage in pro-democratic activism, the churches made a positive impact on the movement because they fostered social capital, creating networks among people, emphasizing civic engagement, and building trust. This dynamic helped shape who the protesters were and how they organized. Moreover, these networks spilled over to connect components of the movement and protect them during periods of harsh state repression, keeping the movement alive and laying the groundwork for its diversification. More specifically, they fostered conditions that helped give rise to the student protest preceding their involvement, then helped protect students as the government responded. They also helped develop other groups of activists, such as labor, and foster intergroup alliances to strengthen the democracy movement. Finally, they broadened the movement’s message and helped anti-government groups earn society’s trust, creating the conditions for broader public involvement.

Church-created social networks kept the protest movement alive when a crackdown could have otherwise extinguished public activism. In addition, they created ways to keep other societal groups involved such that they could continue their fight. These groups then went on to exert more pressure in the late 1970s and 1980s after surviving the introduction of the Yusin government thanks to the churches. This web of church organizations and informal relations thus played an important role in the democracy movement from a more informal, grassroots level.
Notes:
7 Kang, Wi Jo. 1997 ‘Christian opposition to President Park Chung Hee,’ pp. 100 in Sung Bae Park (ed.) Christ and Caesar in Modern Korea, State University of New York Press.
12 Ibid.
17 Ibid., 106.
25 Ibid., 111-114.
26 Chang, Paul Y. 2015 ‘The Emergence of Christian Activism,’ pp. 95 in Protest Dialectics, Stanford UP.
27 Ibid.
28 Ibid., 96.
29 Ibid.
30 Kang, Wi Jo. 1997 ‘Christian opposition to President Park Chung Hee,’ pp. 102 in Sung Bae Park (ed.) Christ and Caesar in Modern Korea, State University of New York Press.

31 Ibid.

Till Death Do Us Part:

Will Longstanding Rivalry Impede the Ethnic Coalition of ISIS and Al Qaeda?

Bianca Pergher
Abstract

According to Dr. Tricia Bacon’s and Dr. Elizabeth Grimm Arsenault’s, “Al Qaeda and the Islamic State's Break: Strategic Strife or Lackluster Leadership?,” the “strategic differences between Al Qaeda and ISIS were not sufficient to cause the split.” The strife that ensued between al Nusra and ISIS caused this complex alliance to rupture. In “ISIS and al-Qaeda—What Are They Thinking? Understanding the Adversary,” Dr. Bernard Haykel states that the jihadi militant movements are consequential to the Sunni revivalism experienced in the brutalized Arab politics of domestically repressed states from Algeria to Iraq, with groups seeking to empower Muslims against Islam’s enemies. Haykel later describes the relationship between ISIS and Al Qaeda as one of contentious competition for resources. Motivated by the work of Dr. Bacon and Dr. Haykel, the author uniquely addresses the remaining question of whether or not, and under what conditions, umbrella networks: ISIS and Al Qaeda, could attempt a merger.

To ascertain the pragmatism of a potential coalition, various alliance literature has been applied into three paradigms; with the first, assessing party coalitions and culture, the second addressing the alliance framework from a business perspective, and the third concentrating on partnerships and mergers. Moreover, the paper focuses on the aforementioned paradigms and various militant group case studies, with a lens that addresses the relational exchange between ISIS and Al Qaeda.
# Table of Contents

- Introduction 114
- Background: The Partnership’s Initial Demise 115
- Consulted Literature 119
- Research Design 126
- Militant Group Case Studies 126
- Expositions 131
- Analysis 133
- Conclusion 134
- Relevant Figures 136
- Notes 141
Introduction

At its zenith, ISIS’s forces controlled a territory the size of Britain, with approximately 10 million captives.¹ Under encampment, captives, most notably the Yazidis of Iraq and Syria, were murdered and enslaved. In 2017, the Iraqi Federal Police reclaimed the city of Mosul, a major jihadi encampment. In early February of 2019, Kurdish-led Syrian Democratic Forces (SDF) vowed to dismantle the remaining caliphate. With help from the U.S.-led international coalition, numerous warplanes and mortar barrages struck the stronghold (see figure 1). Now, with the last Syrian city of Baghouz under siege by the SDF, optimistic citizens may be naïve enough to assume that war on the global jihad is over. As CNN Journalist Ben Wedeman notes, “While some may be tempted to celebrate what is seen as ISIS’s demise, the fertile soil in which ISIS grew and spread is still there.”² With a large base of online followers, potential lone-wolf attackers, and militants working underground to conduct hit-and-run attacks and bombings in Iraq and central Syria, the fight is far from over. One question remains, will the caliphate’s setback be enough to reunite long-term, rivals: ISIS and Al Qaeda?

To better understand whether ISIS and Al Qaeda will likely re-unite, various alliance literature has been applied and assessed into three paradigms; with the first, assessing party coalitions and culture, the second addressing the alliance framework from a business perspective, and the third, concentrating on partnerships and mergers. According to the party coalitions and culture literature, ethnic groups are more likely to coalesce when they have shared cultural history, while the business alliance framework suggests that mergers are formed through delicate cultivation of synergy and pragmatic power-sharing. Moreover, the partnership and merger criteria claim that alliances are formed when groups converge on the basis of ideology. With this understanding in mind, the paper assesses the various paradigms and case studies of militant
groups with a lens that addresses the relational exchange between ISIS and Al Qaeda. Referring to the aforementioned literary framework, the findings indicate that diverging operational structures and a failure to coincide ideological goals, prevents ISIS and Al Qaeda from reuniting.

Background: The Partnership’s Initial Demise

In 1996, Osama bin Laden announced the first fatwa against the U.S. A synthesis conducted by the RAND Corporation has highlighted that a recent “switch from hard to soft targets and from mass casualty to smaller, but more frequent attacks; an increased focus on the economic impact of attacks; [and] greater efficiency in the utilization of loose networks,” has characterized Al Qaeda’s strategy. Subsequent to the 9/11 attacks, Al Qaeda lost its secure base in Afghanistan underpaid protection by the Taliban. Its senior and mid-level commanders were targeted due to increased security, creating a catch 22, whereby future planned ‘spectaculars’ would become unparalleled. As explained by Rabasa et al. in “Beyond Al-Qaeda: The Global Jihadist Movement,” Al Qaeda depends heavily on the subcontracting of militant groups as a de facto force multiplier, “a consequence of the aforementioned loss of haven and leadership—al-Qaeda has been forced to increasingly rely on other terrorist groups who are in some way connected to the network to conduct its attacks.”

According to Rabasa et. al, these outsourced and ongoing attacks, sponsored by subcontracted militant groups, “ensure that the group continued to be viewed as both vibrant and relevant, they also helped to underscore an image of the network as being able to strike at will around the globe.” Under the patronage of the Taliban, Al Qaeda received a safe haven to train militants, indoctrinate them, and simultaneously conduct effective attacks. However, the strategic transition from an identifiable operative core to a loose network of subcontracted militants is in
part reflective of less capable assets and infrastructure (see figure 3). Al Qaeda has now reverted “to other established terrorist affiliates to conduct attacks but also to operate in a more decentralized, cellular fashion.” More recently, Rabasa found that the creation of the virtual jihad and various urban training camps has allowed this network to prevail, igniting a second-generation “wave of terrorism in Morocco, Indonesia, Saudi Arabia, [and] Turkey [that]… is still able to conduct effective attacks despite the loss of training bases in Afghanistan (p. 43).” In addition to using American and European converts for their passports and undetected access through security, Al Qaeda has diversified its resource base beyond the financial assistance it receives from states within the Gulf. Apart from receiving donations from jihad’s numerous supporters, Al Qaeda also commits online fraud and infiltrates charities. Prior to the U.S.’s international coalition against terrorism, Al Qaeda’s central core of senior elites would engage in years of operational planning for ‘spectacular’ attacks. In place of conducting these large-scale, complex ‘spectaculars’, strikes have become smaller and more localized attempts to resuscitate the jihadi cause as part of a global franchise. Al Qaeda’s affiliates have enough endowment to “make operational decisions on their own,” employing shortened planning cycles and bombarding soft targets.\textsuperscript{4}

There were many similarities between ISIS and Al Qaeda, as they are now both parent organizations. However, it was not until drastic disagreements in strategy, approach, and recruitment, initiated their split in the wake of 2014. It was under the mature mediation of Al Qaeda leader, Ayman al-Zawahiri that relations between Al Qaeda and the then-AQI branch were beneficial. Today, the two rival entities compete within the same network of donors, suppliers, and resources for the “soul” of the jihadist movement. To understand what caused this stringent divide, we must assess the fundamental differences in tactics between the parent organizations. Osama bin Laden’s effective leadership aligned a terrorist network that amassed rebel groups for the
global jihadist cause. As the radical son of Saudi elites, Osama bin Laden sought to unite a network of independent, dilettante jihadist fighters with the greater goal of removing their foreign enemy, the United States, from the Middle Eastern region. Militant groups, subdued by oppressive regimes, were swept under Al Qaeda’s umbrella with a myriad of assets, weapons, expertise, training, and funding. However, the death of bin Laden would prove detrimental to the unity of this terrorist network. Unlike bin Laden’s elitist goal to destabilize the West, Abu Musab al-Zarqawi believed the principal enemies of the jihadist movement were Shiites for their false interpretation of Islamic theology and Sunni apostates for their support of the West and corrupt Arab rulers.

At the pinnacle of the U.S.’s invasion into Iraq in 2003, bin Laden lent money to Abu Musab al-Zarqawi, leader of the initial Iraqi Islamic Caliphate, with hopes that he would join Al Qaeda’s network (see figure 2). Although initially reluctant, Zarqawi fastened to bin Laden’s network and formed Al Qaeda’s Iraq branch or “AQI”. From the initial phase of this partnership, there were immediate differences in perspectives, while Zawahiri and bin Laden focused on an anti-West agenda, Zarqawi saw more promise in fighting sectarian wars and Sunni Muslim apostates. In public, bin Laden and Zawahiri embraced this relationship, but it was clear Al Qaeda had absorbed more than it could handle. This became evident when AQI’s violence towards Sunnis provoked their backlash and subsequent partnership with American ground troops. Bin Laden was advised by Al Qaeda spokesman, Adam Gadahn, to sever its ties with AQI, “because of the group’s sectarian violence.”5 When the Syrian civil war broke in 2011, then-AQI leader, Abu Bakr al-Baghdadi, sent fighters to Syria to establish a base organization. Although AQI’s strategic move to send operations to Syria had revived their Iraqi jihadist movement, it ultimately prepared for their secession from Al Qaeda’s main leadership. In charge of connecting the then-parent and child
organizations, Zawahiri let AQI take the reins and build its own name in Syria. When the former-AQI sect coined the name “Jabhat al-Nusra,” Zawahiri tried to correct the situation by claiming Al Qaeda’s affiliation, but Baghdadi refused, causing the inevitable demise of the partnership in 2014.

Almost overnight, Baghdadi’s forces gained significant territory in Iraq, including strategic access to Syria, hydroelectric dams, and lucrative oil refineries. By July of 2014, Baghdadi renamed his operations as the Islamic Caliphate. Perplexed, Al Qaeda member branches were declaring their support for ISIS but Zawahiri understood that Al Qaeda would no longer take center-stage in the global jihadist movement. Primarily, Al Qaeda believed apostate killing sprees were inconclusive to extracting the “root cause” of the Middle Eastern problem, the U.S. Al Qaeda believed that the U.S.’s continued support for corrupt regimes within the Middle East prevented the necessary uprisings for regime change. Essentially, while Al Qaeda focused on the “far enemy,” ISIS prioritized the “near enemy,” such as the Assad regime in Syria and the Abadi regime in Iraq. Additionally, while Al Qaeda collaborates within and outside its terrorist network, the Islamic Caliphate chose to spearhead its organization and remain independent. ISIS uses amateur lone-wolf attackers, targeted by jihadist propaganda to incentivize recruitment. While ISIS consolidates power through expanding territories and coalesced armies, Al Qaeda believed a caliphate was a long-term goal only to be initiated with popular support. Since ISIS originated its power during the civil uprisings in Iraq and Syria, it takes a militaristic approach with mass forces and artillery. Using fear-enticed displays of public beheadings, rape, and crucifixions, ISIS terrorizes populations into their compliance and conversion. Keeping in mind these distinctions, it’s clear that Baghdadi grew tired of Al Qaeda’s slow and steady pace. Moreover, the age difference between the leaders of the two organizations spans the length of twenty years. In addition to differences in tactics, approach, and strategy, ISIS utilizes Twitter, Facebook, and
Instagram to attract the attention of younger, distraught Muslim males. The Islamic State has become the deviant successor group that stemmed and evolved from Al Qaeda but ultimately changed its mission, utilizing brutalist militarized warfare and modernized methods of recruitment. According to Byman and Williams, “ideologically, the sectarianism [ISIS] foments is worsening Shia-Sunni tension throughout the region. So the Islamic State is a much bigger threat to Middle Eastern stability than Al Qaeda ever was.”

**Consulted Literature**

According to Dr. Bernard Haykel’s, “ISIS and al-Qaeda—What Are They Thinking? Understanding the Adversary,” the jihadi militant movements are consequential to the Sunni revivalism experienced in the brutalized Arab politics of domestically repressed states from Algeria to Iraq, with the two groups seeking “to empower Muslims against what they describe as Islam’s enemies, both external and internal.” Haykel describes the relationship between ISIS and Al Qaeda as one of contentious competition for resources. As Haykel explains, though the two groups initially split through differences in approach and, “despite some ideological and operational differences, both Al Qaeda and ISIS tap into a deep vein of resentment, disillusionment, and disenfranchisement, specifically among the Sunni Arabs.” Haykel laments that Al Qaeda’s more elitist view on jihadism’s radical crusade over the West was less relevant to Zarqawi’s Sunni followers, who were mobilized on the ground. Over time, Al Qaeda’s lack of reoccurring ‘spectacular’ threats dimmed its extensive vision, making it “less appealing over time - the attacks of 9/11 were in a sense AQ’s Achilles’ heel since they could not be surpassed.”
Party Coalitions and Culture

In Okamoto’s article, “Organizing across Ethnic Boundaries in the Post–Civil Rights Era: Asian American Panethnic Coalitions,” she discusses the creation of a pan-ethnic Asian American Alliance due to initial insensitive threats towards Chinese Americans. Over time, these remarks were marketed as a threat to all Asians, allowing for the larger goal of collective movement activity. This shared alliance was first used to combat high-levels of imposing anti-Asian violence and sentiment. Eventually, the shared cultural history was utilized to broaden pan-ethnic Asian issues, forming a cross-ethnic coalition. As Okamoto explains, “cross-ethnic coalition events demonstrate that organizations with different interests, goals, and histories can effectively coordinate their efforts for a common cause.”

Okamoto outlines the various conditions that encourage groups to form organizational alliances across ethnic boundaries. The conditions include broadening social forces, sharing racial ideology, influential leaders molding group interests, powerful counter-movements with restrictive legislation, and threats to these organizational movements. Additionally, Okamoto explains the “Resource Hypothesis and Competition Theory,” whereby a coalition will form only under noncompetitive conditions when resources are plentiful. She clarifies that “when racial groups are in competition for scarce resources, racial boundaries are heightened, and coalitions, between different Asian ethnic groups should be more likely.” To test this hypothesis, Okamoto uses a data set of 393 collective action events involving Asian-Americans in thirty metropolitan areas from 1970 to 1998. The author defines a cross-ethnic coalition as a “coordination of efforts between two or more organizations that serve different ethnic populations in the Asian American community.” In relaying Okamoto’s research to this paper, her “Resource Hypothesis and Competition Theory” are instrumental in setting the groundwork for Al Qaeda and ISIS’s potential alliance and
competition framework. If this is correct, ISIS’s current setbacks and loss of its caliphate could level the ‘playing field’ for a coalitional occurrence between the two rival organizations.

**Business Alliances**

According to Wiklund’s and Shepherd’s article, “The Effectiveness of Alliances and Acquisitions: The Role of Resource Combination Activities,” the potential value of an alliance or acquisition that combines resources, “depends on the ability of the firm to discover and conduct productive resource combinations.”\(^9\) The authors list examples wherein the alliance or acquisition was short-lived or less profitable than first assumed. One such example includes the strategic alliance between Apple and IBM in 1992 to create a joint operating system, called Taligent. Despite intentions of success, the alliance dissolved after failed attempts to merge. Although, “examples of alliances and acquisitions that fail to deliver on the intended synergies are mounting, yet they are vigorously pursued by companies around the globe.”\(^10\) Not only do firms differ in their resource portfolios but they vary in their ability to achieve the synergistic benefits of bundling resources. Moreover, resource-sharing significantly impacts the alliance framework, “the value creating potential of the resource portfolio increases when the resources acquired are complementary to the existing ones—the joint use of existing and acquired resources can yield a higher total value than the use of each set of resources independently.”\(^11\) Therefore, achieving synergy takes more than complementary resources, it requires a realized valuation and full exploitation of the resource. A streamlined focus on value creation is an incomplete understanding of the complementary resource’s potential performance. For their study, Wiklund and Shepherd focused on the integration of resource bundling “to realize the potential value of the complementarities.” As they explain, “…productive performance outcomes [depend] on the extent
to which management engages in activities that [bundle] them together within the firm’s resource portfolios.”

Wiklund and Shepherd arrived at three hypotheses, whereby one demonstrates the comparable application to the expounding research on terrorist network alliances. According to the authors, “The more an international firm conducts resource combination activities, the more positive the relationship between (a) engaging in international alliances and performance, and (b) engaging in international acquisitions and performance.” To test their hypotheses, the authors investigated 319 small firms and assessed their resource combination activities to ascertain the effectiveness of alliances and acquisitions. With these hypotheses in mind, the employment of successful resource sharing activities could resemble the inner-workings of a vast terrorist network.

In the “Five Principles to Business Alliances,” Hughes and Weiss outline five key aspects to keep in mind when aligning and merging business organizations. Their first point emphasizes describing the desired relationship, “focus less on defining the business plan and more on how you’ll work together.” This includes a detailed business plan and contract, deciding how the two organizations will interact to better clarify the nature of the partners’ working relationship. In terms of collaboration, firms must know how their counterparts operate, including organizational structure, policies and procedures, and cultural norms to later ensure mutual trust and respect. The second point explains that firms should “develop metrics pegged not only to alliance goals but also to alliance progress.” Frameworks for alliances should include shared goals such as: increased revenue, reduced costs, gains in market share, and the like. Alliance performances should also be measured in terms of “ends” and “means.” The “ends” measurement indicates financial value while the “means” measurement includes the various processes necessary to conduct information
sharing, speed of decision making, and the development of new ideas. It is also essential to highlight the differing expectations and qualitative measures of alliance progress in monthly audits. Without measuring these pay-offs, resources such as time, effort, and investment may be better deployed elsewhere.

The author’s third point suggests that “instead of trying to eliminate differences, [firms should] leverage them to create value.” According to Hughes and Weiss, “companies ally because they have key differences they want to leverage—different markets, customers, know-how, processes, and cultures.” By minimizing the conflict of highlighting differences, companies can focus on reaching agreements. Oftentimes conflict “arises from deep discomfort with differences and assumptions that the same management strategies would work with external partners.” Firms should utilize these unique competencies to “leverage complementary strengths and assets.” The authors also explain that firms can sometimes be indecisive “bureaucratic dinosaurs,” upsetting more “nimble” operational types. Although these behavioral differences can become incompetent and untrustworthy, once these firms acknowledge their areas of strength and weakness, they can effectively respect their differences in a power-sharing structure. Moreover, firms should “go beyond formal governance structures to encourage collaborative behavior.” By nurturing and fostering collaborative behavior throughout the chain of command, firms can promote equal accountability, whereby the firms “develop dispassionate analysis of how both parties contributed to problems that arise in the interdependent relationship.”

Lastly, the authors explain that “companies are not monolithic, yet alliance advice tends to gloss over this basic reality and treat partners as if they were simple, homogeneous entities.” The largest barriers to effective alliances include, “mixed messages, broken commitments, and unpredictable, inconsistent behavior from different segments of a partner organization.” Firms
should devote time to understanding their counterpart’s business culture and “defining the rules of engagement for guided interactions.” Additionally, firms must ensure that different branches don’t have varying levels of willingness to invest and cooperate with the counterpart. In an example, the authors discovered that most executives negotiate firm alliances but “true buy-in from other parts of the enterprise had never been secured.”

In referencing, “Five Principles to Business Alliances,” it is important to keep in mind, that both ISIS and Al Qaeda have different bureaucratic structures, which initially ignited their split. While Al Qaeda is more of a “bureaucratic dinosaur,” calculably planning and weighing the pros and cons of ‘spectaculars,’ ISIS appears to be nimbler with its usage of ‘lone-wolf’ attackers. This represents a difference in strategy that would have to be considered in terms of creating a power-sharing structure.

**Partnerships and Mergers**

In “Approaching Merger: The Converging Public Policy Agendas of the AFL and CIO, 1938–1955,” Cornfield and McCammon assess the eventual merging of the AFL and CIO despite initial apprehension on both sides. The Labor Federation was known for promoting its broad public policy agenda on social welfare, labor law, employment policy, international trade, public education, and other public policies in federal and state legislatures. Both federations subscribed to these public policy objectives before the merger, as evidenced in the wide-ranging resolution introduced at their conventions. In converging their ideological goals, “the merger was expected to concentrate, rather than to disperse, efforts in support of favorable legislation.”13 In citing Gregory Maney’s, “Transnational Mobilization and Civil Rights in Northern Ireland,” the authors argue that the combined working unity of the groups was achieved due to the initial “ideological
congruence between movement organizations.” Moreover, “for coordinated action to occur, groups must have a certain degree of overlapping interests, goals, and core beliefs.”

As Steven Barkan explains in his 1986 Civil Rights Movement study, “groups with pronounced disagreements over policy goals will rarely come together and work jointly to achieve social change.” Interestingly enough, before the eventual merging of the AFL and CIO, the two parties had incomparable differences in organizational structure that led to unsuccessful partnership efforts. “In the years before the 1950s, conflicts between the AFL and CIO over organizational philosophy, union jurisdictions, membership recruitment, and public policy goals, as well as personality conflicts among labor leaders, led to several abortive merger efforts (Seidman 1956).” By studying the agenda from organizational conferences, Cornfield and McCammon found that “…a shift in the ideological orientation of one of the groups that entailed a broadening in the AFL’s policy agenda and its widening embrace of the emerging New Deal welfare state set the stage for organizational merger.” In layman’s terms, the AFL widened its policy agenda to pave a path for the CIO’s shift in ideological orientation. Moreover, the authors discovered that prior to the merger, the AFL had been increasingly criticized for its lack of aggressive organizing and less inclusive model, in comparison to the CIO’s “industrial unionism, in which all production workers of a single industry, regardless of occupation and skill level, belonged to one union.” It appears that years of insurgency within the AFL’s structure led to the eventual merging of the separate groups in 1955. During the period of tension that lasted from 1938 to 1955, the legislative goals of both the AFL and CIO had increasingly coincided. Overall, Cornfield and McCammon note that “the growing convergence in the public policy goals of the two federations resulted mainly from a broadening of the range of public policy domains of interest to the AFL.”
In assessing this source for its application to the terrorist alliance framework, there exists a connection between organizational structures, combined working unity, and ideological congruence, to determine the merging of ISIS and Al Qaeda. When comparing ISIS to Al Qaeda, it appears that one may have to relinquish its bureaucratic structure and less nimble approach to provide for the acquisition of the other, unless coinciding goals are identified.

**Research Design**

In comprehending the conditions for merger among militant groups, case study applications have been utilized to provide a historical context to this qualitative research study. By honing in on these applications, one can better understand the situational rationale that causes radical groups to unite and join forces. For the purpose of this study, two militant group networks, the HuJI and HuM of Pakistan and India and the FARC, ELN, and EPL of Colombia have been analyzed and applied to the aforementioned literature with the goal of ascertaining the conditions in which ISIS and Al Qaeda would likely unite.

**Militant Group Case Studies**

*The HuJI & HuM Militant Groups*

According to Dr. Haykel, despite the recent loss of territory, "reuniting would mean that IS members have to give up on the idea of the caliphate." However, after consulting various case studies, it appears that militant groups capriciously end and mend their relationships, proposing a deviation from the pragmatism suggested by the business alliance literature. Upon completing a case study for the HuJI and HuM militant groups of Pakistan, Khurshchev Singh’s article, “HuJI in India: An Assessment,” explained that the mujahideen forces were “created by the United States..."
and Pakistan to fight against the Soviet occupation in Afghanistan during 1979–1989.” In 1985, the Harkat-ul-Jihadi Islami or HuJI splintered into two groups, the HuJI and the Harkat-ul-Mujahideen (HuM), led by Fazlur Rehman Khalil. After Soviet withdrawal in 1989, the HuM’s predominant focus was on the annexation of the Jammu and Kashmir (J&K) regions, aligning with the radical Deobandi teachings of the Jamiat Ulema-e-Islam, which advocates for a strong pan-Islamic ideology in India. In 1993, Pakistani General, Pervez Musharraf helped to unite HuM and HuJI, forming the Harkat-ul-Ansar or HuA. “While HuM was fully involved in terrorist attacks in J&K, the HuJI was not that active,” causing their eventual split again, in 1997. Additionally, multiple subgroups were formed as different networks and recruits were identified in various areas across Southeast Asia.

By 1992, “HuJI-B” or the Bangladeshi subgroup, was supported by the efforts of Pakistan’s Inter-Services Intelligence (ISI) and Al Qaeda under the leadership of Osama bin Laden. Subsequent to Farooq Kashmiri assuming leadership of the HuM in February of 2000, former members established an off-shoot group, Jaish-e-Muhammad, led by Maulana Masood Azhar (see figure 4). According to Peter Chalk, an expert on terrorism at the RAND Corporation, “although these groups remain organizationally distinct and exhibit subtle ideological and theological differences, they are characterized by an overlapping membership that variously reflects personalities, a common anti-Indian agenda, and shared experiences in overseas militant training camps.” In addition to being offshoots, various external groups, including the “LeT, JeM, and HuM – have been directly linked to bin Laden’s transnational terror network.”

In terms of goals and operations, Khurshchev Singh explains that because the group has various subunits in different countries, multiple objectives and activities have been identified. Although the HuJI was initially created to serve Afghan mujahideen camps during the Soviet-
Afghan War; today, their general mission is to provide defense for Muslims, spread destabilizing terrorism, and wage jihad in captured territories with the hopes of establishing Islamic rule. While their various outfits differ in operation within their respective locations, HuJI –Pakistan, HuJI-Bangladesh and India, don’t appear to clash in their general objectives, as they are operating from separate locations. Moreover, although the various subgroups have well-developed networks and donors, they do not look to outbid each other. This is unlike the competitive behavior of the ISIS-Al Qaeda relationship, which has been witnessed within the scholarly community and previously mentioned by Dr. Haykel.

Simon Bolivar’s CGSB and Member Guerrilla Groups: FARC, ELN, and EPL

The Armed Forces of the Columbian Revolution, or FARC, originally formed in 1966 as a communist consolidation of proletariat defense militia, whose groups dated back to the prior decade. The FARC-dominated territory became known as an “independent republic,” under President Guillermo León Valencia, who coined the term “Tirofijo,” to describe the rogue state that was predominantly “neglected by the national government and plagued by general lawlessness.” During La Violencia, the Colombian Civil War that waged between the conservative and liberal parties between 1948 and 1958, the FARC strengthened its ties with the Marxist-led Communist Party based on its disillusionment with the liberal party. Due to the deliqusce of the Soviet Union, the FARC received a major truncation in its financing. Moreover, its reliance on Marxism provoked a rather anachronistic reputation. Therefore, the FARC began to rely heavily on the guerrilla warfare tactics of kidnapping for ransom, receiving protection rents for illegal drug development and extorting businesses. As Harvey F. Kline explains in “Chronicle of a Failure Foretold: The Peace Process of Colombian President Andrés Pastrana,” this strategy
merely funded the militant group’s efforts and expansion; however, it “did not win them new sympathizers (Kline, 2007, p. 11).”

The promotion of FARC’s Marxist ideology lost touch with public support as the U.S. government increased its sponsorship of the Colombian military, however, this did not negatively impact its enlargement of territory and recruitment. FARC began attacking soft targets, “engaging in actions that harmed civilians, such as land mines and the notoriously inaccurate gas-cylinder bombs, hijacking commercial jets, assassinating elected officials, murdering peace activists, and attacking an upscale family recreation center in the heart of Bogotá.” Moreover, at its zenith FARC drew in a profit of approximately $600 million from its illegal drug trade, kidnappings, and extortion schemes.

When the Army of National Liberation or ELN was officially formed in 1964 by a group of left-wing radical Catholic students, they were inspired by Marxist ideology and the Cuban Revolution. The ELN included a democratic approach to structure, instituting operational meetings to elect position-holders. Initially, the ELN refrained from interfering with the drug trade but after witnessing its lucrative value, the ELN began engaging in kidnappings and the pilfering of oil pipelines. According to Kline, “some of the group asked for and obtained military training and began a series of discussions about a foco (focus) strategy for Colombia.” Moreover, the ELN focused on targeting the coal and oil sector through its kidnapping and extortion stratagem. In the 1990s, ELN suffered from a loss in territory due to increased governmental security. Additionally, U.S. interference in the Colombian military and investment of its energy sector impacted the ELN’s funding and recruitment.

The Ejército Popular de Liberación (EPL, Popular Liberation Army) was established in 1967 as an armed and Mao-influenced subsect of the Colombian Communist Party. Akin to the
ELN, the EPL also operated out of Santander, “an area of the country traditionally affected by disputes over land between landless campesinos and large landowners,” rendering it vulnerable to extortion.\textsuperscript{27} By the 1970s the EPL utilized kidnapping and extortion efforts to finance its militant endeavors. Within the same decade, the M-19 was formed by recusants dismissed by the Communist Party and the FARC. Although the M-19 was initially considered a subsidiary urban guerrilla group, its kidnapping and murder of the National Labor Federation’s leader in 1976 established its claim to fame. By 1990 both the EPL and M-19 entered into peace talks with the Colombian government, allowing for demobilization and disarmament. As a result of this agreement, the former guerrilla groups became legal political parties. The M-19 became the Alianza Democrática M-19 and the EPL became Esperanza, Paz y Libertad, although a minority of its members maintained militancy.

Under President Belisario Betancur in 1984, a Peace Commission was formed to address the guerrilla group activity. According to Kline, Betancur believed involvement in guerrilla groups was “the product of objective circumstances of poverty, injustice, and the lack of opportunities for political participation.” Although Betancur reached an agreement between the FARC and EPL, the ELN had not signed the peace treaty.\textsuperscript{28} The next year, in 1985, “the peace was punctuated by breaches of the ceasefire committed by both sides.”\textsuperscript{29} In 1987, subsequent attempts by the government under then-President Barco had failed to reach an agreement that would result in the disarming and dissolving of the militant groups. Under President César Gaviria, the FARC, ELN, and EPL “entered a coalition through a \textit{Coordinadora guerrillera Simon Bolívar} (CGSB, Simon Bolivar Guerrilla Coordinator) in 1987.”\textsuperscript{30} The CGSB banded together within the same Communist agenda, operating within Colombia to attack lucrative oil rigs.\textsuperscript{31} The coalition mainly formed based off akin doctrine and the collective objective of expanding monetarized and militarized
support. Moreover, the groups stem from the same recruitment of the rural poor to combat Colombia’s wealthy class and oppose U.S. influence. It’s important to keep in mind that the FARC, ELN, and EPL formed relatively around the same time period of the 1960s with a fundamental Communist agenda. While the ELN was more politically motivated and ideologically focused due to its founding by Catholic students, the FARC initially took a more militant and guerrilla warfare based approach. Although the FARC is considered less unified, it does incorporate a central command despite having various factions due to the different operating areas.32

In a similar fashion to Al Qaeda’s more recent subcontracting of rebel groups in homage to the jihadist global mission, the FARC, EPL, and ELN relied on recycling their militant recruits. The CGSB was an effort to consolidate the various groups under one potential power-structure. Had it not been for U.S. sponsorship of the Colombian military and the repeated governmental efforts for peace treaties, the CGSB likely would have remained. However, while the global jihad expands over countries from Afghanistan to India, the militant groups aligned within the CGSB only operate in one country. However, these militant groups do form their own de facto states within conquered territories, acting as separate entities within Colombia.

**Expositions**

In a brief interview with the Director of the Institute for Transregional Study of the Contemporary Middle East, North Africa, and Central Asia at Princeton University, Dr. Bernard Haykel, answered questions that addressed the framework for a possible merger between ISIS and Al Qaeda (AQ) forces. In response, he stated that he sees “AQ and IS as competitors and reuniting would mean that IS members have to give up on the idea of the caliphate, which is not easy to do. Certain individuals in IS might reunite with AQ and give up on the institution of the caliph, but
both organizations have been severely weakened and incapacitated. This is because so many of
their members have been killed or imprisoned and also because their political projects have failed
to bring about the victory that they keep promising to deliver. In my mind, the constitutive elements
that feed global jihadism are still very much present in the region (the poverty, overpopulation,
environmental degradation, lack of opportunity, high expectations and relative deprivation, state
brutalization and the overall disenfranchisement of Sunni Arabs in Iraq and Syria), but at the same
time countries like Saudi Arabia have given up their funding and support of all forms of Islamism
and are now engaged in a campaign of suppression of these activists.” As outlined in Haykel’s
quote, there is a contentious competition between ISIS and AQ for resources within the region.
Various operational goals have been identified, which prevents the two jihadi militant groups from
attempting a merger.

Mr. Asslan Sayyar, the Defense Program Director for South Asia at the United States
Central Command (CENTCOM), has reviewed and provided his perspective on the paper’s
research question. In his comments, he agreed with Dr. Haykel that ISIS and Al Qaeda engage in
continuous outbidding with each other for resources, making themselves more attractive to
potential donors and recruits. The groups compete to gain maximum exposure and funding. Mr.
Sayyar believes reuniting likely will not be an option unless ISIS was immensely desperate. He
claims that although the physical caliphate is in the process of its eradication, the mental and more
virtual platform is one that’s still very much present. Upon acknowledging the aforementioned
business alliance literature, Mr. Sayyar argued that “mergers tend to happen when survival is more
important than growth and for ISIS, it’s not there yet.” Therefore, the research findings now beg
the question: has ISIS been defeated enough to concede and join forces with Al Qaeda, similar to
the amalgamation between the AFL and CIO?
Analysis

In applying the aforementioned literature to terrorist network research, Al Qaeda and ISIS are both international organizations, having multiple franchises in different areas throughout Europe, the Middle East, Africa, and the Maghreb. The effectiveness of their umbrella networks should be assessed based on how well they incorporate resource endowments and moderate their independent branch activities. The initial demise that limited the working relationship of Al Qaeda and its branch in Iraq (AQI), rests in the fact that Bin Laden assumed the homogeneity of the two groups. Bin Laden surmised that AQI, under Baghdad’s leadership, would support Al Qaeda’s strategic operational structure and its overall view to destabilize the West. However, time would prove that Al Qaeda and its branch in Iraq, which later reassembled to form modern-day ISIS, had differing “ends” and “means,” referencing to the aforementioned business alliance literature. For Al Qaeda, its “means” of steady information-sharing for strategic mission operations justified its securement of an “ends” in financial support, increased recruitment, and dynamic network structures. However, ISIS subscribed to a more incompatible and stark nature, whereby its “means” of brutal and terrorizing attacks, fortified an “ends” of increased Islamic territory and on-ground impairment of corrupt Arab regimes and Sunni apostates. In referencing Okamoto’s, “Resource Hypothesis and Competition Theory,” wherein coalitions form under noncompetitive conditions with plentiful resources; as Haykel explains, conditions are competitive and resources are dwindling. Not only have these operational differences become more apparent through the application of the literature but both groups maintain different working structures. Moreover, while ISIS’s network remains loose and nimble, Al Qaeda is more bureaucratic and disciplined in nature. Therefore, a potential coalition is not likely to form under the party coalition and business alliance standpoints.
Through careful inspection of the literature, the business alliance framework relevantly compares to the intertwining networks and motivations of terrorist organizations. One contention lies within the party coalition literature, which suggests that groups of similar cultural background would be most compatible for a potential merger. Although ISIS and Al Qaeda are predominantly Sunni organizations, they’ve previously targeted operatives on both sides, potentially damaging the grounds for a working relationship. Therefore, in accordance with the literature, pertaining to business alliances, partnerships and mergers, and party coalitions, ISIS and Al Qaeda still differ in their goals, operations, and approach, failing to coincide ideological goals and operational structures. While ISIS prefers brutally terrorizing civilians to amass territory for an Islamic caliphate, Al Qaeda prefers to play the waiting game by reserving eventual governance and geopolitical domination.

Conclusion

As discussed and referenced in the business alliance literature, Al Qaeda initially failed to assess for shared goals within its acquired branch in Iraq, or AQI. Al Qaeda assumed ISIS would utilize the same strategies and operational tactics when it first assumed Baghdadi’s group in Iraq. While some alliance frameworks may assess these jihadi militant organizations as homogeneous entities, they ultimately differ in their operational nature, as supported by the business alliance literature. Although ISIS and Al Qaeda may stem from the same radical Islamic network, they vary in their “ends” and “means” approach to the goal of the global jihad. While Al Qaeda utilizes precise and calculated decision-making operations to increase network expansion and financial sponsorship, ISIS employs brutalized guerrilla warfare tactics under a centralized hub to expand territory and recruitment. Most recently, ISIS, under the leadership of Abu Bakr al-Baghdadi, has
released propaganda videos through social media in attempt to revitalize the jihadi network despite the recent loss in caliphate due to increased security in the region.\textsuperscript{33}

In conclusion, while Al Qaeda utilizes its calculable operational planning cycle to effectively engage in offensive attacks, ISIS prefers on-ground, brutal warfare tactics to amass territories. This diverging operational structure, paired with a failure to coincide ideological goals, maintains that the two terrorist organizations will not likely reunite. In the example of the HuJi offshoots from Pakistan, India, and Bangladesh, these groups do not compete as they are operating and recruiting from separate locations. In a similar fashion, the FARC and ELN did not compete despite their fundamental communist agenda because they recruited members from different socio-economic classes. While the FARC recruited destressed campesinos from the rural parts of Colombia, the ELN focused on urban, educated dwellers. Moreover, the CGSB was able to converge these akin doctrines under a centralized power-structure due to this initial lack of competition. Unlike the FARC and ELN, regional conditions between ISIS and Al Qaeda have been competitive as resources continue to dwindle. They continue to outbid each other for scant resources, funding, and sponsors, provided by supporters of the jihadist global movement within the Gulf states. Moreover, Al Qaeda uses calculable operational planning cycles for offensive attacks while ISIS utilizes on-ground, brutal warfare tactics to increase Islamic territory. With these findings, ISIS and Al Qaeda have not converged their diverging operational structures nor have they concentrated their ideological goals to prepare for a potential merger. Furthermore, their continued rivalry in the recruitment of militants, suggests that assets and infrastructure within the terrorist network are diminishing due to an increased crackdown on jihadi franchises.
Relevant Figures

Figure 1.

The Islamic Caliphate

Courtesy of CNN
Figure 2.

Abu Musab al-Zarqawi (left) and Osama bin Laden (right)

Courtesy of BBC
Figure 3.

Al Qaeda’s Network of affiliates and subcontractors

Courtesy of RAND
Figure 4.

Al Qaeda’s South and Central Asian Clusters

Courtesy of RAND
### Merger Analysis

By Bianca Pergher

<table>
<thead>
<tr>
<th>Did they ally?</th>
<th>Different operational structures/procedures prior to the merger?</th>
<th>Effective Resource sharing?</th>
<th>Original apprehension or strife prior to the merger?</th>
<th>Consolidated ideologies to prepare for merger?</th>
<th>Differing goals causing a split? Are they peaceful?</th>
</tr>
</thead>
<tbody>
<tr>
<td>HaJJ/HeM</td>
<td>Yes, HuA, HuJ: mostly inactive</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes, different regional focuses; Yes, they splintered into various groups in 1997, forming the LdL, JeM etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FARC/ELN</td>
<td>Yes, CGSB, FARC: proletariat, guerrilla group, EPL: armed Catholic students</td>
<td>Yes, overlapping membership but varying recruitment</td>
<td>No</td>
<td>Yes, the three groups had similar Communist agendas</td>
<td>No, while the EPL became a political party, the FARC and ELN maintained; Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISIS/AQ</td>
<td>Yes, initially ISIS was AQ’s branch in Iraq</td>
<td>Yes, Bin Laden provided funding to the initial AQ in Iraq branch</td>
<td>Yes, Zarqawi had initial apprehension before merging with AQ</td>
<td>No</td>
<td>Yes, differences in operational strategy and goals; No, both sides have encountered casualties. However, the removal of the older generation may end the rivalry.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFL/CIO</td>
<td>Yes, AFL: unionized each labor sector</td>
<td>Yes, initial apprehension on both sides</td>
<td>Yes, AFL expanded its public policy agenda to merge with CIO</td>
<td>No</td>
<td>No split</td>
</tr>
<tr>
<td>Asian Coalition</td>
<td>Yes, forming a pan-ethnic coalition</td>
<td>Yes</td>
<td>Yes, ultimately anti-Asian sentiment merged the groups</td>
<td>Yes, unifying under one Asian identity</td>
<td>No split</td>
</tr>
</tbody>
</table>

Figure 5.

Militant Group Merger Analysis

By Bianca Pergher
Notes:
2 Ibid.
6 Ibid.
24 Ibid.

27 Ibid.


The Cost of the Culturati:

Studying the Neighborhood Stability Impact

of Cultural District Designations

Prakash Mishra
Abstract

The decision to declare a district for a specific cause is a critical policy decision; making an area an official office park or designated cultural site means it will attract specific types of residents and businesses and require specific amenities. This paper reviews the impact of designating a cultural district as a place-based policy, specifically by developing a measure of neighborhood stability and applying a stress test of neighborhood stability in cultural districts during the Great Recession. The model underpinning the neighborhood stability measure is an optimal stopping time model which frames neighborhood rents as a Brownian motion with drift. This structure imposes minimalist assumptions and develops two reduced form parameters which describe individual preferences for how long to live in a neighborhood. This analysis is in the style of Alvarez et al. (2015). The parameters are then used to test neighborhood stability, with the result that neighborhoods designated specifically as cultural districts are far less likely to experience negative stability (e.g., large amounts of residential outmigration and thus shorter residency spells) with a causal effect size four times larger than the effect size of a recession itself. However, such neighborhoods are also more likely to experience an influx of newer higher income residents after designation, implying the beneficiaries of the new stability may be those who priced out the original creators of the neighborhood's cultural capital.
Table of Contents

Introduction 147
Literature Review 150
Theory 153
Data 165
Calibrations 168
Results and Discussion 170
Conclusion 184
References 186
Appendix 189
Notes 197
Introduction

Urban revitalization efforts have played a major role in the much-documented millennial movement back to the city. One of the common toolkits of the urban revitalization planner is the creation of special districts, such as the Avenue of the Arts in Philadelphia or NoDa in Charlotte. These designations are essentially place-based policies which then serve as signals to entrepreneurs and potential residents that these districts are the intended center of specific economic or social activities. Residents seeking identity-specific (such as in gayborhoods) or industry-specific amenities (the arts district) are more likely to settle in these areas due to the perceived benefits of consumption and production agglomeration, respectively. As a particular example, the Avenue of the Arts in Philadelphia was a major initiative on the arterial Broad Street in the Center City District whose motivation was entirely to condemn blighted housing in the district and replace it with creative industries to drive vitality in a new central business district. Importantly, these districts are a city-level policy, usually the product of a department of City Planning initiative. Such decisions tend to come from one of two sources: the generally higher rate of occurrence of relevant businesses already in the area (an agglomeration pathway) or the occurrence of a political demonstration or relevant exogenous event which is being memorialized in the designation (studied in this paper).

Special districts are often loci for high rates of minority entrepreneurship, and the designation as a district can act as a way to send a clearer signal to new, less informed consumers about the neighborhood’s business and residential communities. This formal designation is thus a critical policy measure that in a sense “brands” a neighborhood to potential consumers. A famous example — the formal designation of the South Beach, Florida Gayborhood was a landmark moment in LGBTQ+ migration into the region. While South Beach was certainly a
queer space before this designation by the city, new in-migration from queer folks who were not South Beach residents remarkably increased after designation. A 1 percentage point increase in the number of same-sex coupled households is associated with an almost 2% to 3% increase in the probability of gentrification.

However, such designations can also attract external, higher-income residents to previously highly local neighborhoods, causing pre-designation residents to get priced out as rents and local amenity prices rise. Then, there is a classic real estate investment cycle problem, where revitalization comes at the expense of current residents. What sets this policy apart from other place-based policy design is persistence of the special district amenity. The designation of a special district is often more lasting than a passing place-based economic policy like a tax abatement. As a result, while an industrial park subsidy displaces one generation of buyers, cultural districting policies could drive a more continuous rate of neighborhood turnover as the wealthiest consumers within the relevant cultural group turnover the cultural district itself.

The question of how “stable” designated cultural districts like Gayborhoods, Chinatowns, or Arts Districts truly are is not, therefore, a straightforward question. On the one hand, one can argue that the designation acts as a time-persistent amenity which anchors individuals of a specific “type” to that district. On the other hand, one can argue that such districts because of their high profile after designation may tend to experience a consistently higher rate of instability as residents and entrepreneurs remain in flux with new tastes, trends, and income brackets. With limited land supply in urban areas, a potential cycle of displacement results.

This paper investigates the question of designated cultural district (DCD) stability, and evaluates whether designation as a special district comes at the cost of long-term neighborhood instability. The extent to which the long-term cost is desirable in the name of urban revitalization
from a purely economic standpoint (admitting the shortcomings of such an analysis) is then discussed. I develop a reduced-form distribution to model the decision-making problem faced by residents of a special district on considering when to leave the district. The model frames leaving the district in terms of an optimal stopping time problem with the operative decision being when to leave conditioned on a stochastic process which models neighborhood rents (a Brownian motion with drift).

I thus provide the following contributions to the field: (1) I develop a direct corollary to results in unemployment dynamics to methods of neighborhood choice. (2) I discuss the creation of a measure of neighborhood stability and the challenges this poses. On this, I develop a measure which controls for individual preferences by developing a measure of the “type” of an individual from the model which gives contribution (1), allowing full identification of the impact of neighborhood features on the resident’s expected lifetime in a neighborhood. (3) I argue that designation as a cultural district contributes to robustness of a neighborhood against exogenous shocks using the Recession as a case study, extending expected resident lifetimes despite the shock by an effect size 3 times as strong as the effect of the Recession itself. However, DCDs also likely raise rents and price out previous residents, so it is an open question as to for whom the neighborhood is more stable; the original creators of the culture of the district, or new, potentially displacing forces.

This paper uses data from RentBureau, a credit bureau dedicated to the multi- family industry which collects rental histories from a network of apartment owners and managers. Individual rental terms are observed, including collections, rental amount, and final payment dates, as a census. This dataset is available from January 1997 until June 2010.
The remainder of the paper is structured as follows. In Section 2, I provide further background on the DCD and literature which mirrors the methodological choices of this paper. The theory and structure of the model is built out in section 3, with some comments on the individual model provided in Appendix A. I begin to develop a notion of a mathematical description for neighborhood stability. Brief data summaries follow in Section 4. Section 5 provides results and discusses the development of a policy model which uses the individual-level reduced form parameters to isolate the causal effect of DCDs on neighborhood stability, isolating a precise definition here as well. Section 6 summarizes our key conclusions.

**Literature Review**

New residents of a neighborhood who do not leave after the first year are more likely to stay in the neighborhood for a long period of time. A marketer might suggest a loyalty effect sets in and keeps residents in their neighborhood. The sociological anchoring effects of community back up this explanation as a potentially causal story. A game theorist may argue that individuals arrive in neighborhoods with imperfect information, and that those who sort out in the first year are those who discover that their information was incorrect to the point that the neighborhood was suboptimal. Both likely hold some grain of truth, suggesting that the hazard rate of migration out of a neighborhood is a mix of heterogeneity and duration dependence. This paper uses the labor model of an optimal stopping time problem for hazard rates in- and out of unemployment and applies it to the resident’s optimal out-migration time problem. We develop nonparametric decompositions of heterogeneity and duration dependence in the manner of. The predecessor to this paper, they frame the decision to take a job as an optimal stopping time problem. They cast wages into a Brownian motion with drift, and I intend to do the same with
local neighborhood rents. Their derivation also allows decomposition of the job-finding hazard into heterogeneity and duration dependence, while developing a nonparametric estimation for a paper of latent sufficient statistics. Studies of separation into unemployment have used this structure to analyze whether recall expectations (expectation of re-employment as in the case of a temporary layoff) muddle duration-dependent signals. Generally, the nonparametric mixed hazards model has been used with time-varying covariates because it is quite easily identified. The method across the literature generally devises a structural evolution for key individual-level parameters, popularly including individual discount rates and some stochastic or fixed formulation for reservation wages. Spinnewiin introduces biased beliefs in to the classic two-period model of unemployment, obtaining sufficient statistics from a modified Baily formula. He proves that in this setting identification of two moments of the individual utility problem is sufficient for policy design. Generally, sufficient statistics approaches estimate wage elasticity to a parameter of interest, such as uninsurance benefit rates.

The sufficient statistics is also quite present in the neighborhood choice literature, to which this paper makes its contribution. The search for such sufficient statistics is in particular still a topic of discussion. Segregation levels, for example, are not able to characterize racial sorting impacts on income and education. Wealth has been estimated to be a key driver for neighborhood relocation and churn (in a model, notably, which does not account for multi-period savings or consumption). I argue that generally, churn is a more apt sufficient statistic than rents alone in the place-based policy literature because such policies tend to price out existing residents and gentrify neighborhoods. Indeed, rents primarily benefit city government tax budgets and those who afford to own local housing capital; gentrification imposes costs by pricing out non-owning, often low-income, households, causing loss of efficiencies from accrued
local knowledge. Owners on the margin, who can afford to own property, can be driven out by higher taxes. Thus, quantifying churn is of major interest when thinking about efficiency. A persistent threat of asset loss can impose well-being effects and generally make individuals less likely to use credit, affecting their ability to make own-optimal consumption choices.

Generally, measures of demand for skilled labor is sufficient to estimate tract demographic, population skill-levels, and housing prices. When analyzing impacts, on displacement, the sufficient statistics tend to come from rents or are indirectly derived from a rent-based measure. This is the perspective this paper will adopt and subsequently defend. This is consistent with literature on neighborhood choice from other perspectives as well. For example, models of private provision argue that an “impatience rate” can be viewed as the source of heterogeneity in deciding to privately provide a public good. Then, the individual level decision in the literature tends to be driven (akin to in the labor literature) by two factors; a timing problem and a rent level.

The experiment in this paper is an impulse response (exogenous shock) experiment which uses changes in hazard rates during a recession to test differences across neighborhoods. A study of the Great Recession on the long-term unemployment rate uses a similar decomposition into duration dependence and heterogeneity to compare pre-Recession and post-Recession hazards. Another paper uses the exogenous shock of plant closures to determine whether there is a significant signaling effect in unemployment duration. Their model is a simply proportional hazards specification with a single-term estimator, however - not a fuller structural approach.

The final relevant literature is the discussion of the anchoring effect of cultural districts. Amenities can act as anchors which fix neighborhoods to certain income levels with less volatility over time when they are persistent. Empirical results on the topic are mixed.
paper questions whether designating an area as a cultural district can act as an anchor point for higher income members of the targeted communities. Actual empirical results around the stability of cultural districts, however, are mixed. An exercise in South Beach, Florida suggests that the queer community before and after the designation of the gayborhood has become more geographically dispersed, and that queer entrepreneurs were gradually priced out by in-migrants from other cities who have taken over the strip.\textsuperscript{28} Theoretically, however, cultural districts like ethnic enclaves should promote agglomeration by overcoming language barriers and work-rule differences in immigrant populations which result in underemployment in other neighborhoods.\textsuperscript{29} City planners have long discussed the merits and drawbacks of anchor-based development which creates anchor institutions that can help prevent fragmentation of communities after development. In the status quo, however, the location of affordable housing is not correlated with proximity to institutional and neighborhood amenities, where anchor-based revitalization is targeted.\textsuperscript{30} This suggests that as an experiment, there is indeed a clean treatment effect when considering specialty districts as possessing potential anchors in the status quo, as opposed to neighboring districts.

**Theory**

I re-formulate the structure set forth in Alvarez et al. (2015). My novel contributions in this theoretical section are formulating and individual choice framework which re-contextualizes this structure to the neighborhood choice setting (Section 3.1) and to aggregate the individual-level data to construct neighborhood-level distribution function variables which make it possible to conduct this analysis with less dense micro-data (Section 3.3). This theory is developed to set forth reduced form parameters to characterize the individual choice problem of determining
when to move out of a given neighborhood. These reduced form parameters, when aggregated to a neighborhood-level distribution, can be used to evaluate neighborhoods on the types of individuals they attract and the relative “stability” of a neighborhood given the mix of renter lifetimes observed in that neighborhood. Only once these measures are derived with a fully specified structure can I begin to craft a causal argument. With this motivation of defining stability as a mixture of individual parameters, I devise the individual choice problem.

*Individual Choice Problem*

I first develop a stylized model of how the resident, $i$, values neighborhood-specific goods (amenities) and all other goods (numeraire consumption). Amenities include cultural and social benefits in designated neighborhoods. The model does not *a priori* assume higher amenity value in such districts, and this piece of the narrative is not tested in this paper. This individual-level set-up will, however, allow us to derive a set of parameters around individual preferences for neighborhoods later in the paper.

Residents maximize an objective utility function across neighborhoods each period, determining where they choose to live. Greater expenditure on rents implies lower expenditure on consumption, and with greater expenditure on rents tends to come a higher utility from amenities (because one can spend more, access is more likely to neighborhoods with tailored local amenities). Tailoring of local amenities occurs through the gradual supply response to shifts in the consumer population of a neighborhood (a lagged Tiebout model of public and private goods provision). This tradeoff is not absolute a low rent can still lead to a decent value of amenities, so I adopt the following relatively flexible utility function for the individual maximization problem.
The last equation refers to the fact that individuals demand some world-clearing nonnegative amenity value at minimum. For the purposes of this paper, we assume that income is equivalent across neighborhoods \( j \). This is obviously not the case, as some individuals will have specific skills that only earn wages in the vicinity of certain workplaces, but one can craft a skill-adjusted measure of amenities to adjust for such a consideration. This model ignores intertemporal income dynamics induced by savings, and does not permit borrowing. In such a setting, the maximized utility should correspond to

\[
\max_j U_{ijt} = c_{ijt} \cdot A_{jt} \tag{1}
\]

\[
\theta_{lt} \geq c_{ijt} + R_{ijt} \tag{2}
\]

\[
A_{jt} \geq A_{jt} \tag{3}
\]

Imagine now the rent-setting problem. Developers charge the individual a rent which should increase in the value of local amenities with a fixed baseline price according to the land value of the neighborhood. A hedonic component that should correlate with income also comes into play; individuals with higher income are more likely to buy homes with the extra bathroom or bedroom. So, willingness-to-pay for rent should adopt a potentially linear form that mirrors the following

\[
R_{ijt} = \omega_j + \omega_A A_{jt} + \omega_\theta \theta_{lt} \tag{5}
\]

One can adjust for imperfect local competition by adding a markup to this rent function which varies with land supply tightness \((\rho_j)\) or a local cost of new construction that might be
connected to land use policies. This markup can be absorbed by the coefficients and components of rent in (5).

Therefore, residents will choose a neighborhood based on the amount of amenity received for rent. Of course, if rents themselves are a function of amenities present, then we can obtain a marginal utility of amenities in a closed form.

$$\max_j U_{ijt} = (\theta_{it} - R_{ijt}) \cdot \frac{1}{\omega_A} (R_{ijt} - \omega_j - \omega \theta_{it})$$

(6)

$$\frac{\partial U_{ijt}}{\partial R_{ijt}} = \frac{1}{\omega_A} \theta_{it} - \frac{1}{\omega_A} (2R_{ijt} - \omega_j - \omega \theta_{it})$$

(7)

Then, individuals will choose to leave a neighborhood for some maximal amenity (which drives a maximal rent an individual is willing to trade off before consumption gets undercut). The notation for this value is $R_{ij}$. Now we borrow a convention from labor economics, however; individuals will not choose a neighborhood unless it meets some minimal level of amenity (and, in turn, rent), $R_{ijt}$. This can correspond to some world-clearing price of a home, the amenity value of the last neighborhood the individual was in, or some other measure of an amenity floor. In labor economics, this is the benefit received during unemployment. This is weakly greater than 0.

We tackle a system with $N$ neighborhoods. In a neighborhood with an amenity level $A_{jt}$, if an individual pays rent on housing of $R_{ijt}$, denote the present value of the neighborhood to the resident as $E(A) - E(R)$. For convenience, we will drop the subscripts in this section, where the subscripts may clutter the analysis.

Individuals will move into the neighborhood as long as $E(A) \geq E(R)$ and $R \leq \theta$ for income $\theta$. Further, there exists a $R$, a lowest-possible rent for living in a neighborhood. Residents choose $A$, the lowest amenity value at which they will live in a neighborhood, which determines an $R$ in
(5). The highest rent is given by $\bar{R}$, corresponding to the rent at which utility from consumption will start to decrease to a suboptimal level if rent rises at a fixed amenity value. Structurally, each individual has an unobservable $A$ and a (thus unobservable) reservation rent derived from this amenity level. They also have an unobserved ceiling rent level, $\bar{R}$.

Structure

Now that the individual level decision problem has been established, consider the market for rents. A complete economic model of neighborhood choice will describe not only how individuals choose where to live but the price response of where they live. Consider the following: an individual resident is in state $j(t)$ corresponding to which neighborhood they live in in period $t$. They experience a state-dependent stochastic process that drives the evolution of individual potential rents over time.

$$d \log(R_{ij(t)} = \mu_{j(t)} dt + \sigma_{j(t)} dB(t)$$

(8)

for standard Brownian motion $B(t)$. The residential state is described by neighborhood and potential log rent. A worker can choose to enter a neighborhood at some fixed cost $\psi$. Workers are risk-neutral with discount rate $r > 0$.

To ensure that the problem is well behaved, apply the condition $r > \mu_{j(t)} + \frac{\sigma^2_{j(t)}}{2}$ for each possible state $j(t)$. If this is violated for some $j(t)$, the expected value of staying in this fixed neighborhood will be infinite. I discuss this restriction further in the appendix and prove it is necessary for stability of the rent process.

The resident will remain in a neighborhood while $j(t) = j$ and $R_{ij(t)} < \bar{R}$, but will churn out the first time this condition is violated. In this way, the above stochastic process can be
connected to the underlying individual decision-maker’s problem. The derived condition is strict as long as \(\psi\) is strictly positive. This encompasses an interpretation of voluntary movement, but a reinterpretation where developers have price setting power will demonstrate the profits for the developer earns a profit \(R_{ij} - A\) (\(A\) is the amenity level that is used to determine the reduced-form \(R\), corresponding to a level of utility from out-of-neighborhood amenities) and can attract new residents for a fixed cost \(\psi\) - then, this can also capture forced eviction and developer-driven mobility.

Residents are described by discount rate \(r\), base amenity utility \(A\), and the parameter space governing the stochastic process for their potential rents. In reduced form, we obtain two parameters, \(R\) and \(\bar{R}\). The distributions are entirely arbitrary across the population.

To determine the length of residency, note that the residency will occur once rents breach the lower threshold, \(R\). The log-rent will follow the set stochastic process and the residency spell continues until the log-rent breaches the upper threshold, \(\bar{R}\). The residency spell is thus the first passage time of this Brownian motion with drift, an inverse Gaussian with density \(f\) according to

\[
f(t, \alpha, \beta) = \frac{\beta}{\sqrt{2\pi t^{3/2}}} \exp \left( -\frac{(\alpha t - \beta)^2}{2t} \right)
\]

where \(\alpha = \mu / \sigma\) and \(\beta = (\bar{R} - R) / \sigma\). The former varies over the reals, while the latter is weakly positive. For nonnegative \(\alpha\), the resident almost surely churns out of the neighborhood in question. But, if \(\alpha\) is negative, they have a probability \(e^{2\alpha\beta} < 1\) of leaving the neighborhood and therefore may never do so.

The shape of the inverse Gaussian is fairly flexible. There are specific characteristics that are worth noting. Hazards at move-in \((t = 0)\) are always 0. It achieves a maximum value at a
finite time $t$, then declines to a long-run limit $\alpha^2/2$. The expected duration of residency is $\beta/\alpha$, with a variance of $\beta/\alpha^3$. Asymptotically, the remaining duration of residency approaches $2/\alpha^2$ (note this can be larger or smaller than the expected residency time at move-in depending on $\alpha$, implying both positive and negative duration dependence on possible). The flexibility of possible behaviors associated with a longer duration of residency allows modeling a dynamic selection problem, where developers in new neighborhoods are likely to cater first to individuals with the highest reservation rents, then the next group, and so forth. This analysis will assume that parameters are time-invariant at the individual level.

To summarize thus far, in order to derive measures of how stable cultural districts are, we have modeled first the decision process of the individual renter choosing a neighborhood. We then show how an arbitrary (Brownian motion) rent process can induce a particular resident to choose when to enter or leave a neighborhood based on an inverse Gaussian distribution. The parameters of this distribution serve as a reduced form description of individual preferences for the neighborhood, and can be used to create a mathematical description of neighborhood stability. In particular, the asymptotic time of residency serves as a way to determine how long a resident is likely to stay in a neighborhood once they have moved in, suggestive of a notion of “retention rate” as in the marketing literature.

Integration: From Individual Parameters to Neighborhood Distributions

As with any model of behavior, this model of individuals and the associated rent market problem must be falsifiable given data. If a population is observed where each individual has some fixed structural parameters $(r, \psi, \underline{A}, \bar{R}, \mu, \sigma)$, the model outputs some reduced form parameters $(\alpha, \beta)$. If each individual is only observed for a single residency in the neighborhood,
the model is non-falsifiable and thus non-testable because a single-spell data-point is perfectly explained by assuming that, if $d$ periods are spent in $j$, then $\sigma_j = 0$, $\mu_j = (R - \bar{R})/d$. This would imply $\alpha, \beta \rightarrow \infty$ and $\beta/\alpha \rightarrow d$.

Unfortunately, unlike in the labor literature, where one observes repeat unemployment spells with nonzero probability the likelihood of observed two spells of residency in the same neighborhood infrequent at best. Such individuals become part of a selection problem as they are likely to have stronger attachments to local communities or resources than those who do not repeatedly enter a community.

This is where this paper diverges from the previous labor literature which motivated the structure of the previous section. Rather than estimating individual optimal stopping times and hazard rates out of neighborhoods, we integrate the distribution of optimal stopping times across all individuals in a neighborhood to determine what the neighborhood-level distribution of hazards looks like. We can then estimate neighborhood-level survival functions rather than individual-level hazards. Then, each neighborhood has structural parameters which are each a distribution of individual-level parameters, $(r, \psi, A, \bar{R}, \mu_j, \sigma_j)$ and reduced-form parameters $(\alpha, \beta)$.

While in some cases, this neighborhood-level aggregation proves an inconvenience, in the case of neighborhood-level comparisons of the impact of a specific designation policy, the aggregation is merely an inconvenience to statistical power.

**Testing the Model**

The model’s falsifiability can be described by two conditions. These conditions will determine whether the inverse Gaussian distribution is an adequate descriptor of the individual resident’s propensity to leave a neighborhood. This is important in my results; only datapoints
which pass this test should be analyzed, and if too little of the data passes this test, then my model will not suit this particular problem of determining the effect of a DCD.

Now, say we have a sample of residents of size $M$ from neighborhood $j$ whose durations correspond to an observed vector $\vec{t}$ of dimension $M$. Each has some set of structural parameters as above. Now, their reduced-form parameters are drawn from a joint neighborhood-level distribution $g$. Then, the density for time of residency (alternatively, the residency “spell”, to mirror the labor literature language):

$$\phi(\vec{t}) = \int \int f(\vec{t}, \vec{\alpha}, \vec{\beta}) g(\alpha, \beta) d\alpha d\beta$$

(10)

This differs markedly from the approach of labor economists whose work originates this model. Rather than identifying these individual-level parameters with any fineness, this approach acknowledges this is not possible with repeated renter data and pools across observations.

Allow $\phi^{(i)}$ to denote the derivative of $\phi$ with respect to $t_i$. From the functional form of the residency spell duration identified in (9), these derivatives satisfy:

$$\phi^{(i)} = \int \int \left( \frac{\beta^2}{2t_i^2} - \frac{3}{2t_i} - \frac{\alpha^2}{2} \right) f(\vec{t}, \vec{\alpha}, \vec{\beta}) g(\alpha, \beta) d\alpha d\beta$$

(11)

Equivalently,

$$\frac{2t_i^2 \phi^{(i)}}{\phi(\vec{t})} = E(\beta^2 | t_i) - 3t_i + E(\alpha^2 | t_i)$$

(12)

There is a degree of precision in the previous incarnation of this model that is lost in the aggregation process. Rather than having the ability to estimate a distribution for $\alpha$ and $\beta$ for the individual, this model must sacrifice the individual-level inference. This is a quirk of residential data; the estimation is unstable otherwise. However, the conditions for model testability are the
same. Both expectations in (12) must be nonnegative. For a constant hazard rate of residency $h$, so that the density of completed spells is $\phi(t) = h^2 e^{-ht} \sum t_i$, these will be

$$E(\beta^2 | t_i) = 3 \frac{\prod t_i}{\sum t_i}$$

and

$$E(\alpha^2 | t_i) = 2h - \frac{3}{\sum t_i}$$

(13)

Then, $E(\alpha^2 | t_i)$ violates nonnegativity when the sum of duration times across individuals, $\sum t_i < 3/(2h)$ or one and a half times the mean duration. Then, constant hazard rates cannot be generated for just any set of residency spells.

If the constant hazard is now unfixed, with a distribution in the population $G$ then the density of these completed spells is $\phi(t) = \int h^2 e^{-ht} G(t) dt$. Then, $E(\alpha^2 | t_i)$ is negative if the ratio of the third moment of $h$ to the second moment is positive. Each moment of the joint distribution of $(\alpha^2, \beta^2)$ can be obtained through the kth partial derivative. For the exact procedure, refer to (Alvarez et al., 2015).

**Nonparametric Identification**

Another consideration of the model demonstrates how to identify the distribution of the reduced form parameters $\alpha$ and $\beta$. Once the data is tested, the model must be able to solve for the exact distribution of theses parameters. Otherwise, the model has been a mathematical hattrick which provides minimal interpretable results.

Using the aggregated densities across individuals observed in the neighborhood, I nonparametrically identify the joint distribution of $(\alpha^2, \beta^2)$ across individuals. Identification is equivalent to identification of the joint distribution of $(|\alpha|, \beta)$. The joint distribution of $(\alpha^2, \beta^2)$ can be nonparametrically identified by comparing results across multiple neighborhoods with
residents who experience a fixed vector $\vec{t}$ (with at least 2 residents being observed in each studied neighborhood). One can compute the joint distribution $g^* (\alpha^2, \beta^2)$ according to

$$\psi (\alpha^2, \beta^2, \vec{t}) = \frac{f(\vec{t}, \alpha, \beta) \hat{g}(\alpha^2, \beta^2)}{\int \int f(\vec{t}, \alpha, \beta) \hat{g}(\alpha^2, \beta^2) d\alpha' d\beta'}$$  \hspace{1cm} (14)

This could be inverted to solve for this joint distribution $g^* (\alpha^2, \beta^2)$. This test should not depend on the sample $\vec{t}$ used to derive the distribution $g^* (\alpha^2, \beta^2)$, a fact which can be used to test sensitivity of the model.

Alternatively, in a more applicable estimation process which mirrors a Newtonian procedure, one can assume a starting distribution of types $g(\alpha, \beta)$. For each of these types, the model provides a density $f(\vec{t}, \alpha, \beta)$ which produces the density of actual durations $\phi(\vec{t})$. In a setting with finitely many types $N = \text{card}\{\alpha, \beta\}$ and sets of durations $T = \text{card}\{\vec{t}\}$, this is the linear system $\phi = F \cdot g$ for likelihood matrix $F : F \in \mathbb{M}^{T \times N}$ and $g \in \mathbb{R}^N$ a vector which indicates the share of each type in the population. This gives $\phi$, the vector of share of each duration in the population. As long as $F$ is invertible, the model is identified. If we assume momentarily $N = T$ (as the other cases will yield many or no solutions), identification is a matter of the rows of the likelihood matrix being linearly independent. As the density of realized durations for one neighborhood is not a linear combination of the density of realized durations for others, the model is identified. In other words, it is unlikely that direct linear dependence is to arise in a highly randomized setting with a large variety of duration types.

**Decomposition of Changes in the Hazard Rate**

As a final feature of the model, I argue it is helpful to separate how much of the neighborhood-level density of hazards (e.g., the distribution of times at which residents are
leaving neighborhoods) can be described as (1) an artifact of heterogeneity, meaning differences are due to differences in some latent aspect of a person that is not described in the model (e.g., individual characteristics rather than neighborhood ones) or (2) an artifact of the structure of the model itself - attributed to the structure of amenity distribution across neighborhoods. Define 

\[ F(t, \alpha, \beta) = \int_0^t f(t', \alpha, \beta) dt' \]

as the fraction of type \((\alpha, \beta)\) residents with spells longer than \(t\) periods.

Then, the distribution of types among those same workers is:

\[
\tilde{g}(\alpha, \beta | t) = \frac{1 - F(t, \alpha, \beta) g(s \alpha, \beta)}{\int \int (1 - F(t, \alpha', \beta') g(s \alpha', \beta') d\alpha' d\beta') s}
\]  

(15)

The density of residual move-out durations for a given type conditional on the unemployment spell lasting at least \(t\) periods,

\[
\tilde{f}(\tau, \alpha, \beta | t) = \frac{f(t + \tau, \alpha, \beta | t)}{1 - F(\alpha, \beta)}
\]

(16)

Finally, then, the density of residual residency duration lasting at least \(t\) periods is

\[
f^*(\tau | t) = \int \int \tilde{f}(\tau, \alpha, \beta | t) \tilde{g}(\alpha, \beta | t) d\alpha d\beta
\]

(17)

The expectation of this density will be denoted \(D^*(t)\). This can be computed directly once the joint distribution of the reduced form \((\alpha, \beta)\) are identified. Now, if we want to decompose the change in expected durations into the effects of heterogeneity and duration dependence. The contributions of heterogeneity and duration dependence are then:

\[
D^*(t) - D^*(0) = \int \int \int_0^{\infty} \frac{d}{ds} (\tilde{f}(\tau, \alpha, \beta | t) \tilde{g}(\alpha, \beta | t)) ds d\alpha d\beta
\]

\[= \int \int \int_0^{\infty} \frac{d}{ds} \tilde{f}(\tau, \alpha, \beta | t) \tilde{g}(\alpha, \beta | t) + \tilde{f}(\tau, \alpha, \beta | t) \frac{d}{ds} \tilde{g}(\alpha, \beta | t) ds d\alpha d\beta
\]

\[=: D^h(t) + D^d(t)
\]
A similar decomposition can also be constructed on the hazard rates. The hazard rate $h(t) = \int \int h(t, \alpha, \beta) \tilde{g}(\alpha, \beta | t) d\alpha d\beta$. This decomposition gives a term for structural duration dependence and a term for heterogeneity, respectively, as:

$$h^{s}(t) = \int \int \int_{0}^{t} \frac{d}{ds} h(s, \alpha, \beta) \tilde{g}(\alpha, \beta | s) d\alpha d\beta$$  \hspace{1cm} (18)$$

$$h^{h}(t) = \int \int \int_{0}^{t} h(s, \alpha, \beta) \frac{d}{ds} \tilde{g}(\alpha, \beta | s) d\alpha d\beta$$  \hspace{1cm} (19)$$

Figure 1: Descriptive histograms for resident move-out times in the entire universe of the dataset. The right panel shows an overall view, while the left compares the hazard rate of residents before and during the recession of 2009.

Data

Tenureship data is obtained from the RentBureau dataset. This Experian-produced product contains payment and residency records for 13 years (1997-2010) for an apartment company’s residents across many zip codes. The data includes rent amounts (which allows testing the underlying rent-based derivation of the individual preference variables) and records of when a resident enters and exits their apartment. The data also presents a detailed account of payment, non-payment, and collections by month through the lease. The dataset excludes subletters and Airbnbs. We consider these both as part of the prevailing renters’ tenure in the current model, acknowledging that this excludes a set of residents from this analysis. Residents
may move within zip codes. Such observations are flagged and removed. The actual population set of city residents therefore will have preferences for shorter term stays in the city than the data suggests.\(^{31}\)

I provide a dataset-level histogram of survivals in the Figure 2. Note that survivals are censored at the two-year mark (24 months on the x-axis), so we must develop a right-censoring correction. The histogram suggests that a large number of individuals churn out at the two-year mark, but the truth is that part of this large bar contains entries which have resided in their apartments for \textit{at least} two years, not exactly two years.

The second subplot compares survivals in the period of the Great Recession, defined to begin in September 2008, to those which occurred beforehand. Off-hand, we generally see the shapes as relatively similar, simply with fewer observations during the recession. I also report the overall joint distribution of pairs of tenures across neighborhoods in a surface plot in (2). Some key observations:

![Figure 2: Joint distribution of pairs of tenure lengths (length of residency) within zip codes.](image)

first, note the joint distribution is generally convex, which reflects the declining hazard rate (likelihood to move out) at longer durations (ignoring the peak at 24 due to right-censoring). Second, the joint density is noisy. This does not appear to be primarily due to sampling variation.
Rather, there is clear persistence of a yearly tenure-ship drop off as this is the likely duration of the average rental lease. The extremely high joint frequencies of multiples of 12 suggests that this is being replicated in the data. This is very much an artifact of the data, and it can be dealt with in two ways. The first; it can be smoothed out using a spectral analysis technique. The second; it can be allowed to remain as is and the model will simply be judged on the surface plot shown for measures of likelihood. Both methods are discussed in the calibration section.

The dataset’s breadth includes several DCD zip codes which are used in this study:

1. Atlanta: Gayborhood, 30318 and Chinatown 30341
2. San Diego: Gayborhood 92103 and Chinatown 92104
3. Los Angeles: Gayborhood 90012 and Koreatown 90005 and 90006
4. New Orleans: The Garden District 70119
5. San Francisco: The Castro, 9411
6. New York City: Riverside District, 10069
7. Tampa: River Arts District, 33602

![Figure 3: Comparisons of tenure and log-rent distributions by allocated control (FALSE) and treatment (TRUE) groups. The right compares length of stay, while the left compares log-rents.](image)

We use these cities as experimental areas, with the DCDs as treated zip codes. There are 2391 zip codes in the total dataset, from which we extract 123 zip codes. Of these, taking proximity effects into account, the zips listed above and any zip codes within a 2-mile catchment are
assigned as treatment. Each zip code has an average of 1340 different residency records. To summarize, I present a simple comparison of log rents and tenure distributions across treatment and control groups in Figure (3).

Generally, there are some initially visible differences across designations in both the tenureship distribution and the rent price distribution. Rent prices appear to be less skewed slightly, in treated districts. The tenure time of a resident appears to follow a similar decrease in skewness. Note that this runs countervailing to an a priori presumption of decreased heterogeneity in treated districts due to cultural similarity. In particular, a homogeneous district is likely to have more unimodal, concentrated distributions. This is clearly not the case. Further exploration and calibration may help to elucidate why this might be happening. With the data in mind, we move to calibrating the model.

Calibration

Testing Compliance

To look at whether the model is able to be fit onto the data, and further whether it is rejected by the data for lack of falsifiability, I conduct a two-stage compliance procedure. I test two possible smoothing algorithms on the data and discuss why I end up rejecting both methods in B. Then, I conduct the tests described in 3.4 to obtain a complete picture of what subsets of the data can be fit with the stochastic model.

Optimization Procedure

The data displays heavy nonlinearities and seasonality, suggesting it is not clean enough for a simple pass through an off-the-shelf convex optimization procedure. I detail the process of
finding a minimum-distance estimator here. First, argue that for a neighborhood with unspecified numbers of residency records, we can obtain a robust estimator of the aggregate density function of \( \alpha_j \) and \( \beta_j \), \( g \), by estimating \( \hat{g} \) on a subset of the possible residency data-points. Rather than using all of the data-points in each neighborhood then, we sample 30 timings from each neighborhood (the cutoff from the common rule derived from the central limit theorem) and estimate \( g^* \) in this space. Note that the density at this point should be relatively stable, and that we should expect to estimate the same density save some sampling variation for various subsets of 30 residency spells. Then, this mirrors the practice of truncating variable-timespell panel data before analysis to a common timespell, but with a sample size across observation areas.

The data exists as some distribution \( \phi(t) \in \mathcal{D}(\mathbb{R}^{30}) \). The distribution of our parameters, \( g \in \mathcal{D}(\mathbb{R}^{30}) \) can be constructed as:

\[
\phi(t) = \int \int \prod_{i=1}^{30} f(t_i, \alpha, \beta) g(\alpha, \beta) d\alpha d\beta
\]

for every such \( t \in \mathbb{R}^{30} \). This is an inner product, so \( \phi = Fg \) for positive, linear \( F \). This is a likelihood function; it is essentially an \( M \times N \) positive matrix whose columns add to 1. Each entry, \( F_{ij} \) is the probability \( Pr\{t_1 \in (t_1(i), t_1(i) + dt), t_2 \in (t_2(i), t_2(i) + dt) | (\alpha, \beta) = (\alpha(j), \beta(j))\} \).

Then, the objective function of the quadratic optimization problem which falls out of the procedure is

\[
\min_{g \in \Delta^N} \|Fg - \phi\|
\]

where \( \Delta^N \) is the distribution of possible types \((\alpha, \beta)\).
To compute $g$ consider the following pre-processing measures:

1. Symmetrize the likelihood so that $\phi(t_1, t_2)$ is the average of the density of $(t_1, t_2)$ and $(t_2, t_1)$.
2. The grid for $\alpha$ is entirely positive as we can only identify the absolute value of $\alpha$.
3. Calculate the likelihood at the final bin as the right-censored likelihood; the likelihood of a residency time greater than or equal to 24 periods.
4. Relax the problem so that we do not require $g$ to have only positive elements and the constraint that the elements sum to one. Scale positive elements of $g$ and have them sum to 1 so that the Karush-Kuhn-Tucker conditions will be satisfied after an iteration.
5. Throw away pairs of $(\alpha, \beta)$ with density below 1 basis point.

We do not require some of the other algorithmic constraints in Alvarez et al. (2015) because of the right-censoring of the data (25 is a relatively low dimensionality for number of periods). The above, particularly symmetrization, are required to avoid some of the pitfalls of the noisiness of data.

**Results and Discussion**

**Summary of Model**

I compile summary statistics of model calibration by treatment and control in Tables 2 and 1, respectively. These treatments are pooled at the label level; in other words, the model does not train on yearly or by-locality groups in these initial tables. I do this for statistical power. At first glance, the estimates do not appear to differ heavily. Generally, there is not excessive heterogeneity; the standard error of $\alpha$ and $\beta$ is not particularly high compared to the mean in any of the groups. However, there is a great deal of heterogeneity in the mean stay itself; standard
errors are generally twice the mean. Note that the mean stay in the treated areas is about half a year longer in raw value.

The cross-sectional variance of the mean duration is also not variable in aggregate; it is around 40 in each group, with the standard error being around 16 (18 amongst the treatment group). The proportion of the variation in realized durations due to individual variation is then around 40 percent (43 percent in the treatment group). The asymptotic duration of a residency spell is about 7.5 months in control areas and 8.2 months in the treated areas. Then, if the individual has lived in an area for a significant amount of time already, there is a negative

<table>
<thead>
<tr>
<th>$E(\alpha)$</th>
<th>$E(\beta)$</th>
<th>$E(\beta / \alpha)$</th>
<th>$Std(\beta / \alpha)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.738</td>
<td>5.832</td>
<td>8.531</td>
<td>16.161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E(2/\alpha^2)$</th>
<th>$E(\beta / \alpha^3)$</th>
<th>$Std(\alpha)$</th>
<th>$Std(\beta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.566</td>
<td>39.664</td>
<td>0.3412</td>
<td>2.153</td>
</tr>
</tbody>
</table>

Table 1: Means of several key statistics across control regions, across all time periods.

<table>
<thead>
<tr>
<th>$E(\alpha)$</th>
<th>$E(\beta)$</th>
<th>$E(\beta / \alpha)$</th>
<th>$Std(\beta / \alpha)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.672</td>
<td>5.921</td>
<td>9.004</td>
<td>18.392</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E(2/\alpha^2)$</th>
<th>$E(\beta / \alpha^3)$</th>
<th>$Std(\alpha)$</th>
<th>$Std(\beta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.173</td>
<td>44.108</td>
<td>0.343</td>
<td>2.180</td>
</tr>
</tbody>
</table>

Table 2: Means of several key statistics across treatment areas, across all time periods.

duration dependence (e.g., they are expected to stay longer the longer they have already stayed). This timeframe is one-third of the total timeframe observed for each record, suggesting the duration dependence effect is quite strong in our data.

I show the fit of the mean type across all neighborhoods on the durations histogram for all observations in New Orleans in Figure 4. The fit is not the best here (nor should it necessarily
be), but notice that the symmetrization has caused the mode to ignore some of the nonstationary behavior around the 12th month of residency. Then, the convex optimization procedure has produced a principled, not overfit, model.

A sample likelihood surface is shown in (5). Clearly, the likelihoods for individual values has a fairly regular shape after the pre-processing above is conducted. This suggests the calibration is comfortable. In an assessment of the goodness of fit, the city-level fits had an average chi-square p-value of 0.403, implying we can comfortably accept the model as a good fit. The variance on this statistic across neighborhoods is 0.199, suggesting that we are generally comfortably within model acceptance standards across individual localities. The asymptotic duration shows differential responses to rent across assignment as well.

![Figure 4: Fit of mean probability distribution (orange) on the New Orleans aggregated data (blue).](image)

![Figure 5: Surface plot of a sample likelihood.](image)
Designated cultural districts tend to have a higher estimated asymptotic duration in wealthier neighborhoods as compared to control groups, suggesting in particular that treatment groups are likely to be income-sensitive. This suggests one of two things is occurring: (1) areas which are of a higher income before designation are more likely to be designated or (2) designated cultural districts are more likely to attract high income residents upon treatment. While we do not have data on this, the inclusion of immigrant communities like Chinatowns and Koreatowns and a Little Puerto Rico seems to suggest that (2) is the more likely explanation, as such communities are generally more likely to have settled with a lower income. Higher-rent treatment areas also generally demonstrate lower proclivity to have heterogeneity, suggesting areas homogenize as these new arrivals enter the neighborhood.

**Variance Decomposition**

I now conduct the exercise described in the theory section which decomposes residual duration of residency into a component explained by the heterogeneity in residents and one explained by duration dependence (negative in this case). Again, I start by comparing these in...
the overall dataset for all periods. Figure (7) compares these. Generally, heterogeneity appears to explain a larger component of the asymptotic variance of residency across neighborhoods.

I map the differences across treatment and control groups by comparing the proportion of variance explained by duration dependence as asymptotic duration dependence increases. These are simple linear estimators. See that the treatment groups tend to have a larger positive duration dependence after the first year, implying that once the first year has passed, the neighborhood is estimated to extract significant more loyalty effects from residents. This also implies that most of the heterogeneity “shakes out” during the first year – as this is when the proportion of variance attributed to heterogeneity is at its highest point – suggesting a compositional sorting of unlike types out of the neighborhood after the first year.

*Yearly Cohorts*

In the next section, I discuss isolating comparisons to same-year and same-city treatment-control groups, whereas the statistics reported in this section are all pooled. This was done because the results in this section are able to share statistical strength across experimental groups, which overcomes the smaller sample size of the treatment group relative to the controls.
The next section is certainly more sample-dependent, so I use bootstrapped variances and standard errors while reporting.

I report the summary statistics of these yearly city-level cohorts in a style mimicking the tables above here in Table 3.

The estimated values of the raw pools are fairly similar, indicating our pooled analysis was likely a decent approximation of the true underlying patterns. In particular, proportion of variance attributed to heterogeneity remains around 40% and the standard errors are similar on most counts.

![Figure 8: Increase in variation attributed to duration dependence with increasing estimated asymptotic duration in treatment and control groups](image)

Table 3: Means when calculated in year and city-specific pools.
Neighborhood Stability: Causality

Thus far, much of this paper has been a replication of the previous work of Alvarez, Borovickova, and Shimer. The novel contribution of the previous sections was simply a matter of building a context for the application of their model within the experiment of residency spells and neighborhood choice. I now contribute the second novel piece of this paper, the application to measuring the effect of a designated cultural district’s anchoring amenity on the hazard rate out of a neighborhood. I will use this to test whether cultural districts have a significant impact on out-migration during an overall period of increased outmigration, the Great Recession. While the previous sections present enough nuance to develop some conclusions about potential impact of DCDs on neighborhood stability, a more causal argument is possible given the Recession as an evenly applied exogenous shock (though perhaps not evenly deep across its application, as some cities and zips were hit harder than others - more on this momentarily).

Defining “Neighborhood Stability.”

The distribution \(g(\alpha, \beta)\) encapsulates and expresses in reduced form the distributions of neighborhood-aggregated individual parameters, \((r, \psi, A, \bar{R}, \mu, \sigma)\). The model for the anchoring effects are simple. Then, we can use measures from our reduced form analysis as individual-level terms. In particular, we can define a notion of neighborhood stability if a neighborhood tends to attract individuals for a longer period of time (conditioned on them staying for some mean duration - this is the definition of the asymptotic duration quantity reported in this paper). Note that by using this definition of asymptotic duration, we can be flexible on a cutoff by neighborhood; if a neighborhood has a lower baseline mean duration, it can still have a high asymptotic duration for individuals who survive that mean duration. This is a good measure of stability because it emphasizes two points: (1) local tunability, and (2) robust estimation of
heterogeneity across individuals. However, it faces a critical flaw; there is no estimator for people who wish to leave a neighborhood but cannot due to the fixed costs associated with moving. Put more succinctly, it does not capture financial entrapment due to structural differences in access to credit across space.\textsuperscript{32}

We can use the asymptotic duration as a measure of neighborhood stability with an important caveat. Asymptotic duration of a residency spell may be significantly determined by wealth of an individual, as wealthier residents are less likely to be priced out. This poses a problem as to whether the neighborhood’s amenities are a causal force for the outcome of stability or whether the individuals they attract are; attracting more stable individuals might be less attractive as an indicator of neighborhood stability as it is more subjective to time-varying tastes. Then, it is difficult from our data alone to identify the source of stable neighborhoods as a demand-side or supply-side push, but we can identify stability itself with some degree of comfort. There is a second problem with identifying treatment effect as well, namely that treatment alone may capture the effect of differences in provision of time-varying amenities across neighborhoods. To address this, the use of an aggregated rent term is advised. However, rents can be seen as endogenous to treatment. I discuss this in the modeling process.

Arriving at a Causal Model.

Design a simple model of difference-in-differences which assesses how well treated\textsuperscript{(DCD)} neighborhoods weather the impulse shock of the Great Recession

\[
\frac{2}{\alpha_{jk}^2} = c(j) + \tau_k + \delta_0 DCD_j + \delta_1 R_k + \delta_2 R_k DCD_j + \gamma P_j + \varepsilon
\]  

(22)

where \( c(j) \) is a city-level fixed effect that captures spatial differences in tenure-ship patterns, and \( \tau_k \) are differences caused by time dynamics which do not get captured in evolving individual
traits. $R$ is the variable representing the onset of the recession. $k$ is a cohort-level subscript, to account for changing distributions over time. Cohorts are defined as individuals who move into a zip code in the same year for the time being. I capture expectation of local amenities and local conditions through the logarithmic mean of rent prices, $P_j$ for the time being; though it is an imperfect measure far and away, it is highly correlated with most local amenities. In this way, it forms a natural fixed effect measure for zip-code level variation. Then, $\delta_2$ serves as a fully identified difference-in-differences measure after local fixed effects are taken into account. I argue this differencing is necessary primarily due to the use of rent alone as the controlling covariate. I report the coefficients in table 5, as well as some associated measures of the model’s appropriateness and goodness of fit. The difference-in- differences estimator is not significant in this regression; the treatment with designation does not significantly protect individuals from the effects of the recession on their asymptotic duration. Interestingly, the recession is also not particularly significant in this regression. However, the coefficient on the treatment effect is significant at the 10% level. This is interesting in the sense that it reinforces the significance of the treatment, but it does not tell us that the treatment itself is a causal response in the face of a recession.

I try two more models which are somewhat nested by this model. The first removes the time dummies to test whether they are too redundant given inclusion of the recession variable. The condition number of the covariance matrix for the first model is particularly high, suggesting the time dummies (largely insignificant, admittedly) are too collinear with the recession variable given the sample size. The second variation on the model also eschews time dummies. This version further relaxes the assumptions of parallel slopes (already partially satisfied by the inclusion of a rent-correlated fixed effect). Heterogeneity in the neighborhood might be an
important source of information as neighborhoods which are less uniform in their responses to shocks will have different outcomes post-recession.

Looking at the model reports in Appendix B suggests the most viable model from a log-likelihood perspective is that without time controls but with a control for heterogeneity. This model suggests that in fact, the DD estimator and the estimator for the treatment area both amount to an insignificant effect. The heterogeneity, however, dominates this regression. This suggests that an unconditional analysis of the asymptotic residency duration relies primarily on the heterogeneity of the locality.

I present one final causal analysis, in which I conduct a two-stage procedure. In stage 1, I regress the asymptotic duration on the estimated mean $E(\frac{\beta}{\alpha})$ of residency spell types in the zip code within cohort $k$:

$$E(2/\alpha^2) = \zeta E(\frac{\beta}{\alpha}) + \kappa$$  \hspace{1cm} (23)

In this case, we know the theoretical slope of this regression, but variation $\kappa$ can be entirely attributed to non-individual effects on the asymptotic duration. In other words, conditional on the type of the individual, I extract the component of asymptotic duration whose variance is not explained by any measure of individual preferences for neighborhood features. This measure $\kappa$ I now regress against the same second-stage linear model as performed best in the previous analysis:

$$\kappa = c(j) + \delta_0 DCD_j + \delta_1 R_k + \delta_2 R_k DCD_j + \gamma_1 F_j + \gamma_2 h^k_j + \epsilon$$  \hspace{1cm} (24)

By this, I suggest that these uncorrelated components of the asymptotic duration are determined by zip-level added “loyalties” derived from homogeneity, local culture, and the city itself. Argue that this component is no longer endogenously determined with rents. While rents may affect the mean type of individual to visit a neighborhood, they are not determined
simultaneously with the residual duration of an individual type. This model is not comparable off of log-likelihood to previous models as it is not nested. However, notice that here the recession now has a direction that intuitively makes sense - it decreased the de-individualized component of asymptotic duration by 0.25 months on average, significant at 10%. The DD estimator is now nearly significant at 10%, with an increased expected asymptotic duration of 0.81 months. This suggests that the effect of a designated cultural district must be de-individuated. In particular, the component of variation corresponding to an individual’s type when conditioned strengthens the linear model overall, including significance of local fixed effects.

The interpretation of this final model is that individuals have some baseline taste for amenities and rents, $\alpha$ which determines their individual type as in our original model. To isolate how a local neighborhood affects individual decision-making on the margin, we must control for their individual tastes for that neighborhood in the first place, represented by this calculated type, and look at how the residual varies with observables.

*Price Effects*

A counterfactual that could prove not only that the model has the intended result but also help analyze a policy situation is to consider how much of price is determined by individual tenureship characteristics. In particular, we can look at differences in rents between treated and non-treated areas and ask how much of the variation in this difference is potentially due to individual sorting as opposed to the anchoring amenity itself. Some key questions include:

1. Are treated neighborhoods attracting a different type of individual, or are they able to keep the same type of individual better?
2. Are treated neighborhoods significantly more or less expensive than their non-treated counterparts?

3. How much of variation in price in treated neighborhoods can be explained by individual types of people?

Here we again treat rents as endogenous with conditional expectations of key individual-level variables. A simple comparison of mean type across treated and non-treated neighborhoods answers the first counterfactual question: individuals who are attracted to treated neighborhoods are generally not different in type, as 

\[ E(\beta/\alpha|DCD) - E(\beta/\alpha|CONTROL) \]

is not significantly different from zero. This is actually rather encouraging; it tells us that both neighborhoods have renters with similar rent response to changes in price, but the marginal effects of community we derived above that explain the residual duration time of an individual are better explained by treatment effects. This suggests that the same type of person is more likely to stay longer in a treated area than an un-treated one.

Treated areas would seem likely to have a larger rent given the presence of a significant anchoring amenity. A quick comparison of means suggests rents are 27% higher in the treated sample, with significance at 99% confidence. A quick difference-in-differences that simply looks at treatments before and after the recessionary shock indicates there are little grounds for causality, however. Then, this may be due to sampling and we do not have a strong case for item (2) above.

The final question is perhaps the most interesting. The counterfactual poses the question as to the elasticity of this price volatility to the local type distribution of individuals. In other words, how much of local type is responsible for determining prices and how much of price is derived from local amenities and the “basket of goods” provided by the neighborhood?
In theory, this component of variation should be nonzero as suppliers should be elastic to mover preferences. The suppliers are likely to have some prior about reservation rents and reservation amenity levels and can use this to create the distribution of offered prices, $P_j$. This distribution is weakly centered in the range $\min_{i\in I_j} R_i$ to $\max_{i\in I_j} R_i$ - the maximal and minimal rent tolerances of a neighborhood’s current residents. Put simply, the distribution of houses supplied should be concentrated in the levels of rent with nonzero demand. As the region attracts individuals of heterogeneous types, we therefore expect the price distribution to become more variable. This supplier elasticity can be an important variable for policy decision-making which allows for assessment of supplier responsiveness to consumers. Such elasticities have been used to assess consumer protection and predatory lending, as potential examples of applications.

There are two ways to test transmission of individual types to prices. One is a simple variance decomposition of $\text{Var}(P_j)$. Another considers the possibility that suppliers uses previous-period types to project current-period demand. Consider the simple measure $\|\Delta(P_j)\|_2$ which describes the magnitude of the change in rents over the course of a year. This is a simple measure of price volatility that eliminates the need to predict positive or negative price shifts, focusing just on the scale of the shift in prices. We expect that this second measure is going to capture a more realistic picture of supplier behavior than the former, as instantaneous variance should not be predicted by instantaneous types. Consider first the model for a simple difference in price changes:

$$
\Delta(\log(P_j))_t = \mu_1 E_i(\beta/\alpha)_{t-1} + \mu_2 E_i(2/\alpha^2)_{t-1} + \mu_3 E_i(h^b(r))_{t-1} + c(j) + R_t + \epsilon
$$

The model regresses price changes on the mean type of individual in the previous period, adding a city-level fixed effect and a recession control. The coefficients are reported in table (4).
The model suggests that about 8% of the variation in log-price changes are accounted for by the variation in individual types. The table suggests, further, that as the expected mean time increases, the rent variability is decreasing, yet as asymptotic duration increases, rent variability increases. This seems like a contradiction until one remembers that both are correlated with $E(\alpha)$.

The analysis suggests that $\alpha$ tends to drive up rent variability - this makes sense, as alpha for the individual is driven by the ratio of the drift of their rent stochastic process over its variance. This regression becomes a proper test of the model behavior, and the model passes with appropriate significance that we can feel comfortable with the stochastic underlying process that defines the reduced-form parameters. Further, the model tells us that the proportion of variance corresponding to heterogeneity, when rising by a single percentage point, can drive up rent variability by 2.5%. This is a significant impact that suggests that suppliers are highly responsive to variation in individual types in the previous years as a proxy for rent-setting in the current year with an elasticity of 2.49.

|                | coef  | std err | t    | P>|t| | [0.025 | 0.975 |
|----------------|-------|---------|------|------|--------|--------|
| Intercept      | 0.1933| 0.058   | 3.314| 0.001| 0.079  | 0.308  |
| C(San Diego)   | 0.0013| 0.100   | 0.013| 0.990| -0.196 | 0.199  |
| C(Los Angeles) | -0.0080| 0.036  | -0.226| 0.821| -0.078 | 0.062  |
| C(New Orleans) | 0.2091| 0.337   | 0.621| 0.535| -0.452 | 0.871  |
| C(New York)    | 0.0293| 0.049   | 0.595| 0.552| -0.067 | 0.126  |
| C(Tampa)       | 0.0778| 0.040   | 1.930| 0.054| -0.001 | 0.157  |
| $R_i$          | -0.0739| 0.030  | -2.422| 0.016| -0.134 | -0.014 |
| $\mu_1$       | -0.0155| 0.006  | -2.446| 0.015| -0.028 | -0.003 |
| $\mu_2$       | 0.0196| 0.006   | 3.369| 0.001| 0.008  | 0.031  |
| $\mu_3$       | 0.4153| 0.123   | 3.384| 0.001| 0.174  | 0.656  |

Table 4: Model estimated for (25).

This counterfactual analysis thus provides several additional insights: first, that the interpretation of the stochastic process is reflected in the data even though rent data itself is not used in the calibration of the model. Second, that the individual type of persons across
neighborhoods is not different just because of designation. Yet third, that a significant piece of rents across neighborhoods is determined by the distribution of individual types in that neighborhood.

Conclusion

The result of the key regression in (24) suggests that designation of districts as a policy choice may significantly increase the lifespan of a neighborhood in the face of a recession, if only marginally in effect size. The cultivation of this type of loyalty can have significant effects on city budgets in moments of crisis by increasing the likelihood of having a stable tax base, and can further drive longer-lasting city growth. Yet, this analysis suggests a complexity to the narrative of the designated cultural district. DCDs make neighborhoods more robust to productivity shocks, in a manner which is masked when they are considered in aggregate (see Tables (2) and (1)). Put simply, micro-data and deeper local insights reveal variation that is masked in the overall distribution.

The income dynamics of designation are yet unexplored, however. It is clear that designated cultural districts upon designation tend to drive up rents (simple correlation analysis suggests a bump of 27% over neighbors) and attract potentially wealthier individuals, which in turn helps drive increased asymptotic duration in such localities. However, it is unclear whether displacement from DCDs after designation may actually drive out the highest-loyalty residents; the counterfactual result as to who drives the response to recession cited above is difficult to test without more complete records. If the creators of cultural capital have left post-designation, perhaps these neighborhoods would have been even more robust to recession - or less so. Thus, a causal tension still persists.
An important result of this paper is further the application of the labor economics paradigm to the neighborhood choice literature, or a component of the neighborhood choice process. The model is robustly corroborated by rental data as discovered in the section discussing price effects. This test provides useful elasticity estimation that can suggest how well landlords respond to individual renters. By studying when people leave neighborhoods, this analysis suggests, economists can learn a great deal about where neighborhoods fail their residents. Further, the analysis is conducted on neighborhood-level histograms of micro-data and is still able to garner remarkable insight on distributions of important sufficient statistics. This approach therefore suggests that the model might be further empowered with more granular and descriptive micro-data to get useful individual-level results as well. The additional power can be used to test further place-based policy effects and questions impacting stability analysis.
References


S. Becker and E. Jahn. Labor market signaling and unemployment duration: Evidence from Germany. 2015.


Appendix

A Expanding on the Individual-Level Problem

With the given stochastic process for rents, the present value of rents will satisfy the Bellman-Jacobi-Hamilton equations:

\[ rE_j(R) = \exp(R) + \mu_j E'_j(R) + \frac{\sigma_j^2}{2} E''_j(R) \quad \forall R \quad (26) \]

The solution of this set of parallel equations across all neighborhoods, \( r \).

\[ rE_j(R) = \frac{\exp(R)}{r - \mu_j - \sigma_j^2/2} + \sum_{k=1}^{N} u_k \exp(\lambda_{jk} R) \quad (27) \]

where the poles have opposite sign and are the roots of the \( N \)-dimensional equation \( r = \lambda_j (\mu_j + \lambda_j \sigma_j^2/2) \). A solution to this equation should also satisfy the equations:

\[
\begin{align*}
R_{ij} &< \bar{R}_{ij} \\
E_j(R) &< E_k(R) \quad \forall k \neq j \quad (28) \\
E_j(R) &< E_k(R) + \psi \quad \forall k \neq j \quad (29) \\
E'_j(R) &< E'_k(R) \quad \forall k \neq j \quad (30) \\
E'_j(R) &< E'_k(R) \quad \forall k \neq j \quad (31) \\
E''_j(R) &< E''_k(R) \quad \forall k \neq j \quad (32) \\
E''_j(R) &< E''_k(R) \quad \forall k \neq j \quad (33)
\end{align*}
\]

The conditions require value functions to be continuous and differentiable at the boundaries. Finally, we have two conditions which ensure no bubble around employment or unemployment:
These equations ensure that for an arbitrarily low rent, the value functions all converge to the value of moving to another neighborhood and if the rent increases without bound then the value function converges to the value of living in neighborhood \( j \).

The no-bubble conditions (34) imply that all \( k : k \neq j \) in (27), \( u_k = 0 \). Otherwise, the expected value of neighborhood \( j \) will diverge relative to other neighborhoods as rent grows asymptotically large positively or negatively. Then, the value functions will be, abusing notation slightly:

\[
r E_{j(t)}(R) = \frac{\exp(R)}{r - \mu_j + \frac{\sigma^2_j}{2}} + u \exp(\lambda_j R)
\]

with

\[
\lambda_j = \frac{-\mu_j - \sqrt{\mu^2_j + 2r \sigma^2_j}}{\sigma^2_j}
\]

Then, we end up with \( 4N - 4 \) equations in the unknowns \( (u_j, R, \overline{R}) \), \( \forall j \in \{1, \ldots, N\} \) from the conditions in (28). The values \( u_j \) must be positive since it is feasible to stay in one neighborhood forever or never arrive in that neighborhood for all time.

These arguments collectively demonstrate that there exists a unique fixed cost which optimizes the width of inaction \( \overline{R} - \underline{R} \). This section merely extended the initial argument of the base paper to an \( N \)-dimensional setting with total equivalence across possible classes, rather than
having an employment-unemployment asymmetric value function. See Alvarez for more information on this.

B Filtering

*High-Pass Filter*

I proposed using a filter in the frequency domain (essentially a high-pass filter around the annual frequency to penalize excessive changes in slope) to target one of the specific causes of the noisiness of the empirical data. The second proposal comes from Alvarez et al. (2015), which is the use of a Hodrick-Prescott filter.

I start with the Fourier analysis approach. Taking the Fourier transform and removing the annual component of the data results in the transformation shown in Figure 9. Recall the original histogram shown in Figure 4 for comparison. The filtering operation essentially wipes out most of the noise corresponding to the peaks at 12 months and 24 months (minus the right-censoring effect at 24 months), though the semi-annual peaks are still present (as we do not filter out in six-month increments). The data captures the essential patterns of the original histogram

![Figure 9: High-pass filtered histogram of frequency of moves.](image)
Figure 10: Hodrick-Prescott filtered histogram of frequency of moves.

without being too noisy. However, cutting out some of the frequencies in the upper range has also left some of the other information out of the final analysis; the histogram is less granular and I believe the filtering leaves too simplified a model in place. In fact, the slope after the first year shows why this filtering operation cannot stand; it is far too steep a dropoff and far too deep of one as well. The time domain has lost too much information in the process. Raising the high-pass filter barrier does not fix this problem.

_Hodrick-Prescott Filtering_

The second option, the Hodrick-Prescott filter, was used in the study which precedes this paper. It has come under immense scrutiny for similar reasons behind the rejection of the high-pass filtering exercise above; it simply cleans up the data too much to the point that trends which may be spurious become emphasized. From (10), this seems to be exactly what results. In particular, the filter essentially creates a clean sine wave out of the data. This clean functional result is interesting in its own right, but may lead to an easy overfitting by most models and makes the censoring corrections less easy to conduct. The filter embeds this information directly into the trendline, so using the trend seems disingenuous.
I rejected both of these procedures before conducting the analysis, and the following two sections discuss alternate ways of ensuring the data is workable while making fewer assumptions about the validity of parts of the data or where the information in the data exists. A carte blanche filtering operation is likely too liberal given that the structural model already places an assumed stochastic structure on the model, but it was important to discuss given preceding literature.

C Regression Summary Tables

C.1 Base Model

| Dep. Variable: | asymptotic duration | R-squared: | 0.138 |
| Model:         | OLS                  | Adj. R-squared: | 0.110 |
| Method:        | Least Squares        | F-statistic:   | 4.978 |
| Date:          | Wed, 27 Mar 2019     | Prob (F-statistic): | 1.89e-12 |
| Time:          | 17:16:49             | Log-Likelihood: | -1776.3 |
| Df Model:      | 21                   |            |      |

| Omnibus:       | 101.625              | Durbin-Watson: | 1.257 |
| Prob(Omnibus): | 0.000                | Jarque-Bera (JB): | 563.989 |
| Skew:          | 0.496                | Prob(JB):      | 3.40e-123 |
| Kurtosis:      | 7.265                |            |      |

C.2 Model without Time Coefficients

| Dep. Variable: | asymptotic duration | R-squared: | 0.100 |
| Model:         | OLS                  | Adj. R-squared: | 0.087 |
| Method:        | Least Squares        | F-statistic:   | 7.705 |
| Date:          | Wed, 27 Mar 2019     | Prob (F-statistic): | 9.37e-12 |
| Time:          | 17:59:37             | Log-Likelihood: | -1791.7 |
| Df Residuals:  | 696                  | BIC:       | 3656. |
| Df Model:      | 9                    |            |      |

| Omnibus:       | 128.755              | Durbin-Watson: | 1.274 |
| Prob(Omnibus): | 0.000                | Jarque-Bera (JB): | 809.435 |
| Skew:          | 0.646                | Prob(JB):      | 1.71e-176 |
| Kurtosis:      | 8.084                | Cond. No.     | 296. |
C.3 Model with Heterogeneity Controls

| Dep. Variable: | asymptotic duration | R-squared: | 0.119 |
| Model:         | OLS                | Adj. R-squared: | 0.104 |
| Method:        | Least Squares      | F-statistic:   | 7.994 |
| Date:          | Wed, 27 Mar 2019   | Prob (F-statistic): | 3.78e-13 |
| Time:          | 18:01:03           | Log-Likelihood: | -1612.5 |
| Df Residuals:  | 653                | BIC:          | 3303. |
| Df Model:      | 10                 |              |      |
| Omnibus:       | 56.970             | Durbin-Watson: | 1.202 |
| Prob(Omnibus): | 0.000              | Jarque-Bera (JB): | 302.572 |
| Skew:          | -0.026             | Prob(JB):     | 1.98e-66 |

C.4 Two Stage Model

| Dep. Variable: | $\kappa$ | R-squared: | 0.115 |
| Model:         | OLS      | Adj. R-squared: | 0.100 |
| Method:        | Least Squares | F-statistic: | 7.669 |
| Date:          | Wed, 27 Mar 2019   | Prob (F-statistic): | 1.56e-12 |
| Df Model:      | 10                 |              |      |

| coef  | std err | t    | P>|t| | [0.025 | 0.975 |
|-------|---------|------|-------|---------|-------|
| C(city)[Atlanta] | -2.4593 | 1.050 | -2.342 | 0.019 | -4.521 | -0.398 |
| C(city)[College]   | -1.8884 | 1.105 | -1.710 | 0.088 | -4.057 | 0.281 |
| C(city)[Los Ang]   | -3.1059 | 1.127 | -2.755 | 0.006 | -5.319 | -0.892 |
| C(city)[New Ori]   | -5.2551 | 1.682 | -3.124 | 0.002 | -8.558 | -1.952 |
| C(city)[New Yor]   | -4.3181 | 1.250 | -3.454 | 0.001 | -6.773 | -1.863 |
| C(city)[Tampa]     | -3.2019 | 1.056 | -3.033 | 0.003 | -5.274 | -1.129 |
| C(treatment)       | 0.0091  | 0.299 | 0.030  | 0.976 | -0.577 | 0.596 |
| C(recession)       | -0.2566 | 0.153 | -1.676 | 0.094 | -0.557 | 0.044 |
| DD                 | 0.8137  | 0.498 | 1.633  | 0.103 | -0.165 | 1.792 |
| lrent              | 0.3439  | 0.153 | 2.244  | 0.025 | 0.043  | 0.645 |
| propvar            | 1.7055  | 0.637 | 2.677  | 0.008 | 0.454  | 2.957 |

| Omnibus:     | 171.883 | Durbin-Watson: | 1.748 |
| Prob(Omnibus): | 0.000  | Jarque-Bera (JB): | 1023.804 |
| Kurtosis:    | 8.736   | Cond. No.      | 303. |
Table 5: Table of coefficients for the difference estimator is denoted, “DD.”

|          | coef   | std err | t     | P>|t| | [0.025] | [0.975] |
|----------|--------|---------|-------|------|--------|--------|
| C(city)[Atlanta] | 0.9908 | 2.048   | 0.484 | 0.629 | -3.030 | 5.012  |
| C(city)[College]  | 2.2074 | 2.151   | 1.026 | 0.305 | -2.016 | 6.431  |
| C(city)[Los Ang]  | -0.9419 | 2.148   | -0.439 | 0.661 | -5.159 | 3.275  |
| C(city)[New OrI]  | -5.8997 | 3.047   | -1.936 | 0.053 | -11.883 | 0.883  |
| C(city)[New Yor]  | -2.0203 | 2.357   | -0.857 | 0.392 | -6.647 | 2.607  |
| C(city)[Tampa]    | 0.1777 | 2.053   | 0.087 | 0.931 | -3.853 | 4.208  |
| C(year)[T.1998]   | -0.5781 | 1.335   | -0.433 | 0.665 | -3.200 | 2.044  |
| C(year)[T.1999]   | -0.7332 | 1.353   | -0.542 | 0.588 | -3.390 | 1.923  |
| C(year)[T.2000]   | -1.4647 | 1.269   | -1.155 | 0.249 | -3.955 | 1.026  |
| C(year)[T.2001]   | -3.4705 | 1.246   | -2.784 | 0.006 | -5.918 | -1.023 |
| C(year)[T.2002]   | -1.9115 | 1.216   | -1.572 | 0.116 | -4.299 | 0.476  |
| C(year)[T.2003]   | -2.6072 | 1.172   | -2.225 | 0.026 | -4.908 | -0.306 |
| C(year)[T.2004]   | -2.1911 | 1.152   | -1.902 | 0.058 | -4.453 | 0.071  |
| C(year)[T.2005]   | -2.0074 | 1.140   | -1.760 | 0.079 | -4.247 | 0.232  |
| C(year)[T.2006]   | -1.2788 | 1.137   | -1.124 | 0.261 | -3.512 | 0.954  |
| C(year)[T.2007]   | -0.6642 | 1.142   | -0.582 | 0.561 | -2.907 | 1.578  |
| C(year)[T.2008]   | -0.1897 | 0.390   | -0.486 | 0.627 | -0.955 | 0.576  |
| C(year)[T.2009]   | -0.2031 | 0.388   | -0.524 | 0.600 | -0.964 | 0.558  |
| C(year)[T.2010]   | -0.2920 | 0.386   | -0.756 | 0.450 | -1.050 | 0.466  |
| C(treatment)      | 0.8125 | 0.472   | 1.721 | 0.086 | -0.114 | 1.739  |
| C(recession)      | -0.6848 | 0.832   | -0.823 | 0.411 | -2.318 | 0.949  |
| DD                | 0.6184 | 0.837   | 0.738 | 0.461 | -1.026 | 2.263  |
| lrent             | 1.1841 | 0.272   | 4.346 | 0.000 | 0.649  | 1.719  |

Table 6: Table of coefficients for the model in (22) without time coefficients. The basic difference-in-difference estimator is denoted, “DD.”

|          | coef   | std err | t     | P>|t| | [0.025] | [0.975] |
|----------|--------|---------|-------|------|--------|--------|
| C(city)[Atlanta] | -1.8984 | 1.743   | -1.089 | 0.277 | -5.322 | 1.525  |
| C(city)[College]  | -0.6772 | 1.854   | -0.365 | 0.715 | -4.317 | 2.963  |
| C(city)[Los Ang]  | -3.7848 | 1.867   | -2.027 | 0.043 | -7.451 | -0.118 |
| C(city)[New OrI]  | -8.7279 | 2.877   | -3.034 | 0.003 | -14.376 | -3.080 |
| C(city)[New Yor]  | -4.9805 | 2.106   | -2.365 | 0.018 | -9.115 | -0.846 |
| C(city)[Tampa]    | -2.6470 | 1.761   | -1.503 | 0.133 | -6.104 | 0.810  |
| C(treatment)      | 0.7074 | 0.476   | 1.485 | 0.138 | -0.228 | 1.643  |
| C(recession)      | 0.6736 | 0.264   | 2.550 | 0.011 | 0.155  | 1.192  |
| DD                | 0.6999 | 0.847   | 0.826 | 0.409 | -0.964 | 2.364  |
| lrent             | 1.3635 | 0.261   | 5.228 | 0.000 | 0.851  | 1.876  |
Table 7: Table of coefficients for the model in (22) without time coefficients and an added term controlling for heterogeneity in neighborhoods. The basic difference-in-difference estimator is denoted, “DD.”

|            | coef | std err | t     | P>|t| | [0.025 | 0.975 |
|------------|------|---------|-------|-----|--------|--------|
| C(city)[Atlanta] | -0.0440 | 1.656  | -0.027 | 0.979 | -3.297 | 3.209 |
| C(city)[College] | 1.0109 | 1.743  | 0.580 | 0.562 | -2.411 | 4.433 |
| C(city)[Los Ang] | -1.7796 | 1.778  | -1.001 | 0.317 | -5.272 | 1.713 |
| C(city)[New Orl] | -6.4528 | 2.654  | -2.431 | 0.015 | -11.664 | 1.242 |
| C(city)[New Yor] | -3.0860 | 1.972  | -1.565 | 0.118 | -6.959 | 0.787 |
| C(city)[Tampa] | -0.9030 | 1.665  | -0.542 | 0.588 | -4.173 | 2.367 |
| C(treatment) | 0.5707 | 0.471  | 1.211 | 0.226 | -0.355 | 1.496 |
| C(recession) | 0.4780 | 0.242  | 1.979 | 0.048 | 0.004 | 0.952 |
| DD | 0.7748 | 0.786  | 0.985 | 0.325 | -0.769 | 2.319 |
| Irent | 1.3453 | 0.242  | 5.563 | 0.000 | 0.871 | 1.820 |
| propvar | -3.8544 | 1.005  | -3.835 | 0.000 | -5.828 | -1.881 |
Notes:
1 (Couture and Handbury, 2017)
2 (Bounds, 2006)
3 (Kanai and Kenttamaa-Squires, 2015)
4 (Christafore and Leguizamon, 2018)
5 (Ratiu, 2013)
6 (Redfearn, 2009)
7 (Kubrin and Ishizawa, 2012)
8 (Lee and Lin, 2017) for more on persistent amenities
9 (Deng et al., 2003)
10 (Temkin and Rohe, 1998)
11 (Anenberg, 2016)
12 Alvarez et al. (2015).
13 (Nekoei and Weber, 2015)
15 Spinnewijn (2015)
16 (Kroft and Notowidigdo, 2016)
17 (Bayer and McMillan, 2005)
18 (Bayer et al., 2016)
19 (Givord et al., 2013)
20 (Shaw and Hagemans, 2015)
21 (Sakizlioglu, 2014)
22 (Edlund et al., 2015)
23 (Liu et al., 2017) (Furman and Orszag, 2015)
24 (Bhattacharya et al., 2017)
25 (Kroft et al., 2016)
26 (Becker and Jahn, 2015)
27 (Lee and Lin, 2017)
28 (Kanai and Kenttamaa-Squires, 2015)
29 (Kaplan, 1998)
30 (Silverman et al., 2015)
31 (Pedersen, 2001)
32 (Schuetz et al., 2018)
33 (Lee and Lin, 2017)
Token Representation?

Measuring the Impact of Female Reservations in Panchayati Raj Institutions
in Elections to State and National Legislatures

Surbhi Bharadwaj
Abstract

Reservations have long formed a fundamental tenet of affirmative action in India. Quotas for representation of various disadvantaged groups proliferate across public educational institutions and government jobs. However, elections to public office have largely escaped such quotas, except those that are caste-based. A shift in this status quo occurred in 1992 with the establishment of the Panchayati Raj system of grassroots governance. 34% of all seats under Panchayati Raj Institutions (PRIs) were to be reserved for women under the 73rd amendment. Another constitutional amendment passed in September 2009 increased PRI quotas for women to 50%. This paper seeks to examine if the 2009 increase in mandated female representation in PRIs from 34% to 50% had an impact on female candidates’ performance in Lok Sabha and Vidhan Sabha elections, where there continue to be no quotas for women. I seek to understand if reservations indeed change underlying voter behavior—do descriptive changes in representation at one level motivate changes at another? Given that largely the same electorate votes for candidates at differing levels of government, it is likely that change in attitudes towards candidates at one level impacts elections to higher levels of government.

This paper contributes to literature by examining the upward impact of quotas in PRIs on elections to levels of government that do not have reservations for female candidates, thus giving a better indication of underlying change in general voter behavior. I found largely insignificant or slightly negative effects of female reservation on performance of female candidates in Lok Sabha and Vidhan Sabha elections immediately succeeding the imposition of these reservations in 2009. Even as female representation may signal the progressive nature of a society, there is some evidence of a backlash effect that may go away (or is simply difficult to assess over time). There is hazard in arguing the quotas can solve the problems of regressive social norms—
mandating 34 to 50% quotas in communities where women take up less than 10% of political seats can lead to unforeseen effects, and reduce the quota to a form of token, statistical representation for women. However, the initial hesitancy and backlash can lead the way to a positive equilibrium and more equitable society in the long run, achieving change that would have otherwise taken decades to attain organically.
Table of Contents

Introduction .................................................. 202
Background .................................................. 203
Literature Review .......................................... 208
Methodology .................................................. 212
Results & Analysis ......................................... 222
Shortcomings .................................................. 230
Conclusion .................................................... 231
References .................................................... 233
Appendix ....................................................... 235
Notes .......................................................... 242
Introduction

Reservations have long formed a fundamental part of affirmative action policy in India. Quotas for representation of various disadvantaged groups—stratified along lines of caste, economic status, linguistic identity and gender—proliferate across public educational institutions and government jobs. However, elections to public office have largely escaped such quotas, except caste-based ones.¹ A major shift in this status quo occurred in 1992, when the incumbent coalition government passed the 73rd Amendment to the Constitution, thereby establishing the Panchayati Raj system of grassroots governance which mandated 1/3rd reservation of seats for women. The amendment institutionalized a third level of governance in India, below the existing national (Lok Sabha) and state legislatures (Vidhan Sabhas) (Figure 1).² Panchayati Raj Institution’s (PRIs) major responsibilities were to administer local infrastructure (public buildings, water, roads) and identify targeted welfare recipients. [4]

In a landmark move, 34% of all seats in PRIs were to be reserved for women. The policy promised to radically change Indian political dynamics at time when less then 7.6% of Lok Sabha members and 3.8% of candidates in the general elections identified as female. Another constitutional amendment passed in September 2009 increased PRI quotas for women from 34% to 50%. By the general elections in 2014, 19 of India’s 28 states and union territories had implemented the 50% reservation (Figure 2).³
Figure 1: Structure of governance in India

Figure 2: States that implemented 2009 amendment for 50% reservation in PRIs
This paper will seek to examine if the 2009 increase in mandated female representation in PRIs from 34% to 50% impacted female candidates’ performance in Lok Sabha and Vidhan Sabha elections, where there continue to be no quotas for women. In doing so, I seek to understand if reservations indeed change voter behavior. One can expect that reservations for women have a positive impact on voter attitudes towards female policymakers, improving female candidates’ performance in elections. Given that largely the same electorate votes for candidates at differing levels of government, it is likely that change in attitudes towards candidates at one level impacts elections to higher levels of government. I will seek to test this hypothesis through my analysis.

While existing literature on the topic has examined the change in voter attitudes towards female candidates at the same level of government (i.e. Panchayati Raj), this paper contributes to literature by examining the upward impact of quotas in PRIs on elections to higher tiers of government that do not have reservations for female candidates, thus giving a better indication of underlying changes in general voter behavior.

**Background**

Historically, attitudes towards female empowerment in India, as in most of the world, have often been regressive. The medieval ages bore witness to practices like sati—where widowed wives were burnt on the pyre alongside their dead husbands—dowry, child marriage, and rampant discrimination, compounded by oppressive societal structures like the caste system. 19th and 20th century India saw significant strides towards greater female representation in public life. Several of the independence movement’s greatest leaders were women. The likes of Sarojini Naidu, Annie Beasant, and Vijaya Lakshmi Pandit broke the conventional mold society crafted for women. They appeared in public life, campaigned before masses and advocated for Indian
freedom from British colonial rule. Upon independence in 1947, the country adopted a Constitution more progressive than the society it was drafted for. The Constitution enshrined universal adult suffrage, preceding several powerful democracies in the West, and by 1966, Indira Gandhi was elected as the first female head of state, an achievement yet to be realised in U.S.A.

![Child sex ratio, Indian states](image)

**Figure 3: State-wise sex ratio, Census 2011**

However, these singular achievements do not reflect broader trends in society. While the status of women in Indian society has progressively improved since independence, 21st century India faces several challenges including sexual harassment, female foeticide, and dismal levels of female participation in public and professional life. The country ranked 130th on UNDP’s gender inequality index, which considers metrics such as female labor force participation, education, maternal mortality and representation in Parliament. [11] The female literacy rate is 65.46% compared to 82.14% for men. [1] While women form over 80% of the informal sector employed...
in agriculture and dairy, they account for less than 30% of the population engaged in paid work.

[9]

Attitudes towards gender also vary dramatically across the country. This is reflected in the range of sex ratios across states (Figure 3). Sex-selective abortion and female infanticide are especially rampant in North Indian states. Haryana has only 834 female children per 1000 males—one of the most skewed sex ratios in the world. Similar geographical discrepancies can be noted in the number of cases of dowry deaths reported (Figure 15). [19] India has also gained international infamy for incidents of violence against women. A spate of brutal rape cases in 2012 mobilized large parts of the country to advocate for female empowerment. These social movements had political ramifications, with parties tweaking their platforms to appeal to the popular mood, and coincided with the institution of 50% reservation in PRIs. Conversation in the public sphere on issues of women’s empowerment thus took on a renewed urgency between the 2009 and 2014 elections.

Female political participation has increased since independence but representation in legislative houses remains abysmally low. The rate of voter turnout among women for Lok Sabha elections in 1962 was 46.63% (compared to 63.63% for men) and rose to 65.63% in 2014 (versus 67.09% for men) (Figure 4). This change occurred as female representation peaked in the 16th Lok Sabha, elected in 2014, where 11.5% of seats were won by female candidates and 7.6% of contestants were women (Figure 4). This places India at the 151st position in the UNDP rankings based on proportion of female members of Parliament, behind countries like Iraq, Jordan and Senegal, that perform worse on other metrics. [11] However, broader measures of political participation, such as representation and engagement at state level elections, as measured by the World Economic Forum, place India 9th worldwide, indicating greater overall female political
participation as compared to several OECD countries. [6] This may point to some sort of glass ceiling for female representation—while female candidates are well received at lower levels of government, that often have greater provisions for female affirmative action, they face significant barriers in advancing to the highest echelons of power. Considering trends from past elections, one can note that while the number of female candidates has risen significantly, more than doubling between 1999 and 2014, the number of elected female MPs has stayed relatively stable. The overall gender composition of the House has thus not changed. In 2014, only 368 out of 548 constituencies saw at least one female candidate contesting, and only 63 female candidates won a seat. However, women tend to have a higher success rate, defined as the ratio of the number of winners versus number of candidates for each gender. In 2014, 9.4% of all female candidates ultimately won a seat compared to 6.3% of male candidates (Figure 6). Given the high barriers to female candidacy, it may simply be the case that female candidates are stronger on average than their male counterparts which motivates their decision to run in an otherwise fraught environment. They may also have greater funds and backing from the parties whose platforms they contest on.

<table>
<thead>
<tr>
<th>Year</th>
<th>Candidates</th>
<th>Female candidates</th>
<th>% Female candidates</th>
<th>Winners</th>
<th>Female winners</th>
<th>% Female winners</th>
</tr>
</thead>
<tbody>
<tr>
<td>1967</td>
<td>2405</td>
<td>71</td>
<td>2.952</td>
<td>527</td>
<td>32</td>
<td>6.072</td>
</tr>
<tr>
<td>1971</td>
<td>2789</td>
<td>91</td>
<td>3.263</td>
<td>520</td>
<td>27</td>
<td>5.192</td>
</tr>
<tr>
<td>1977</td>
<td>2439</td>
<td>74</td>
<td>3.034</td>
<td>542</td>
<td>19</td>
<td>3.506</td>
</tr>
<tr>
<td>1980</td>
<td>4675</td>
<td>147</td>
<td>3.144</td>
<td>555</td>
<td>39</td>
<td>5.607</td>
</tr>
<tr>
<td>1984</td>
<td>5312</td>
<td>164</td>
<td>3.087</td>
<td>514</td>
<td>44</td>
<td>8.560</td>
</tr>
<tr>
<td>1989</td>
<td>6159</td>
<td>203</td>
<td>3.296</td>
<td>529</td>
<td>31</td>
<td>5.860</td>
</tr>
<tr>
<td>1991</td>
<td>8759</td>
<td>334</td>
<td>3.813</td>
<td>526</td>
<td>40</td>
<td>7.605</td>
</tr>
<tr>
<td>1996</td>
<td>13806</td>
<td>609</td>
<td>4.361</td>
<td>545</td>
<td>43</td>
<td>7.890</td>
</tr>
<tr>
<td>1998</td>
<td>4767</td>
<td>279</td>
<td>5.853</td>
<td>545</td>
<td>45</td>
<td>8.257</td>
</tr>
<tr>
<td>1999</td>
<td>4648</td>
<td>290</td>
<td>6.239</td>
<td>543</td>
<td>51</td>
<td>9.392</td>
</tr>
<tr>
<td>2004</td>
<td>5481</td>
<td>359</td>
<td>6.550</td>
<td>547</td>
<td>47</td>
<td>8.592</td>
</tr>
<tr>
<td>2009</td>
<td>8070</td>
<td>560</td>
<td>6.939</td>
<td>543</td>
<td>58</td>
<td>10.68</td>
</tr>
<tr>
<td>2014</td>
<td>8301</td>
<td>670</td>
<td>8.071</td>
<td>548</td>
<td>63</td>
<td>11.50</td>
</tr>
</tbody>
</table>
Figure 4: Female electoral participation and representation, 1968-2009

Figure 5: Female candidacy and wins, 1963-2014
Literature Review

Quotas are a popular system of mandated affirmative action across the world, especially with regards to gender. These may manifest in terms of voluntary self-imposed party quotas, legislated candidate quotas (which mandate a certain proportion of candidacies are reserved for women), and reserved seats (which only women can contest). Of the countries that currently have gender quotas, 61% have voluntary party quotas, 38% have legislated candidate quotas, and 20% have reserved seats. [10] Gender quotas were first introduced in Nordic countries in the 1970s, by setting a 40% target for “representation of both sexes on electoral lists.” This was quickly followed by similar schemes in Latin America (legislated candidate quotas) and other Western European countries (voluntary party quotas). Such policy measures can have drastic effects in short periods of time. For example, Rwanda, where only 18% of members of parliament were female in the 1990s, introduced 30% quotas for women in 2003. Since the change Rwanda has gone on to elect

Figure 6: Success rate of candidates, 1963-2009
Parliaments comprised of 56% and 64% women in the 2008 and 2013 elections respectively. [18] Most countries in the top 20 rankings for female representation in the national legislature exercise some form of quota policy for women. [17] Thus quotas actively change the descriptive make ups of legislative houses, which can spur substantive change in policy making as well. In 2018, Clayton and Zitterberg found that increased female representation led to greater relative spending on public health, and decreased expenditure on military. [16]

In the Indian context, female political participation has been the subject of much research, the large democracy affording a fertile testing ground for regression based analyses. The 1992 introduction of reservation in PRIs involved randomized allocation of Panchayat seats to women. This forms an important largescale natural experiment that can isolate the causal effects and impacts of female representation in local leadership. Male and female political leaders often make different policy decisions, which can have ramifications for the constituencies that elect them. Further literature has examined the impact of quotas for women in female political participation, both as candidates and as voters in elections.

In a report for the World Bank, Rohini Pande (2012) lists popular arguments for and against gender-based quotas: [10]

**Pro-quota:**

1. Improves descriptive and substantive representation
2. Reduces taste-based discrimination, as a result of regressive social norms and perceptions in the electorate
3. Increases information about the performance and competence of female candidates, countering prevailing social norms and leading to the role model effect as more women are inspired to participate in the political process
4. Improves investment in women and education, as female policy makers are more familiar with the demands of their constituents

**Anti-quota:**

1. May lead to crowding out of other minorities, as women from the most privileged groups stand to benefit

2. Worsens allocation of leadership positions if female candidates are less competent or experienced than male counterparts

3. Worsens attitudes among voters if they believe their choices are restricted as a result of the quota, or severely contradict prevailing social norms. This may lead to a backlash effect against female candidates

The election of female policy makers has been associated with increased provision of public goods [4], lower levels of corruption [3] and greater investment in education. [2] Considering the 1992 reservation, Chattopadhyay and Duflo (2004) studied the impact of policy making by female leaders on the provision of public goods. [4] They noted that investment in drinking water and provision of public goods significantly increased in seats reserved for women. Female policy makers thus take decisions that more closely reflect issues that are relevant to villagers. This indicates that female holders of political office are as, if not more, capable as their male counterparts. Clots-Figueras (2012) and Beaman et al (2012) noted that educational attainment rates improved in seats reserved for women councilors. [2] [8] They found that the “gender gap in adolescent educational attainment was erased” and closed by 20 and 32% in parental and adolescent aspirations, respectively. [4]
These positive associations should be able to counter prevailing misogynistic prejudices among voters against female candidates. Observing the success of female Panchayat members may thus improve information about female competence in policy making and motivate voters to vote for female candidates in state and general elections. While exposure may alter social perceptions, Rudman and Fairchild (2004) argue that it may also lead to a “backlash effect.” [12] A mandated quota may be seen as restricting voter choice or violating voters’ sense of identity with regards to a traditionally male-dominated profession. This could lead to female candidates underperforming on average as voters react to a policy deemed unfair. However, Beaman et al (2009) note that through the exposure effect, in the long-run quotas can improve the precision of voters’ information about the expected effectiveness of future female leaders thereby reducing statistical discrimination. They write, “mandated exposure can improve the perception of women leaders’ effectiveness by reducing the risk associated with electing a woman.” [15] They study this phenomenon in PRIs, noting that despite an initial backlash in the first succeeding election cycle, after ten years of quotas, women are more likely to “stand for, and win, elected positions in councils required to have a female chief councilor in the previous two elections.” They argue that exposure to female councilors thus improves perceptions of female leader effectiveness and weakens stereotypes about gender roles in the public and domestic spheres. Representation further encourages future female candidacy. Bhalotra et al (2011) note that women winning office leads to a 9.2% increase in female candidates fielded by major political parties in subsequent elections. [13] While they argue that this is a result of incumbent female candidates being more likely to contest again (as opposed to an entry of new female candidates) there is also an observed reduction in party bias against female candidates.
While the literature seems promising these studies have focused on the change in voter behavior in successive elections within one level of government—the Panchayati Raj. In this paper, I will examine the impact of quotas across different levels of government.

**Methodology**

*Model*

I use a regression model to isolate the treatment effect of the implementation of 50% reservation in some states after the 2009 and before the 2014 general elections to the Lok Sabha (Figure 2). I estimate the following model:

\[
y_i = \alpha_{female} + \delta_{female} \times reservations_{50} + \gamma_{ie} + \epsilon_i
\]

Here, the treatment effect is denoted by \( \delta \) and gives us the impact of female reservations on the dependent variable. I consider the following three dependent variables:

1. **mean vote share**: a continuous variable denoting the percentage of vote share received by a candidate in a given race. This can help isolate changes in female candidate’s performance, even if the reservation does not actually lead to them winning more seats. As noted below, I modify this variable to account for competition in different electoral races by subtracting the mean vote share for a candidate in a given election.

2. **Win**: a binary variable that takes on values of 0 and 1 depending on whether the candidate wins the seat. I run probit regressions on this variable due to its binary nature.

3. **Position**: discrete variable that gives the rank of a candidate in a given race and, like mean vote share, can provide a more accurate assessment of performance of female candidates

\( \gamma_{ie} \) consists of dummy variables to account for other observed measures and fixed effects, to better isolate the effect of reservations.
In particular, I include interactions between the female dummy and variables like incumbency and turnout to isolate any gender-specific effects. For example, the anti-incumbency effect may be stronger against women or female candidates’ performance may change with change in voter turnout, as less motivated or politically active members of the electorate exercise their suffrage. I also include party interactions with year to account for general voting trends for or against one party in a given election cycle. For instance, it may be possible that more women win in an election cycle. However, if the party that is favored regardless has more female candidates, this does not point us to a causal impact of the reservation—more females are elected simply due to their association with the popular party.

This analysis could be made more robust by controlling for socio-economic variables such as literacy rate, sex rate, female work force participation etc., which all point to the community’s prevailing perceptions of gender norms. However, I was not able to find data that is granular enough to improve the model’s fit. I thus rely on the assumption that these underlying socio-economic variables would not have changed significantly over the five-year time span between the 2009 and 2014 general elections. To ensure that female candidacy did not differ significantly before the imposition of the reservation, I further run a differences t-test on the treatment and control constituencies before the increased reservation was enacted as a robustness check. While the proportion of female candidates is slightly lower in the treatment group, the difference is not significant (Figure 7).
I further conduct this analysis through three approaches:

1. a **naive regression including all seats contested** and candidates winning more than 5% of the vote share in the 2009 and 2014 elections

2. considering only those constituencies that are **geographically contiguous and lie on borders of pairs of states that did and did not implement the 50% increase in reservation** (Figure 16). I assume these groups of contiguous constituencies have similar baseline attitudes towards women, reducing bias from unobserved differences, and allowing us to more accurately determine the treatment effect

3. while the first two approaches consider the nationwide general elections to the Lok Sabha, one may argue that upward impacts of reservations at the Panchayat level may not reach the highest level of public office within five years of quota implementation. I thus consider the **performance of women in Vidhan Sabha elections before and after the imposition of the additional reservation**. This estimate may be more biased since it is harder to isolate underlying trends from the treatment effect of the reservation as all constituencies in a given state experience the increase in reservation and I simply estimate the change in performance across the two Vidhan Sabha elections, with some

---

**Figure 7:** Pre-treatment (2009) difference in female candidacy in treatment and control groups

---

**Table:**

<table>
<thead>
<tr>
<th>Group</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2166</td>
<td>.0743306</td>
<td>.0056375</td>
<td>.2623687</td>
<td>.0632752 .0853859</td>
</tr>
<tr>
<td>1</td>
<td>5904</td>
<td>.0675813</td>
<td>.0032673</td>
<td>.2510473</td>
<td>.0611763 .0739863</td>
</tr>
<tr>
<td>combined</td>
<td>8070</td>
<td>.0693928</td>
<td>.002829</td>
<td>.2541367</td>
<td>.0638473 .0749384</td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>.0067493</td>
<td>.0065158</td>
<td>-.0060257</td>
<td>.0195242</td>
</tr>
</tbody>
</table>

**t-test male if year == 2009, by(policy_50) unequal**

Two-sample t test with unequal variances

\[
\text{diff} = \text{mean}(0) - \text{mean}(1) \quad t = 1.0358
\]

Satterthwaite’s degrees of freedom = 3710.15

Pr(T < t) = 0.8498 \quad Pr(|T| > |t|) = 0.3003 \quad Pr(T > t) = 0.1502
controls. It is thus difficult to conclude whether the model captures the causal effect of reservations, or simply reflects an underlying trend due to other factors.

Data

The primary datasets I used are sourced from the Trivedi Center for Political Data at Ashoka University. The datasets note all candidates who contested seats in state and general elections, and their performance in terms of votes and margins. [5] I carried out some fuzzy matching across candidate names to generate incumbency variables. I manually coded in geographically contiguous constituencies for the second part of my analysis. The dummy variable \( \text{reservation}_{50} \) accounts for all states that implemented 50% reservation for women in PRIs following the September 2009 legislation.

Reducing Noise

Candidates contesting elections in India have to pay a security deposit. Candidates that fail to secure more than 1.6th of the votes cast in the constituency lose their deposit. In the 2014 general elections, 84.8% of the 8,301 contesting candidates lost their deposits (Figure 8). Digging deeper into the vote share distribution, one can note that over 74.32% of candidates secured less than 1% of the vote in the 2014 elections (amounting to 6,169 of 8,301 candidates contesting). One can assume that candidates who secure such small percentages of the vote share were never serious contenders in the first place. Their performance is largely random and not indicative of underlying voter attitudes. I seek to reduce the noise in the data and better isolate the treatment effect of the reservations on serious contenders to these positions by dropping all candidates who obtain less than 5% vote share.
I only consider candidates running for elections during general elections in 2009 and 2014. A small number of seats, comprising 1.3% of candidates in the dataset, are contested through by-elections between the 2009 and 2014 general elections (Figure 9). By-elections are conducted to fill seats that become empty usually due to the death or resignation of the sitting MP. However, these are small in number and to keep comparisons consistent, I drop these candidates from the dataset.

Figure 8: Distribution of vote shares won by candidates across the 2009, and 2014 general elections

**By-Elections**

By-Elections
Normalizing Candidate Vote Shares

To analyze regressions considering vote share as a dependent variable, I modify the variable by each electoral race to make the coefficients more comparable, as the difficulty of an electoral race may depend on the number of competent candidates contesting (Figure 17, 18). For example, a 5% vote share gain in a two-way electoral race implies a differing degree of candidate strength compared to a 5% gain in a race with more than four candidates. If there were greater number of competent candidates in a race, gaining the same amount of vote share could require substantially larger amounts of effort. To account for this, I modify the vote share won by each candidate in every unique electoral race (by year and constituency) by subtracting vote share in that electoral race from the candidate’s vote share.\textsuperscript{9}
\[ \text{mean \_ vote \_ share}_i = \text{vote \_ share}_i - \overline{\text{vote \_ share}}, \]  

(2)

where, \( \overline{\text{vote \_ share}} = \frac{\sum_{i=1}^{N} \text{vote \_ share}_i}{N} \),

\( N = \) number of candidates,

and \( \sum_{i=1}^{N} \text{vote \_ share}_i = 100 \)
<table>
<thead>
<tr>
<th></th>
<th>(1) mean_vote b/p</th>
<th>(2) mean_vote b/p</th>
<th>(3) mean_vote b/p</th>
<th>(4) win b/p</th>
<th>(5) win b/p</th>
<th>(6) win b/p</th>
<th>(7) Position b/p</th>
<th>(8) Position b/p</th>
<th>(9) Position b/p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>main</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2.267</td>
<td>0.598</td>
<td>-0.781</td>
<td>0.010</td>
<td>-0.088</td>
<td>-0.153</td>
<td>-0.011</td>
<td>0.096</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td>(0.605)</td>
<td>(0.886)</td>
<td>(0.846)</td>
<td>(0.980)</td>
<td>(0.823)</td>
<td>(0.700)</td>
<td>(0.968)</td>
<td>(0.731)</td>
<td>(0.539)</td>
</tr>
<tr>
<td>Female * Reservation 50</td>
<td>1.825</td>
<td>2.729</td>
<td>0.589</td>
<td>-0.092</td>
<td>0.045</td>
<td>-0.038</td>
<td>-0.052</td>
<td>-0.113</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.335)</td>
<td>(0.110)</td>
<td>(0.732)</td>
<td>(0.571)</td>
<td>(0.785)</td>
<td>(0.827)</td>
<td>(0.664)</td>
<td>(0.306)</td>
<td>(0.890)</td>
</tr>
<tr>
<td>Incumbent</td>
<td>7.647***</td>
<td>7.507***</td>
<td>7.879***</td>
<td>0.453***</td>
<td>0.569***</td>
<td>0.597***</td>
<td>-0.484***</td>
<td>-0.480***</td>
<td>-0.493***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Female * Incumbent</td>
<td>-0.644</td>
<td>-0.997</td>
<td>-1.915</td>
<td>-0.095</td>
<td>-0.118</td>
<td>-0.137</td>
<td>0.087</td>
<td>0.105</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>(0.766)</td>
<td>(0.604)</td>
<td>(0.312)</td>
<td>(0.618)</td>
<td>(0.541)</td>
<td>(0.497)</td>
<td>(0.496)</td>
<td>(0.368)</td>
<td>(0.282)</td>
</tr>
<tr>
<td>Caste</td>
<td>-2.424***</td>
<td>-2.584***</td>
<td>-2.503***</td>
<td>0.001</td>
<td>-0.006</td>
<td>-0.014</td>
<td>-0.002</td>
<td>0.007</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Turnout Percentage</td>
<td>0.034</td>
<td>0.044*</td>
<td>0.108**</td>
<td>0.003</td>
<td>0.004*</td>
<td>-0.006***</td>
<td>-0.007***</td>
<td>-0.010***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
<td>(0.046)</td>
<td>(0.004)</td>
<td>(0.188)</td>
<td>(0.039)</td>
<td>(0.263)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Female * Turnout</td>
<td>-0.011</td>
<td>0.013</td>
<td>0.051</td>
<td>0.003</td>
<td>0.005</td>
<td>0.006</td>
<td>-0.002</td>
<td>-0.003</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.871)</td>
<td>(0.849)</td>
<td>(0.438)</td>
<td>(0.628)</td>
<td>(0.473)</td>
<td>(0.325)</td>
<td>(0.720)</td>
<td>(0.476)</td>
<td>(0.235)</td>
</tr>
<tr>
<td>2014</td>
<td>0.538</td>
<td>0.377</td>
<td>0.237</td>
<td>-0.026</td>
<td>-0.026</td>
<td>-0.024</td>
<td>0.018</td>
<td>-0.017</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.341)</td>
<td>(0.584)</td>
<td>(0.741)</td>
<td>(0.609)</td>
<td>(0.704)</td>
<td>(0.739)</td>
<td>(0.625)</td>
<td>(0.722)</td>
<td>(0.508)</td>
</tr>
<tr>
<td>INC * 2009</td>
<td>8.719***</td>
<td>6.746***</td>
<td></td>
<td>0.628***</td>
<td>0.582***</td>
<td>-0.577***</td>
<td>-0.484***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC * 2014</td>
<td>-0.815</td>
<td>-3.572***</td>
<td></td>
<td>-0.699***</td>
<td>-0.764***</td>
<td>0.122*</td>
<td>0.253***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.318)</td>
<td>(0.000)</td>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.030)</td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BJP * 2009</td>
<td>3.382***</td>
<td>1.379</td>
<td></td>
<td>0.248**</td>
<td>0.204*</td>
<td>-0.250***</td>
<td>-0.159*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.143)</td>
<td></td>
<td>(0.004)</td>
<td>(0.021)</td>
<td>(0.000)</td>
<td>(0.017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BJP * 2014</td>
<td>11.720***</td>
<td>10.113***</td>
<td></td>
<td>1.078***</td>
<td>1.079***</td>
<td>-0.678***</td>
<td>-0.599***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>22.006***</td>
<td>18.733***</td>
<td>18.752***</td>
<td>-0.696***</td>
<td>-1.032***</td>
<td>-1.095***</td>
<td>2.613***</td>
<td>2.839***</td>
<td>2.702***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.004)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>R-Square</td>
<td>0.054</td>
<td>0.142</td>
<td>0.228</td>
<td>0.042</td>
<td>0.118</td>
<td>0.152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Cases</td>
<td>3263.000</td>
<td>3263.000</td>
<td>3263.000</td>
<td>3263.000</td>
<td>3263.000</td>
<td>3263.000</td>
<td>3263.000</td>
<td>3263.000</td>
<td>3263.000</td>
</tr>
</tbody>
</table>
Table 2: Regression on geographically contiguous constituencies

<table>
<thead>
<tr>
<th></th>
<th>(1) mean_voted</th>
<th>(2) mean_voted</th>
<th>(3) mean_voted</th>
<th>(4) win</th>
<th>(5) win</th>
<th>(6) win</th>
<th>(7) Position</th>
<th>(8) Position</th>
<th>(9) Position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b/p</td>
<td>b/p</td>
<td>b/p</td>
<td>b/p</td>
<td>b/p</td>
<td>b/p</td>
<td>b/p</td>
<td>b/p</td>
<td>b/p</td>
</tr>
<tr>
<td>Female</td>
<td>-16.949</td>
<td>-5.990</td>
<td>-7.470</td>
<td>-0.968</td>
<td>-0.900</td>
<td>-0.217</td>
<td>0.154</td>
<td>-0.555</td>
<td>-0.468</td>
</tr>
<tr>
<td></td>
<td>(0.186)</td>
<td>(0.581)</td>
<td>(0.498)</td>
<td>(0.482)</td>
<td>(0.948)</td>
<td>(0.878)</td>
<td>(0.865)</td>
<td>(0.491)</td>
<td>(0.562)</td>
</tr>
<tr>
<td>Female * Reserve 50</td>
<td>-3.325</td>
<td>-2.010</td>
<td>-1.206</td>
<td>-1.126*</td>
<td>-0.846</td>
<td>-0.939</td>
<td>0.383</td>
<td>0.398</td>
<td>0.326</td>
</tr>
<tr>
<td></td>
<td>(0.453)</td>
<td>(0.580)</td>
<td>(0.742)</td>
<td>(0.038)</td>
<td>(0.143)</td>
<td>(0.100)</td>
<td>(0.192)</td>
<td>(0.234)</td>
<td>(0.207)</td>
</tr>
<tr>
<td>Incumbent</td>
<td>7.010***</td>
<td>7.160***</td>
<td>7.312***</td>
<td>3.419*</td>
<td>5.333*</td>
<td>5.577*</td>
<td>-0.501***</td>
<td>-0.516***</td>
<td>-0.526***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.031)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Female * Incumbent</td>
<td>0.910</td>
<td>-1.280</td>
<td>-0.906</td>
<td>0.703</td>
<td>0.400</td>
<td>0.467</td>
<td>-0.144</td>
<td>0.110</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>(0.842)</td>
<td>(0.723)</td>
<td>(0.806)</td>
<td>(0.218)</td>
<td>(0.439)</td>
<td>(0.376)</td>
<td>(0.965)</td>
<td>(0.677)</td>
<td>(0.643)</td>
</tr>
<tr>
<td>Caste</td>
<td>-1.399</td>
<td>-1.686</td>
<td>-0.573</td>
<td>0.069</td>
<td>0.069</td>
<td>0.162</td>
<td>-0.086</td>
<td>-0.068</td>
<td>-0.142</td>
</tr>
<tr>
<td></td>
<td>(0.210)</td>
<td>(0.102)</td>
<td>(0.679)</td>
<td>(0.471)</td>
<td>(0.544)</td>
<td>(0.265)</td>
<td>(0.226)</td>
<td>(0.325)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Turnout Percentage</td>
<td>-0.070</td>
<td>-0.076</td>
<td>0.009</td>
<td>-0.000</td>
<td>0.001</td>
<td>0.003</td>
<td>-0.003</td>
<td>-0.002</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.330)</td>
<td>(0.263)</td>
<td>(0.948)</td>
<td>(0.975)</td>
<td>(0.907)</td>
<td>(0.810)</td>
<td>(0.540)</td>
<td>(0.574)</td>
<td>(0.882)</td>
</tr>
<tr>
<td>Female * Turnout</td>
<td>0.358</td>
<td>0.183</td>
<td>0.206</td>
<td>0.023</td>
<td>0.012</td>
<td>0.016</td>
<td>-0.010</td>
<td>0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
<td>(0.333)</td>
<td>(0.278)</td>
<td>(0.320)</td>
<td>(0.608)</td>
<td>(0.524)</td>
<td>(0.537)</td>
<td>(0.933)</td>
<td>(0.947)</td>
</tr>
<tr>
<td>2014</td>
<td>1.808</td>
<td>-4.943**</td>
<td>-5.772**</td>
<td>0.080</td>
<td>-1.001***</td>
<td>-1.061***</td>
<td>-0.102</td>
<td>0.343**</td>
<td>0.315*</td>
</tr>
<tr>
<td></td>
<td>(0.252)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.586)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.353)</td>
<td>(0.008)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>INC * 2009</td>
<td>6.203**</td>
<td>5.713*</td>
<td>0.302</td>
<td>0.316</td>
<td></td>
<td></td>
<td>-0.360*</td>
<td>-0.354*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.012)</td>
<td>(0.160)</td>
<td>(0.148)</td>
<td></td>
<td></td>
<td>(0.034)</td>
<td>(0.037)</td>
<td></td>
</tr>
<tr>
<td>INC * 2014</td>
<td>8.647***</td>
<td>8.182***</td>
<td>0.596</td>
<td>0.644*</td>
<td></td>
<td></td>
<td>-0.528***</td>
<td>-0.523***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.052)</td>
<td>(0.035)</td>
<td></td>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>BJP * 2009</td>
<td>3.508</td>
<td>3.078</td>
<td>-0.030</td>
<td>-0.014</td>
<td></td>
<td></td>
<td>-0.088</td>
<td>-0.102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.158)</td>
<td>(0.900)</td>
<td>(0.952)</td>
<td></td>
<td></td>
<td>(0.631)</td>
<td>(0.581)</td>
<td></td>
</tr>
<tr>
<td>BJP * 2014</td>
<td>23.644***</td>
<td>23.371***</td>
<td>2.627***</td>
<td>2.680***</td>
<td></td>
<td></td>
<td>-1.452***</td>
<td>-1.454***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>23.503***</td>
<td>21.833***</td>
<td>21.916***</td>
<td>-0.724*</td>
<td>-0.912*</td>
<td>-0.965</td>
<td>2.675***</td>
<td>2.747***</td>
<td>2.484***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.005)</td>
<td>(0.026)</td>
<td>(0.020)</td>
<td>(0.361)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>R-Square</td>
<td>0.042</td>
<td>0.296</td>
<td>0.306</td>
<td></td>
<td></td>
<td></td>
<td>0.042</td>
<td>0.221</td>
<td>0.215</td>
</tr>
<tr>
<td>Number of Cases</td>
<td>425,000</td>
<td>425,000</td>
<td>425,000</td>
<td>425,000</td>
<td>425,000</td>
<td>425,000</td>
<td>425,000</td>
<td>425,000</td>
<td>425,000</td>
</tr>
<tr>
<td></td>
<td>(1) Assam</td>
<td>(2) Andhra Pradesh</td>
<td>(3) Chhattisgarh</td>
<td>(4) Gujarat</td>
<td>(5) Jharkhand</td>
<td>(6) Madhya Pradesh</td>
<td>(7) Maharashtra</td>
<td>(8) Odisha</td>
<td>(9) Uttarakhand</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>------------</td>
<td>--------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>(0.568)</td>
<td>(0.319)</td>
<td>(0.782)</td>
<td>(0.158)</td>
<td>(0.933)</td>
<td>(0.352)</td>
<td>(0.273)</td>
<td>(0.374)</td>
<td>(0.042)</td>
</tr>
<tr>
<td><strong>Female * Reservation 50</strong></td>
<td>-1.306</td>
<td>1.783</td>
<td>-4.218</td>
<td>4.001</td>
<td>0.444</td>
<td>-1.174</td>
<td>3.368</td>
<td>5.837*</td>
<td>2.921</td>
</tr>
<tr>
<td></td>
<td>(0.603)</td>
<td>(0.358)</td>
<td>(0.187)</td>
<td>(0.139)</td>
<td>(0.900)</td>
<td>(0.457)</td>
<td>(0.123)</td>
<td>(0.044)</td>
<td>(0.362)</td>
</tr>
<tr>
<td><strong>Incumbent</strong></td>
<td>5.594***</td>
<td>3.368***</td>
<td>-2.435</td>
<td>4.937***</td>
<td>5.440***</td>
<td>0.266</td>
<td>9.687***</td>
<td>9.056***</td>
<td>6.002***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.006)</td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.007)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td><strong>Female * Incumbent</strong></td>
<td>-1.190</td>
<td>-0.032</td>
<td>7.946**</td>
<td>-3.175</td>
<td>2.589</td>
<td>0.529</td>
<td>-0.346</td>
<td>0.437</td>
<td>0.862</td>
</tr>
<tr>
<td></td>
<td>(0.645)</td>
<td>(0.988)</td>
<td>(0.009)</td>
<td>(0.107)</td>
<td>(0.577)</td>
<td>(0.829)</td>
<td>(0.889)</td>
<td>(0.878)</td>
<td>(0.769)</td>
</tr>
<tr>
<td><strong>Caste</strong></td>
<td>0.230</td>
<td>-0.247</td>
<td>0.669</td>
<td>0.138</td>
<td>-0.395</td>
<td>-0.379</td>
<td>-0.115</td>
<td>0.019</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.576)</td>
<td>(0.581)</td>
<td>(0.106)</td>
<td>(0.748)</td>
<td>(0.381)</td>
<td>(0.186)</td>
<td>(0.772)</td>
<td>(0.964)</td>
<td>(0.999)</td>
</tr>
<tr>
<td><strong>Turnout Percentage</strong></td>
<td>-0.012</td>
<td>-0.051</td>
<td>-0.134**</td>
<td>-0.038</td>
<td>0.001</td>
<td>-0.092**</td>
<td>-0.027</td>
<td>-0.022</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(0.672)</td>
<td>(0.074)</td>
<td>(0.002)</td>
<td>(0.479)</td>
<td>(0.993)</td>
<td>(0.009)</td>
<td>(0.313)</td>
<td>(0.740)</td>
<td>(0.682)</td>
</tr>
<tr>
<td><strong>Female * Turnout</strong></td>
<td>0.036</td>
<td>0.086</td>
<td>0.062</td>
<td>-0.167</td>
<td>-0.015</td>
<td>-0.085</td>
<td>0.056</td>
<td>-0.304</td>
<td>-0.300*</td>
</tr>
<tr>
<td></td>
<td>(0.647)</td>
<td>(0.302)</td>
<td>(0.782)</td>
<td>(0.105)</td>
<td>(0.951)</td>
<td>(0.428)</td>
<td>(0.603)</td>
<td>(0.273)</td>
<td>(0.019)</td>
</tr>
<tr>
<td><strong>year== 2007.0000</strong></td>
<td>0.322</td>
<td>-3.439***</td>
<td>0.775</td>
<td>-10.570***</td>
<td>2.401</td>
<td>1.919*</td>
<td>0.777</td>
<td>-0.977</td>
<td>-1.667</td>
</tr>
<tr>
<td></td>
<td>(0.705)</td>
<td>(0.302)</td>
<td>(0.603)</td>
<td>(0.000)</td>
<td>(0.060)</td>
<td>(0.028)</td>
<td>(0.231)</td>
<td>(0.472)</td>
<td>(0.138)</td>
</tr>
<tr>
<td><strong>control_INC_yr1</strong></td>
<td>5.610***</td>
<td>8.646***</td>
<td>16.459***</td>
<td>14.280***</td>
<td>1.116</td>
<td>11.449***</td>
<td>7.752***</td>
<td>5.619***</td>
<td>10.327***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.008)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.014)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-1.557</td>
<td>4.191*</td>
<td>-4.825</td>
<td>-1.415</td>
<td>-2.527</td>
<td>-3.428</td>
<td>-1.834</td>
<td>2.100</td>
<td>-5.306</td>
</tr>
<tr>
<td></td>
<td>(0.530)</td>
<td>(0.036)</td>
<td>(0.192)</td>
<td>(0.707)</td>
<td>(0.556)</td>
<td>(0.188)</td>
<td>(0.295)</td>
<td>(0.229)</td>
<td>(0.126)</td>
</tr>
<tr>
<td><strong>R-Square</strong></td>
<td>0.160</td>
<td>0.114</td>
<td>0.489</td>
<td>0.221</td>
<td>0.118</td>
<td>0.406</td>
<td>0.148</td>
<td>0.276</td>
<td>0.195</td>
</tr>
<tr>
<td><strong>Number of Cases</strong></td>
<td>1290.000</td>
<td>1753.000</td>
<td>485.000</td>
<td>1287.000</td>
<td>618.000</td>
<td>1376.000</td>
<td>2022.000</td>
<td>972.000</td>
<td>716.000</td>
</tr>
</tbody>
</table>
Results and Analysis

Impact of 2009 Amendment for 50% Reservation

The results from my model yield largely inconclusive or negative evidence regarding the causal impact of quotas for women on their performance in Lok Sabha elections. Under the naive regression (Table 1), I get insignificant results across my three dependent variables. The treatment effect on vote share is weakly positive while dependent variables win and position also pick up some negative impact of the reservation.

The more telling analysis follows from the second, constrained dataset (Table 2). Analyzing only neighboring constituencies in the treatment and control groups accounts for underlying differences in prevailing social norms, that are not captured by the variables in the model. I further include dummy variables to account for the various groupings of these geographically contiguous constituencies (Figure 16). These models yield better fit and more consistent coefficients. The treatment effect is seen to be negative across the models considering vote share and win probability as the dependent variables and positive for the model considering position—indicating worse performance by female candidates in states with 50% mandated reservation in PRIs. Model (4) yields a significant result at a 5% significance level. Since this is a probit model, I compute the marginal effects of the coefficients to enable interpretation (Figure 21). The marginal effects indicate that female candidates contesting from reserved constituencies have a 38.6% lower likelihood of winning the election. However, this significance gets diminished once party and geography controls are added, though the coefficient remains negative, indicating some degree of backlash effect. It may seem that in geographically contiguous constituencies, an increase mandated representation at the PRI level for the treatment constituencies, while the status quo is maintained in the control constituencies, leads to a backlash against female candidates.
contesting Lok Sabha seats in the treatment constituencies, even as underlying social norms and perceptions remain largely the same across the two groups.

Running a naive differences t-test on the treatment group shows a significant 1.07 percentage point increase in female candidacy between the 2009 and 2014 elections (Figure 10). However, this is marginally less than the 1.13 percentage point increase seen nationwide (Figure 11). This marginal difference concurs with the results of the regression—the effect of the reservation seems to be insignificant and marginal at best. However, digging into this change further, we see that there is significant fluctuation in changes across the control group, while the change in the treatment states is consistently positive (though of smaller magnitude) (Figure 12). Much of the control states’ outperformance is driven by an almost 10% increase in female candidacy in Uttar Pradesh, which did not implement the reservation but is the biggest state in India and holds 80 Lok Sabha seats.
. // change over time for constituencies that imposed 50% reservation
. keep if policy_50 == 1
(4323 observations deleted)

. ttest female, by(year) unequal

Two-sample t test with unequal variances

<table>
<thead>
<tr>
<th>Group</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>5904</td>
<td>0.675613</td>
<td>0.0032673</td>
<td>0.2510473</td>
<td>0.6711763 0.679083</td>
</tr>
<tr>
<td>2014</td>
<td>6144</td>
<td>0.6752078</td>
<td>0.0034273</td>
<td>0.2686457</td>
<td>0.671569 0.680865</td>
</tr>
<tr>
<td>combined</td>
<td>12048</td>
<td>0.6738412</td>
<td>0.0023707</td>
<td>0.2662149</td>
<td>0.668942 0.776881</td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>-0.010765</td>
<td>0.0047351</td>
<td></td>
<td>-0.0199801 -0.001428</td>
</tr>
</tbody>
</table>

diff = mean(2009) - mean(2014)
$t = -2.2611$
Satterthwaite's degrees of freedom = 12036.6

Ha: diff < 0  Ho: diff != 0  Ha: diff > 0
Pr(T < t) = 0.0119  Pr(|T| > |t|) = 0.0238  Pr(T > t) = 0.9881

Figure 10: Change in female candidacy rates in treatment constituencies

. ttest female, by(year) unequal

Two-sample t test with unequal variances

<table>
<thead>
<tr>
<th>Group</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>8070</td>
<td>0.6693928</td>
<td>0.002829</td>
<td>0.2541367</td>
<td>0.6638473 0.749384</td>
</tr>
<tr>
<td>2014</td>
<td>8301</td>
<td>0.6687132</td>
<td>0.0029599</td>
<td>0.2724105</td>
<td>0.6748522 0.865741</td>
</tr>
<tr>
<td>combined</td>
<td>16371</td>
<td>0.6751329</td>
<td>0.0026063</td>
<td>0.2636136</td>
<td>0.670944 0.779173</td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>-0.0113204</td>
<td>0.0041162</td>
<td></td>
<td>-0.0193805 -0.0032522</td>
</tr>
</tbody>
</table>

diff = mean(2009) - mean(2014)
$t = -2.7502$
Satterthwaite's degrees of freedom = 16341.3

Ha: diff < 0  Ho: diff != 0  Ha: diff > 0
Pr(T < t) = 0.0030  Pr(|T| > |t|) = 0.0060  Pr(T > t) = 0.9970

Figure 11: Change in female candidacy rates across all constituencies

224 of 275
Impact on Vidhan Sabha Elections

Each state in India also holds elections for the state legislature and chief ministership every five to six years. However, these elections cycles are not synchronized and multiple states go to election each year. I sought to estimate a similar regression model for several states that conducted elections before and after the adoption of the 2009 change (Table 4). I carried out these regressions on vote share of candidates, to note change in performance of female candidates. The initial results here are not very promising—the treatment effect seems largely insignificant and there is variation in the sign and magnitude of the coefficient across the states. However, the two models that yield best fit, Chhattisgarh and Madhya Pradesh, also have negative coefficients on the reservation impact, in agreement with my results from the earlier analysis.

Change in Female Vote Share and Candidacy Since 1992
Current literature on the topic of mandated female representation suggests that quotas lead to a backlash effect in the short run (one election cycle) but after ten years the observed effects are positive as voters become acquainted with the competence of female administrators and underlying biases are altered. [15] The results from my regression based analysis are in agreement with the backlash effect though the strength of the effect is hard to isolate due to the several other interacting factors when considering voter attitudes between PRIs and elections to higher levels of government. The effect has been consistently negative even as the magnitude is hard to determine. However, in my earlier analyses I simply consider one succeeding election cycle, that occurred within five years of the change in policy. It may be the case that the backlash effect goes away after several years and female reservations ultimately lead to a new positive equilibrium. The instatement of PRIs in 1992 could thus allow me to examine the impact of reservations over time.

In 1992, mandated female representation was imposed in PRIs across the country, increasing female quotas from 0% to 34%. The impact here would thus also be much more profound than a simple increase in already existing quotas of the kind that occurred in 2009. To examine this theory, I run Bonferroni tests on the aggregate female vote share, female candidacy and female win percentage to note changes across election cycles.

The first general election after the instatement of the PRI occurred in 1996. The simple difference in means indicates that women actually fared worse in these elections (Figure 13), garnering lesser aggregate vote share even as the proportion of female candidates contesting increased significantly by 48.29% (Figure 14). While this is a naive estimate and does not control for many factors, it is telling that the number of female candidates fielded could increase so significantly within one election cycle. This result seems to agree with both the role model effect—female candidates are inspired by seeing other women in strong leadership and are motivated to
contest elections—as well as the backlash effect—even as more female candidates contest they are able to amass even less of the popular vote share. Over time, the female vote share percentage increases significantly, and is up 1.7 percentage points by the 1998 general elections. Interestingly, a similar backlash effect is not seen between the 2009 and 2014 general elections, when female vote share increases by 2.7 percentage points. However, this naive differences-in-mean estimate may simply be picking up general trends towards female empowerment in society. The period between 2009 and 2014 general elections was particularly important for the Indian feminist movement as noted above. Several incidents of sexual assault sparked off nation-wide protests and conversations regarding female empowerment in the country, leading to parties altering their platforms to appeal to sentiments in the electorate.
. oneway vote_share_percentage year, bonf tab

<table>
<thead>
<tr>
<th>Year</th>
<th>vote_share_percentage</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>7.0531286</td>
<td>16.027798</td>
<td>521</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>6.978232</td>
<td>16.325419</td>
<td>543</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>7.6234743</td>
<td>17.028384</td>
<td>544</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>8.7467772</td>
<td>19.448282</td>
<td>543</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>9.3462615</td>
<td>18.508317</td>
<td>543</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>9.0329834</td>
<td>18.079394</td>
<td>543</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>11.722259</td>
<td>19.162195</td>
<td>518</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8.6318482</td>
<td>17.892497</td>
<td>3755</td>
<td></td>
</tr>
</tbody>
</table>

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>8655.38654</td>
<td>6</td>
<td>1442.56442</td>
<td>4.53</td>
<td>0.0001</td>
</tr>
<tr>
<td>Within groups</td>
<td>1193155.59</td>
<td>3748</td>
<td>318.344607</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1201810.97</td>
<td>3754</td>
<td>320.141442</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparison of (sum) vote_share_percentage by Year (Bonferroni)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>1.000</td>
<td>.074897</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>1.000</td>
<td>.570346</td>
<td>.645242</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>1.000</td>
<td>1.76855</td>
<td>1.69365</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>2004</td>
<td>.760</td>
<td>1.000</td>
<td>2.36803</td>
<td>2.9313</td>
<td>1.72279</td>
<td>.599484</td>
</tr>
<tr>
<td>2009</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>2.05475</td>
<td>1.40951</td>
<td>.286206</td>
</tr>
<tr>
<td>2014</td>
<td>2.376</td>
<td>2.68928</td>
<td>2.97548</td>
<td>4.09878</td>
<td>4.74403</td>
<td>4.66913</td>
</tr>
</tbody>
</table>

Figure 13: Change in aggregate vote share of female candidates in constituencies
. oneway female year, bonf tab

<table>
<thead>
<tr>
<th>Year</th>
<th>Summary of (sum) female Mean</th>
<th>Std. Dev.</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>0.64478764</td>
<td>0.89770923</td>
<td>518</td>
</tr>
<tr>
<td>1996</td>
<td>1.1277778</td>
<td>3.5732939</td>
<td>540</td>
</tr>
<tr>
<td>1998</td>
<td>0.51571165</td>
<td>0.74425759</td>
<td>541</td>
</tr>
<tr>
<td>1999</td>
<td>0.53703704</td>
<td>0.80150515</td>
<td>540</td>
</tr>
<tr>
<td>2004</td>
<td>0.66481481</td>
<td>0.87887122</td>
<td>540</td>
</tr>
<tr>
<td>2009</td>
<td>1.037037</td>
<td>1.1192855</td>
<td>540</td>
</tr>
<tr>
<td>2014</td>
<td>1.2407407</td>
<td>1.2474119</td>
<td>540</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Mean</th>
<th>Std. Dev.</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.82495345</td>
<td>1.6449658</td>
<td>3759</td>
</tr>
</tbody>
</table>

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>294.324952</td>
<td>6</td>
<td>49.0541587</td>
<td>18.64</td>
<td>0.0000</td>
</tr>
<tr>
<td>Within groups</td>
<td>9874.49441</td>
<td>3752</td>
<td>2.63179489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10168.8194</td>
<td>3758</td>
<td>2.70591255</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bartlett's test for equal variances: chi2(6) = 2.7e+03 Prob>chi2 = 0.000

Comparison of (sum) female by Year (Bonferroni)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>.48299</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>-.129076</td>
<td>-.612066</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>-.107751</td>
<td>-.590741</td>
<td>.021325</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.000</td>
<td>0.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>.020027</td>
<td>-.462963</td>
<td>.149103</td>
<td>.127778</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.000</td>
<td>0.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>.392249</td>
<td>-.090741</td>
<td>.521325</td>
<td>.5</td>
<td>.372222</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.003</td>
</tr>
<tr>
<td>2014</td>
<td>.595953</td>
<td>.112963</td>
<td>.725029</td>
<td>.703704</td>
<td>.575926</td>
<td>.203704</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.822</td>
</tr>
</tbody>
</table>

Figure 14: Change in female candidates contesting as a proportion of total candidates
Shortcomings

My main regression model considers the 2009 natural experiment which involved an increase in quota from 34% to 50%, and not an instatement of quotas from scratch. The observed impacts on voter behavior may thus appear relatively muted since some mandated exposure to female policymakers has already occurred due to the 34% reservation that prevailed from 1993 till the policy change in 2009. After accounting for this fact, my paper seems to address the question of whether an increase in mandated representation further alters voters’ perceptions and prevailing social norms. As per the exposure and role model effect theories, and corroborated by the results yielded by analyses, I hazard a guess that change in voter behavior due to an increase in quotas is incremental at best.

This paper’s biggest shortcoming is that I was not able to study the 1993 natural experiment when quotas for women went from 0% to 34% across the entire country. This would have allowed for a more direct comparison of no-quota versus quota scenarios, as well as allowed me to consider more election cycles to note the long-term impact of mandated representation. However, I faced several limitations in carrying out such an analysis. Firstly, the 1993 policy change involved all states and constituencies in India. This makes it difficult to isolate the causal treatment effect from underlying societal trends towards female empowerment that would have existed regardless of the quotas. Secondly, on a more logistical note (and an issue that can surely be solved with adequate resolve), I could not conduct more long-term analyses since I did not have constituency level data for variables such as literacy rate, sex ratio, female infanticide rates, female workforce participation etc., that would have affected attitudes female candidates as well. I was able to do away with this data in the 2009-2014 analysis under the assumption that underlying variables would not have changed significantly in a five-year time span. However, any analysis exceeding
ten-year or more than three election cycles, must need to control for these socio-economic variables.

Further precision could have been achieved by mapping PRI districts to Lok Sabha constituencies and accounting for the exact areas that had female reservation and during which election cycles. Female reservation is randomly allocated across PRI seats and often changes between election cycles. My analysis aggregated such effects at the state level and thus yields a lower magnitude of the treatment effect. This would also allow the analysis to be in line with existing literature, which has noted that voters become more amenable to voting for female candidates if their PRI seat has been reserved for women over two election cycles. [15]

Conclusion

In this paper, I sought to analyze the impact of an increase in mandated female representation in PRIs, on the performance of female candidates in elections to the Lok Sabha, which does not have similar quotas for women. I found largely insignificant or slightly negative effects of female reservation on performance of female candidates in Lok Sabha and Vidhan Sabha elections immediately succeeding the imposition of these reservations in 2009. However, my analysis does not account for longer run changes in voter attitudes as a result of reservation and this forms a compelling topic for future research.

While the impact of quotas on voter behavior remains unclear, the positive benefits of women in policy making, especially in societies where they are historically underrepresented, has been empirically proven. India intermittently sees calls for reservation at higher levels of government as well, where female representation remains dismal. Even as female representation may signal the progressive nature of a society, there is some evidence of a backlash effect that may
go away (or is simply difficult to assess over time). In this context, a concept from psychology appears particularly pertinent—moral licensing. The term describes a phenomenon in which “past good deeds can liberate one to engage in behaviors that are immoral, unethical or otherwise problematic, behaviors that one would otherwise avoid for fear of appearing immoral.” [14] Pop psychologist Malcolm Gladwell describes moral licensing in the context of the feminist movement, noting the sexist backlash against Australia’s first female Prime Minister Julia Gillard after her election. [7] He argues that by electing a female head of state, the Australian populace had proved its progressiveness and then felt free to go back to its sexist ways. While Gladwell’s argument is fraught and several other factors, especially political ones, may have played into Gillard’s criticism, this critique did manifest itself in extremely sexist terms. Phenomena like moral licensing may thus negate any short run positive gains to be made from mandated female representation. There is hazard in arguing the quotas can solve the problems of regressive social norms—mandating 34 to 50% quotas in communities where women take up less than 10% of political seats can lead to unforeseen effects—and reduce the quota to a form of token, statistical representation for women. However, the initial hesitancy and backlash can lead the way to a positive equilibrium and more equitable society in the long run, achieving change that would have otherwise taken decades to attain organically. The question of whether quotas are indeed the best means to increase female representation thus remains open to debate.
References


Appendix

Figure 15: Cases of dowry deaths reported across the country
Figure 16: Geographically contiguous constituencies across states that did and did not implement 2009 amendment.
Figure 17: Normalization: Distribution of vote_share_percentage
Figure 18: Normalization: Distribution of \( \text{vote\_share\_percentage} - \overline{\text{vote\_share\_percentage}} \)
Figure 19: Distribution of total vote shares in constituencies
. oneway fwin year, bonf tab

<table>
<thead>
<tr>
<th>Year</th>
<th>Summary of (sum) fwin</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>1991</td>
<td>.07677543 (521)</td>
<td>.26649066</td>
</tr>
<tr>
<td>1996</td>
<td>.07918969 (543)</td>
<td>.27028359</td>
</tr>
<tr>
<td>1998</td>
<td>.08272059 (544)</td>
<td>.27571295</td>
</tr>
<tr>
<td>1999</td>
<td>.09392265 (543)</td>
<td>.29199007</td>
</tr>
<tr>
<td>2004</td>
<td>.08655617 (543)</td>
<td>.28144284</td>
</tr>
<tr>
<td>2009</td>
<td>.106814 (543)</td>
<td>.30916143</td>
</tr>
<tr>
<td>2014</td>
<td>.1160221 (543)</td>
<td>.32054672</td>
</tr>
<tr>
<td>Total</td>
<td>.09179894 (3780)</td>
<td>.28878012</td>
</tr>
</tbody>
</table>

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>.70716674</td>
<td>6</td>
<td>.117861124</td>
<td>1.41</td>
<td>0.2050</td>
</tr>
<tr>
<td>Within groups</td>
<td>314.4386</td>
<td>3773</td>
<td>.083339147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>315.145767</td>
<td>3779</td>
<td>.083339358</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bartlett's test for equal variances: $\chi^2(6) = 31.9376$  Prob>$\chi^2 = 0.000$

Comparison of (sum) fwin by Year (Bonferroni)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.005945</td>
<td>.003531</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.017147</td>
<td>.014733</td>
<td>.011202</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.009781</td>
<td>.007366</td>
<td>.003836</td>
<td>.007366</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.030039</td>
<td>.027624</td>
<td>.024093</td>
<td>.012891</td>
<td>.020258</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.039247</td>
<td>.036832</td>
<td>.033302</td>
<td>.022099</td>
<td>.029466</td>
<td>.009208</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Figure 20: Change in female win percentage
Figure 21: Marginal effects for probit coefficients in Model (4), Table 2
Notes:

1. 85 (15.6%) and 48 (8.9%) seats of 543 constituency Lok Sabha are reserved for candidates from Scheduled Castes (SC) and Scheduled Tribes (ST) respectively. SC and ST groups comprise 16.6% and 8.6% of the country’s population, as per the 2011 national census. [1]

2. The entire country participates in general elections to the national Lok Sabha, each state has an assembly election to a Vidhan Sabha and every municipality has its own Panchayats.

3. In the same year, the 108th Amendment to the Constitution, proposing 50% reservation of women in the Lok Sabha, failed to pass.

4. However, the authors observed no changes in the female roles and participation in the labor market, arguing that the change in attitudes was primarily due to a “role model effect”.

5. I only include dummies for the INC and BJP here since these are the two dominant parties that contest every seat in the Lok Sabha.

6. Publicly available data can only be obtained from the national census that occurs every ten years (2001, 2011, 2021) and is available at the state-level.

7. States that have implemented reservation Assam, Andhra Pradesh, Bihar, Chhattisgarh, Gujarat, Himachal Pradesh, Jharkhand, Kerala, Karnataka, Madhya Pradesh, Maharashtra, Odisha, Rajasthan, Sikkim, Tamil Nadu, Tripura, Uttarakhand, West Bengal”

8. amounting to about 25,000 INR or 350 USD for the Lok Sabha elections, and lesser sums for state assembly elections.

9. As a sanity check, I looked over the distribution of total sum of vote shares in constituencies (which in theory should be 100% in all constituencies). An overwhelming majority of constituencies report 100% votes, with some falling in the 97-99% range (Figure 19). The discrepancy can probably be attributed to rounding or minor measurement errors—the modification of the vote share variable in practice thus matches the theory.

10. Models 3, 6, 9 include state-wise fixed effects.

11. Position signifies rank so a positive coefficient indicates worse performance (i.e. lower rank) as the first ranked candidate is the winner.
Empowerment, Resistance and the Birth Control Pill:
A Feminist Analysis of Contraception in the Developing World

Abigail Trombley
Abstract

As access to contraception and family planning becomes more integral to understandings of socioeconomic development, discourse surrounding the topic frames women’s ‘unmet need’ for contraception as a problem affecting poor, traditionally feminine, third world women. This paper will challenge this classic liberal discourse by examining notions of power and capitalist goals of development, and through an analysis of twelve major NGOs and IGOs. It will use notions of “problem framing” and “solution framing” to construct women in developing countries as active subjects capable of choice and creating change. This alternative view examines the implicit binary of modern, western women with contraceptive access, and poor, tradition-bound, third world women, receiving contraceptives as a gift in order to generate a more productive population, thus reinforcing the capitalist systems that disenfranchise women.
Table of Contents

Introduction 246
Methodology 249
Discursive Framing of Contraception in the Developing World 250
Theory in Action: Contraception and Reproduction in Post-Colonial Puerto Rico 262
An Alternative Vision: Women as Active Subjects 266
Contraception as the “Master’s Tool” and Concluding Thoughts 270
Notes 273
Introduction

As early as the 1920’s scientists made the transformational discovery of the ability of synthetic hormones to prevent ovulation. It was not until the early 1960’s that the contraceptive pill, designed for women to control their fertility, entered the market and fundamentally impacted the lives of women and the family structure.¹ The narrative of the development of the contraceptive pill focuses on timelines of biochemical advances and efforts of feminist pioneers such as Margaret Sanger.² However, the mainstream discourse obscures the notable timing of the development of the contraceptive pill as occurring concurrently with U.S. involvement in the developing world.³

As the population of a nation became seen as a vital resource to the state, greater attention was placed on the control of which segments of the population were reproducing. As population control soon became seen as key to a nation’s modernization and prosperity, contraception became a favored tool of development in the global south. The contraceptive pill that had the ability to revolutionize women’s control over their bodies immediately became intimately entangled with issues of economic development, the modernization imperative, and colonial control.

Since then, the advent of contraception is commonly understood as a benchmark in the development of the socioeconomic well-being of women and their families. The majority of studies on the use of contraception focuses on its frequently documented connection to socioeconomic development. However, within the literature there are contrasting conclusions when delineating the relationship between contraception use and women’s autonomy in fertility decisions; some studies show that increased education and economic independence are associated with increased use of contraception.⁴ Other studies counter this; they have shown that even when women’s autonomy is increased via socioeconomic development, men often remain in control of decisions regarding fertility due to pervasive traditional norms.⁵ Therefore, studies on the
connection between economic development and women’s autonomous choice to use contraception are far from conclusive; nevertheless, the presentation of contraception as a socioeconomic tool of development holds a consistent presence in the discourse regarding contraception in the developing world. Alternatively, an analysis informed by a variety of feminist theories allows the use of contraception to be viewed as women’s assertion of autonomy over their reproduction, and therefore as an act of resistance in a patriarchal society.

This alternative view of contraception is essential for recognizing the potential of contraception to allow women to reclaim their agency and transform the societies they live in. Yet, this is not the way contraception has been presented in discourse. Women’s decisions to limit their reproduction have been written about in regard to demographic, health and economic concerns and consequences. However, a feminist analysis illuminates an alternative viewpoint that examines the resistant nature of a woman’s decision to control her fertility, the conditions that determine whether she has the ability to do so, and the implications for making this choice in societies in which woman’s autonomy over their fertility is resistant to the patriarchal order. This examination of contraceptive use by women in the developing world utilizes postcolonial feminist theory as the starting point for analysis. Postcolonial feminist theory destabilizes traditional theories in which the colonized subject is absent or rendered inferior, shifting focus to the experience of the colonized to articulate and understand the subject’s lived experience.

Informed by postcolonial feminist theory, along with insights of other feminist theories which have enriched and evolved postcolonial feminist thought, I argue that discourse around contraceptive use in the developing world must construct women as active subjects and include a discussion of women’s individual agency to resist cultural norms in the choice to use contraception. While contraception initially presents as an opportunity of liberation for women, I
analyze discourse from organizations that advocate for women’s use of birth control in the developing world, as well as deliver contraceptive services themselves, in order to uncover the dominance of liberal, capitalist assumptions therein. Postcolonial feminist scholarship has exposed such assumption for being, at best, inapplicable and, at worst, further oppressive. I examine the implications of these assumptions within the context of contraception as an act of resistance by an active subject.

I argue that the discourse of economic development stems from a liberal feminist perspective that focuses on economic empowerment as a crucial element of gender equality. This is significant according to a Foucauldian conception of discourse used in this analysis, which reveals discourse to be a mechanism of power, possessing the ability to affect the subject whom is objected to it. This narrow focus neglects broader elements of a woman’s being. It privileges a western, capitalist perspective that ties women’s value to capital production and accumulation, further binding them to the patriarchal capitalist system that historically has oppressed women. Further, it perpetuates the binary between modern, “liberated,” western women who use contraception, and traditional, oppressed women for whom contraception is delivered as a gift from the western saviors in their mission to “improve” the developing world by achieving goals informed by progressive values. The aim of this study is not to deny that contraception is a fundamental human right, as well as an important element of women’s reproductive and maternal health. Rather, the aim is to uncover a problematic binary in which contraception is marshalled as a defining marker between the “first world” and “third world” woman. Discourse around contraception in the developing world informed by a feminist analysis must present women as active subjects, whose choice to control her reproduction in a world in which fertility choices are dominated by men is an act of resistance.
Methodology

In order to expose the liberal, capitalist assumptions in the discourse surrounding contraception in the developing world, I engage in a critical discourse analysis of published material from eight non-governmental organizations (NGOs) and four intergovernmental (IGOs) which provide contraception to women in the developing world (see Appendix). The overrepresentation of NGO’s in this selection is a telling representation of the “NGO boom” since the 1980s. This reflects the way hegemonic neoliberalism has diminished the role of the state and government-related institutions, placing greater trust in private organizations such as NGOs. NGOs are viewed as market-based actors with greater capacity for efficiency than government-related IGOs. I selected the data set based on the following criteria: each organization has a country of origin in the global north, the majority or all of their “target” countries are in the global south, and their services include the distribution and/or education of contraceptive methods. These contraceptive methods may include oral contraceptives, long-acting reversible methods such as the IUD, and abortion, but do not necessarily include all of these methods. The material for the analyzed discourse includes published material released by the organization. Most of this material was sourced from the website of the organization and includes official published reports as well.

I divide this published discourse into two broad categories which I have labeled “Problem Framing” and “Solution Framing.” Problem framing refers to discourse used to name and describe the issue of unmet need for contraception in the developing world. Solution framing refers to the discourse describing the organization’s proposed solution and rationalization for its efficacy and appropriateness. The analysis is attentive to the three levels of discourse analysis identified by Norman Fairclough: discursive practice, text analysis, and social practice. Discourse practice will inform my understanding of which discourses are called upon to inform these organizations’
perspectives. Text analysis will illuminate how this discourse is translated into agenda-setting. Lastly, analysis of the social practice will reveal the impact of this discourse on the construction of social and cultural relations and structures in the developing world.

Problem framing discourse falls under Fairclough’s category of discursive practice as it explains how the problem presents in the minds of those thinking about and working on it. Solution framing discourse provides the material for text analysis as it describes how the proposed solution is translated in the work of these organizations. Solution framing is also the material for analysis of social practice as the broader cultural impact of these policies and programs is considered. Fairclough’s method frames my consideration of the discourse as it is analyzed with a feminist analysis drawing upon a range of contributions to feminist theory. The following section discusses the problem and solution framing of contraception using Fairclough’s tools of analysis. The next section exemplifies these findings through the case of contraceptive delivery in Puerto Rico. Following, I discuss the alternative conception of contraception in the developing world that is centered around women as active subjects, for whom contraception is a tool of resistance in a patriarchal society. I conclude with summarizing remarks and a call for reconceptualization of contraception to guide further policy.

Discursive Framing of Contraception in the Developing World

Problem Framing

In this section I use Fairclough’s discursive practice to investigate how the lack of contraceptive access in the developing world is framed as a problem, by exposing the narratives relied upon to explain this issue in the discourse of western organizations. A lack of available contraception is referred to in the literature as women’s “unmet need” for contraception. Unmet
need is defined as “nonuse of contraception among women stating a desire to avoid pregnancy.”

As a postcolonial feminist, Mohanty problematizes western feminists’ construction of the “third world woman” as a homogenous category with uniform challenges and desires. Prevalent in the discourse is naming of “lack of access” to contraception, with the singular and universal solution being access or “met need.” In this way, these organizations’ conception of unmet need contributes to the “discursive homogenization and systemization of oppression” of developing world women by reducing the explanation of women’s fertility decisions to contraception access.

A frequent theme in the discourse revolves around what the World Health Organization describes as “cultural or religious opposition” to contraception in the developing world. Population Council also refers to “family or community opposition to family planning.” Similarly, the Bill and Melinda Gates Foundation speaks of “cultural and knowledge barriers.” While cultural perspectives on the use of contraception do vary, it is problematic to organize those variations along a binary that perpetuates what Mohanty describes as the “third world difference.”

Eurocentric assumptions construct a binary between the western and third world woman described by Mohanty:

This average third-world woman leads an essentially truncated life based on her feminine gender (read: sexually constrained) and being ‘third world’ (read: ignorant, poor, uneducated, tradition-bound, religious, domesticated, family-oriented, victimized, etc.) This, I suggest, is in contrast to the (implicit) self-representation of western women as educated, modern, as having control over their own bodies and sexualities, and the ‘freedom’ to make their own decisions.

This binary is relevant to the issue of contraception when the problem is reframed as women’s reproductive control as a tool of economic empowerment and therefore indicator of socioeconomic development. Unmet need, in these organizations’ discourse, is framed as a
problem of cultural “backwardness” which opposes modern participation of women in the capital labor market. A capital-producing, industrialized, liberalized and globalized society is framed as the universal progressive ideal, and the problem rests in unmet need for contraception as an obstacle in society’s inevitable trajectory.

Marxist feminist theory is useful in understanding how this framing is problematic as it explains how the sexual division of labor produced by capitalism and industrialization contributes to women’s subordination. For Mohanty, a postcolonial critique of capitalism draws upon Marxist feminist theory to explain the way these phenomena situated wage labor as the “real ‘productive role’” for men. The housewife is relegated to a “maintenance” role which produces no exchangeable value in a capitalist society. This strand of feminism points to the root of this binary in capitalist structures which historically have oppressed women by confining them to the domestic sphere, where they cannot produce or acquire capital. In supplement, postcolonial feminism explains the necessity of historical and cultural context, which further complicate the narrative put forth in the problem framing by western organizations. For example, it neglects to explain the way being the head of the household may not be an oppressive role for women, instead a choice not previously available to them. Therefore, the problem of unmet need is a reductionist explanation which positions “met need” as the gateway to modern societal development. This imagination should be questioned in light of the way it further entrenches the creation of the monolithic third woman in the developing world.

Another way in which the problem framing of unmet need for contraception in the developing world is problematic is the way it creates what Mohanty calls “politics of location.” Politics of location refers to the “historical, geographical, cultural, psychic and imaginative boundaries” which substantiate political definitions and self-definitions for western feminists.
Politics of location is constructed frequently when feminists are working cross nationally and across racial differences and unintentionally reproduce internal power hierarchies. This occurs in congruence with a universalization and essentialization of third world women’s needs, desires and experiences, as well as the patriarchal societies they inhabit. Consider UNESCO’s International Planned Parenthood Foundation’s characterization of “misinformation, prejudice and gender discrimination” as well as “violence” at the “community level.” Marie Stopes International also writes of reducing the “stigma” in these communities and the Bill and Melinda Gates Foundation also speaks of “cultural and knowledge barriers.” Pathfinder International states explicitly the need to “talk to male religious leaders, and other key players to eliminate all biases and misconceptions.” The Population Council describes the problem of women having to “face family or community opposition,” similar to the “cultural or religious opposition” cited by the World Health Organization.

These characterizations prove problematic in several ways. Firstly, they construct a universal depiction of societies in the developing world which are community or kinship based, bounded by tradition, and subject to religious domination. Secondly, these “backward” communities are constructed in contrast to the liberal and liberated world occupied by western feminists, along with the staff of these organizations. In this way, Mohanty’s politics of location are in full force. Politics of location are problematic in the way they rely upon collectivities based on broad identity categories, eliminating both cultural and historical relativity, as well as the agency of the individual. Mohanty argues that “historicizing and locating” this agency is a necessary alternative to narratives of gendered oppression which rely upon universal characterizations and assumptions. In the case of access to contraception, the act of resistance taken by a woman choosing to control her reproduction via contraception in a world dominated by
patriarchy should be deemed as so. Concerningly, this notion is lost when subsumed under the grand narrative of contraception as a tool of development to lead developing societies out of the “backward,” traditional conditions described in this discourse. The reconceptualization of reproductive control as an act of resistance is explored in later sections.

In summary, problem framing provides the substance for what Norman Fairclough labels discursive practice. Discursive practice examines the way the text is produced, specifically which discourses are called upon in its production. In the discourse examined, the problem of unmet need for contraception draws upon a tradition of thought which envisions “progress” and “modernization” as achieved through liberal conceptions of the individual, community and the state. This is evident through the binary constructed between modern, liberated women in the western world who have access to contraception, in comparison to oppressed women in the developing world who are denied access to contraception due to “backward” illiberal cultures. This viewpoint stems from these organizations’ origins in the western world, in which liberal, modern, industrial, secular and capitalist societies are viewed as the goal of historical and economic evolution.

Solution Framing

Solution framing explores the ideas put forward to address the issue of unmet need for contraception. I utilize Fairclough’s lenses of text analysis and social practice to examine the relationships between the producers of the discourse, subjects of the discourse, and the actions these relations are translated into. A feminist analysis often draws upon Foucault’s ideas regarding the production of power as Foucault offers an alternative to resource theories of power viewed as something men possess, and women are suppressed with. Instead Foucault focuses on how power
is produced through knowledge and discourse. Foucault describes objects of power, those whom power is wielded upon, who possess concepts which describe their distinctive character traits. In our setting of contraceptive delivery, the objects of this setting are women who live in developing countries. The concepts relating to these objects can be recalled in Mohanty’s description of the “Third World Woman”–a woman who is distinctly traditional, uneducated, power, oppressed, religious and domestic. For Foucault, modes of authority represent the subjects which wield exercised power. In our setting, modes of authority consist of the organizations which promote contraception use in the developing world: non-governmental and inter-governmental organizations that are involved in the production and distribution of contraceptive tools. Importantly, these are the groups which have authority in determining the moral standard regarding reproduction. This means these groups, which represent significant economic, political and cultural power globally, uphold the belief that controlled reproduction is essential to liberation and prosperity.

Lastly Foucault names the lines of strategic action which define the subjects through which power is exercised. Foucault sees power as productive, meaning it is able to produce subject identities, whom internalize their position. In this sense, this is a power which figurately and physically transforms people. In our setting, lines of strategic action consist of the ways in which contraception is promoted in the developing world: the conception of contraceptive methods, their distribution and advocacy in communities it is brought into. Foucault conceptualizes power as a force that is deployed to facilitate specific outcomes, processes and practices. Further, it may be deployed by both dominate and subordinate forces which both possess agency, even though they both experience outcomes of exercised of power. This provides a framework to view the way in which developing world women’s bodies are regulated through fertility control.
Foucault believed that sexuality in modern society has a close connection with structures of power, and his “repressive hypothesis” argues that sexuality tends to be oppressed. It is important to not equate the repression of sexuality itself with contraception. However, the repressive hypothesis is useful in understanding how the control of fertility became a desired standard in a modern, liberal society. Nevertheless, contraception is considered here as a tool of that Foucault labels “biopower.” He explains biopower has two levels: the first exists on the individual level, on which political significance is placed in the defining of medical norms of a healthy body. In the case of reproduction, “health” is a major theme of the discourse. Medical norms defined by western organizations are translated into controlled fertility. The World Health Organization indicates this desirability through its discourse of “[limiting] size of families, “invest more in each child” and “slowing unsustainable population growth and the resulting negative impacts on the economy.” USAID writes that contraception “mitigates the impact of population dynamics on natural resources and state stability.” This discourse reveals limited fertility as desirable according to this western organization’s values and dovetails into the second level of biopower.

This level exists at the national level, at which the nation’s population is viewed as a “resource” that the nation holds the duty of safeguarding. In the case of contraception in the contemporary era, the conception of the “nation” as used by Foucault is eroded by the primacy of transnational organizations and groups as the subjects wielding biopower in a political era of dominant neoliberalism (rather than statism). However, given these groups’ unilateral national origins in western countries, the national disparity between countries promoting contraception and those nations which become the objects of biopower holds some national significance.

The significance of biopower is the way in which objects of its disciplinary control internalize its projected norms, resulting in a monitoring of their own behavior and a self-
formulation of subjects of their own knowledge. This is the process occurring in the solution framing of discourse around contraception. Foucault himself was hesitant of sweeping teleological narratives and instead closely examined tools masked as “inventions of efficiency” which, in reality, serve to form a new system of control, the objects of this control being human bodies. The discourse of “efficiency” is prevalent in the solution framing of contraception. This discourse resembles elements of what Deborah Stone describes the “market model” of policy. Stone argues that the discourse of economics and society as a market is hegemonic in the formulation of policy. The problem with the market model, Stone describes, is the way it reduces relationships to individuals trading with one another in pursuit of maximum individual gain. In this way, the market model is born out of the liberal conception of individualism, the ideological imaginary that dominates the western world.

As stated, the discourse around contraception consists of the language of market efficiency. Consider the language which forms the Bill and Melinda Gates Foundation’s solution framing around family planning in the developing word. For example, the Foundation places a great emphasis on the cost-effectiveness of contraception, stating it is one of the “most cost-effective investments a country can make;” citing the fact that “every dollar spent on family planning can save governments up to 6 dollars.” PSI echoes the assertion of family planning as a “proven and cost-effective solution.” In this case, we can directly see the imagination of contraception as the national level of biopower described by Foucault. Contraception is framed as a market-friendly solution to aid the government in achieving this level of social control.

The discourse analysis reveals an emphasis on “choices” in the method of contraception. Marie Stopes International advocates offering women the “widest range of methods” while “reducing restrictions.” PSI speaks of “commodity stockouts,” the need to “create demand for
family planning and improve service delivery, “[utilize] existing wholesale and retail distribution infrastructure” and “[create] franchised networks.” AWID writes of the importance of “ensuring access to RH [reproductive health] commodities they [women] want when they need them.” Additionally, PATH speaks of “innovative financial models” that focus on digital health and data to, as the Population Council would say “prepare the market” for contraception in the developing world.

These examples illustrate the concepts of discourses around contraception to be market-centered. This is a result of contraception discourse of organizations originating in the western world’s reliance on capitalist, neoliberal principles. Attentive to Fairclough’s level of analysis “discursive practice,” it is necessary to understand which discourses are relied upon to produce the discourse utilized by the organizations. As a postcolonial feminist, Mohanty writes that careful attention must be paid to cultural and historical contexts. The western and capitalist context utilized to produce this discourse is then translated to the methods utilized to wield power. Cooper adds “ideology” to the modes of power of knowledge and discipline that Foucauldian feminists emphasize. In this example, ideology is a powerful force in generating outcomes and processes, as it informs the goals and strategies of organizations aiming to deliver contraception. The danger of this lies in its neglect of a postcolonial feminist analysis which would reveal that the privileging of western, capitalist ideology ignores the cultural and historical context in which this contraceptive delivery is taking place, as well as reinforces the hierarchy in which this ideology is deemed superior.

Feminist writers have critiqued an emphasis on “choice” as it “hides the operations of power that construct choices” and neglect to acknowledge or address the patriarchal structures which perpetuate oppression. Additionally, it emphasizes a women’s ability to make a “choice,”
while denying the way that the selection was created without that woman’s input.\textsuperscript{52} A study in Oman, where the government has provided free contraception to all married couples since 1994, found that fertility rates did not decrease, as patriarchal culture structures continued to dominate fertility decisions.\textsuperscript{53} In this case, a misplaced emphasis on the ability to “choose” contraception obscured the root cause of women’s lack of fertility control. In consideration of Fairclough’s text analysis which analyzes how discourse is translated into agenda-setting, a rhetoric that relies heavily on choice will make little progress in achieving women’s reproductive liberation.

These sentiments are evident again in the discourse’s desired outcomes of increased use of contraception in the developing world. Pathfinder International begins generally by stating that contraception means “better health, more girls in school, greater resources for families, communities and nations.”\textsuperscript{54} PSI echoes by stating that “women with the ability to control their fertility have better access to education and employment opportunities, bringing economic benefits and improved livelihoods.”\textsuperscript{55} USAID writes that contraception “[improves] women’s opportunities for education, employment, and full participation in society” and “reduces poverty by contributing to economic growth at the family, community and national levels.”\textsuperscript{56} The World Health Organization states that access to contraception is essential for “supporting development” and an “opportunity for women to pursue additional education and participate in public life, including paid employment.”\textsuperscript{57}

The element of this discourse of particular consequence rests on its belief in contraception as a vehicle to encourage women’s participation in the labor market. Kalpana Wilson argues that liberal feminism has evolved into “neoliberal feminism” in the area of development. “Women in Development” (WID) approaches translated liberal feminism’s focus on “including women” to women’s inclusion in the production and accumulation of capital via the labor market in
development models. Wilson writes that WID initiatives “focused on women’s education, training and access to technology which would make them more productive and improve their access to the market.” Clearly, contraception is viewed as a line of strategic action to accomplish this aim. Evident in the above-mentioned discourse is the promotion of contraception of a tool that allows women to stay in school and participate in paid labor activity with the ultimate goal of economic development. Wilson notes this approach rests on the neo-classical economic assumption that “individuals are utility maxing and economic growth comes from the exercise of individual choice…supported by the free market” (Stone’s market model).

Postcolonial and Marxist feminism informed the subsequent “Women and Development (WAD) movement, which challenged WID’s approaches by pointing how women’s exploitation is an inherent element which sustains the process of capitalism and imperialism. Later, Gender and Development (GAD) offered an even more comprehensive counter to WID by explaining how the emphasis on inclusion of women in the paid labor market reproduces gendered social inequalities. This is because WID approaches only recognize labor performed in the capital labor market, ignoring labor produced by women in the domestic sphere as it does not result in the production of capital. WID reproduces these relations by assuming equal gender distribution in the household in which each individual seeks to maximize individual utility.

WID approaches seek to transform the domestic, uneducated, powerless “third world woman” through market processes and contraception is a line of strategic action to accomplish this aim. We return to Foucault’s claim that these inventions of efficiency create a new system of social control which ultimately exerts power over human’s bodies. Several studies have complicated the linear narrative that contraception directly produces autonomous benefits for women, and instead supports GAD’s assertion that gendered social inequalities exist intra-
household. For example, Mason points that in some societies, children are essential social capital for women. This is especially true when a woman marries into extended-kin households, where they cannot gain security or respect until she has a child. Additionally, lack of fertility control is not always related to social and/or economic oppression; in western African societies, women are expected to be both self-supporting and have high fertility desires. As mentioned previously, a study in Oman showed that traditional and community influences remained strong, despite access to free contraception. The study used a bivariate link to also measure women’s education and employment [economic autonomy]. The fertility rate in Oman remained high, revealing that autonomy doesn’t always increase women’s choice in contraception in the face of remaining strong traditions of men making decisions regarding fertility.

The significance of these studies is the way they complicate the narrative embraced by the western organizations examined here and the WID discourse they rely upon. Solution framing represents both Fairclough’s text analysis and social practice. Text analysis, through wording, metaphors and other linguistic characteristics, explains the relationship and specifically interactional control between the creator and subject of the text. Further, it illuminates how these texts understand events and social relations and construct their own version of reality perception of these subjects. We have established the connection between these discourses and the power they wield in establishing these western organizations as modes of authority. Finally, social practice explains the relationship between the discourse itself and the way in which it is disseminated, informing social and cultural relations and structures. We have explored how these discourses produce biopower which is used to exert control of women in the developing world’s bodies, the impact of which is explored in the following section. This use of biopower occurs within the
practice of contraceptive delivery which masks the way in which power is exerted through subsuming it under the practice of development.

**Theory in Action: Contraception and Reproduction in Post-Colonial Puerto Rico**

The case of post-colonial Puerto Rico is useful in demonstrating the *problem framing* of overpopulation and the *solution framing* of contraception as a socioeconomic tool of development. Briggs explains the history of contraceptive delivery in Puerto Rico which occurred concurrently to Puerto Rican nationalist struggles against colonialism and U.S intervention. She introduces two nationalist parties holding opposing positions on the use of contraception in Puerto Rico. The first was the nationalist movement of Albanizu, which relied upon Catholic values and rhetoric to frame contraception as an attempted genocide of Puerto Ricans by the American missionaries and aid workers which were promoting and distributing it. This group held up motherhood as a glorified symbol of nationhood and Puerto Rico itself.66

The opposing nationalist group consisted of liberal professionals; these were Puerto Rican and North American nurses, social workers, and feminists who contrastingly saw contraception as key to modernizing the nation and increasing its inhabitant’s wealth and quality of life.67 Briggs describes this group as “feminists equally possessed of a modernizing nationalism,” and birth control was central to this vision.68 Like American feminists, this group sought to connect maternalism, the state and social advancement; they promoted birth control as crucial for producing healthier and wealthier families in which “healthy mothers are able to educate a small number of children and transmit values of progress and modernization.”69 These feminists set up birth control clinics which targeted working-class women.
The efforts to improve the well-being of working-class Puerto Rican women through reproductive health and choice may seem benign, and a preferable alternative to Albanizú’s glorification of motherhood and female fertility as central to the nation. However, the dark side of this modernizing, liberalizing project was its support of “progressive eugenics,” which the movement saw as positively linked to infant and maternal health.\textsuperscript{70} Progressive eugenics became another tool for the colonial gaze, that saw the control of native women in their colonies to be essential to successful rule, in particular women’s forced adoption of the imperial country’s cultural values.\textsuperscript{71} Progressive eugenics supported these efforts through restricting the reproduction of native women in the name of development and modernization.\textsuperscript{72}

At this point, it is necessary to examine the origins of contraception in the developing world, which was the practice of eugenics. In the mid-19\textsuperscript{th} century, the British Malthusian League began to advocate for birth control as a remedy for excess reproduction which they viewed as a cause of poverty. Briggs describes how this study of population democracy emphasized “class, racial, and geographical differences with respect to reproduction.”\textsuperscript{73} Specifically, Malthusians were concerned with rising fertility rates in non-Anglo-Saxon, non-white populations. In Puerto Rico, the term “overpopulation” was framed as a problem of excessive sexuality and fertility among working-class Puerto Rican women. This \textit{problem framing} of overpopulation saw it as responsible for poverty, delinquency, homelessness, prostitution, and disease.\textsuperscript{74} In Puerto Rico, Clarence Gamble (the heir to Proctor & Gamble at the time), actively pursued a eugenics project through which he sought to limit reproduction among working-class and non-Anglo-Saxon women.\textsuperscript{75} In the 1930s, sterilization become available, and the Eugenics Sterilization Board authorized the use of involuntary sterilizations, with the clear goal of limiting reproduction among poor, dark-skinned Puerto Rican women.\textsuperscript{76}
The discussion of race as central to the eugenics movement is critical. Black feminism has drawn attention to the way in which women of color experience a “double-discrimination” of race and gender. Crenshaw coined the term “intersectionality” to name how this combination of racial and gender-based oppressed is an “intersectional experience greater than the sum of racism and sexism.” Further, an intersectional approach criticizes mainstream feminism for treating race and gender as exclusive categories, arguing this further marginalizes the experiences of women of color. The contributions of critical race feminists such as Crenshaw are crucial in examining the racism integral to reproduction control. This is evident in publications released on the topic in the 1930s. For example, an article titled “Overpopulation or Underpopulation? - A Review of Conflicting Opinions” was published in the *Journal of Heredity*; the author Guy Burch concluded that the “worst quarter” of the population was growing, and native-born whites were decreasing in numbers. Similarly, Lothrop Stoddard wrote in *The Rising of Color* that high fertility rates among non-“Nordic” and non-“Anglo-Saxons” were a threat to society. These titles represent a few among many examples of how the issue of race was central to fears about rising population and poverty among non-white population.

An intersectional feminist analysis reveals that both sexism and racism intersect with other forms of inequality. The issue of reproductive control could not be more exemplary of the way in which racism intersects with classism and sexism. As intersectional feminism illuminates, women of color experience concurrently various forms of oppression as a “complex interaction of race, gender and class that is more than the sum of its parts.” Roberts, a black feminist, argues a harmful mythology exists which devalues black women as mothers. She explains this myth has developed and perpetuated in three ways. The first was the condition of slavery, in which black women were denied reproductive autonomy, as reproduction had capital value to slave owners.
The second experience is the way in which the child welfare system disproportionately takes children away from black mothers. Lastly, is the history of forced sterilization of black women, and the encouragement of government-funded sterilizations of black and poor women which continues today. While these primarily speak to experiences of black female Americans, they illustrate the devaluation of black mothers which began with eugenics and has continued to evolve and inform reproductive policy, as well as contraceptive delivery to women of color in the developing world. Eugenics projects by western scientists in developing countries, which later evolved into birth control programs, was an act of racism, sexism and classism in the way it was designed to limit the reproduction of poor women and women of color.

The timing of the development of the oral contraceptive pill further reveals its racist origins. Briggs states by following scientific and technological breakthroughs, the development of the contraceptive pill should have begun in the 1940’s, rather than 1956. Studies occurring as early as the 1920’s had proved that estrogen and progesterone could be used to prevent ovulation, and more than several studies that followed indicated methods to utilize this finding to create contraception for humans. However, the timing of the oral contraceptive tool coincided with the beginning of U.S. involvement in development in the postcolonial world. Therefore, Briggs argues that the framing of overpopulation as a global social problem was a “critical precondition” for the development of the contraceptive pill. This was the beginning of the viewpoint that birth control was a key tool in economic development. This belief permeated the mindset of the liberal nationalist Puerto Ricans, evidenced by the founding of the Population Association in Puerto Rico. This was a group of mainly feminists who operated on the conviction that modernization, technology and science were the keys to social progress. Therefore, the contraceptive pill was itself a “product” of postcolonial development.
The example of competing nationalisms in Puerto Rico – one in which maternity is glorified as an integral symbol of the nation, the other which views limited fertility as crucial to modernization – represents the way in which postcolonial struggles encompassed a battle over the regulation of black, brown, and working-class women’s bodies. It also illustrates the way in which the development and promotion of birth control by western countries in the developing world has a racist, classist history when considered in relation to its roots in the practice of eugenics. An intersectional feminist analysis seeks to understand the various forms of oppression that may be operating on a woman simultaneously. Racism and classism were at the center of contraceptive delivery because before it was “development,” it was eugenics.

The acknowledgement of this history is crucial given the American origin of the organizations examined in this paper. Nine out of twelve organizations whose discourse was analyzed in this paper originate from the United States. This example illustrates racial conflict present in the United States, but these tensions are not absent in other western and historically imperialist nations. These organizations cannot easily be separated from the racialized history of their country of origin and its connection to the evolution of birth control. The implication of an analysis informed by an intersectional feminist analysis complicates the problem framing of unmet need for contraception and subjects the solution framing of contraceptive delivery to be viewed as a contemporary mode of colonial engagement.

An Alternative Vision: Women as Active Subjects

The discourse used by the organizations examined in this paper unilaterally view contraception as a socioeconomic solution to aid development in the global south. Women’s bodies become victims to the mission of neoliberalization and women’s empowerment serves as a
disguise for labor participation. This discourse is problematic in the way it perpetuates the phenomenon Mohanty calls *originary power relations*. *Originary power relations* refer to binary power structures in which men are the “unilateral” and “undifferentiated” source of power. On the other hand, women occupy the roles of a cumulative reaction to this power; the binary between men as the powerful, and women as the powerless, is created.86 Mohanty points that western feminist literature recreates this binary wherein western women occupy the “subject” role of power, and the less educated, poorer, women of color “objects” in the global south are the powerless.87 In this way, colonial relations are recreated, and women in the developing world are constructed as powerless and their political agency is denied.

Contraception is undoubtedly a human right and a vital tool for women to embody full autonomy. However, the discourse around contraception must focus less on its function as a tool of socioeconomic development, and instead on its potential to transform women as subjects of resistance. Claudia Leeb explicitly recognizes the connection between power and discourse and the ways in which discourse “constitutes subjectivity itself.”88 Her work focuses specifically on the way in which discursive subject formation occurs in capitalist modes of production, and I examine the case of contraceptive delivery in light of her analysis.

Foucauldian-inspired feminist discourse examines the construction of the subject as created through subjection to discourse.89 However, as Leeb points out, this is problematic as it denies the feminine subject the opportunity to engage in constructing themselves as an active subject. An active subject is “capable of expanding its own powers and projects, an alternative to the reactive subject which defines itself merely in terms of what it opposes.”90 In the case of contraception delivery, women in the developing world are constructed in “object” status through the binary of originary power structures which situate western women as modern, liberated, and in possession
of full reproductive control. However, the discourse around contraception can be viewed alternatively through Leeb’s concept of the moment of the limit, which refers to the way the discourse cannot completely define a subject, as the real and non-identity create a gap in the discourse. Identity thinking, as defined by Adorno, lumps the needs and desires of white, middle to upper class, western, heterosexual women to make up the whole subject of “woman.” However, this renders women who do not fall into this category as non-identical. In our case, the non-identical consists of women in the developing world whose reproductive needs and desires are defined with the falsely whole “woman” identity.

Thus, it is evident that discourse does not have total control over constructing subjects. Therefore, existing within the discourse around contraception is the opportunity to understand women in the developing world’s use of contraception instead as an expression of active resistance in a patriarchal society. Leeb writes of this potential, “In this moment of the limit, the possibility of a political subject with the capacity of transformative agency emerges.” Thus, contraception becomes an area in which non-identity women – women of the developing world – transform themselves as active subjects.

It is important here to avoid the trap described by Mohanty in constructing a “Third World Woman,” meaning, assuming all women in the developing world are using contraception for the same reasons and desires. This can be evaded through understanding the woman of the developing world as what Leeb calls a political subject-in-outline. By this she means the subject “moves within a tension of a certain coherence (the subject) necessary for agency, and permanent openness (the outline) necessary to remain inclusive.” This conceptualization of the active subject accomplishes two important tasks. Firstly, it does not presume a subject based on identity to compose of an exclusive collective body. At the same time, it retains the subject’s agency.
Therefore, women in the developing world are constructed as active subjects who are non-exclusive yet have the capacity to transform the world around them.

Women using contraception in the developing world are not necessarily conscious of the task of becoming active subjects. This process is nonetheless essential for understanding the opportunity for resistance to the dominant forces of patriarchy and capitalism which contraception presents. If the discourse regarding contraception fails to acknowledge women as active subjects, women’s empowerment through contraception is rendered hollow, as its discourse further binds them to the capitalist, neoliberal structures which led to their oppression. Alternatively, framing the use of contraception as an act of resistance transforms the woman into an active subject with the ability to resist objectification.

Organizations which work in contraceptive delivery can allow room for women to be active subjects in their choice to use contraception by giving the developing world women they are working with full subject status in their conceptualization, strategy and interaction. In practical terms, this would look like the abandonment of their paternal, we-know-best nature which presumes contraceptive use to be the proven solution to developing world women’ liberation. Instead, the woman herself - her needs, desires, concerns and aspirations - would be centered in their discourse. This includes a recognition of variation among individual women, and that the homogeneous group of ‘women in developing world’ organizations refer to in their discourse does not exist.\textsuperscript{97} Thus, women as active subjects are “centered around a notion of being,” that is not constructed as part of or in opposition to a universal entity.\textsuperscript{98} Importantly, for this to happen, the market-centered discourse that characterizes fertility choices as rationally-made calculations must cease to “ignore the intricate webs in which women live their lives.”\textsuperscript{99} Instead, it must acknowledge women’s full lives with relational experiences in the communities in which they live, along with
their own desires for reproductive control which cannot be reduced to market-based considerations.

This also means that contraceptive delivery would need to be designed in light of an intersectional feminist analysis described aptly by Johnson-Odin when she writes, “It is not that black or other Third World feminists take a position against contraception, but that they seek to frame the discussion in a context which incorporates the impact of race and class on reproductive issues.” At first glance, contraceptive delivery in the developing world is seemingly logical in appealing to feminist notions of equality and autonomy. However, in order to enact fundamentally feminist policy, it is essential organizations offering contraceptive delivery design their programs in consideration of these contributions derived from various feminist analyses. When this occurs, organizations can provide contraception in a truly feminist manner that is empowering for the women they serve.

**Contraception as the “Master’s Tool” and Concluding Thoughts**

The promotion and delivery of contraception in the developing world must be considered in a feminist context with a lens sensitive to issues of gender, race and class. As we have discovered, the problem framing of unmet need of contraception is focused on the ways in which unintended pregnancies prevent full participation of women in the paid labor market. In this way, overpopulation provided a socioeconomic explanation for poverty in the developing world. The issue with this framing is dual-fold. Firstly, it abdicated both global capitalism and colonialism of any responsibility in producing structures which have impoverished the global south. Secondly, it perpetuates these structures through the encouragement of female participation in the capitalist system which oppresses them.
Another significant issue with the problem framing of contraception is the way in which it relies upon the essentialization of non-western culture which is rendered as “backward.” This discourse reproduces *originary power relations* which construct women in the developing world as powerless. Painting women as agentless reinforces their *object* status in a patriarchal society, denying their choice to use contraception as an expression of self-determined active subjectivity.

The solution framing of contraception in the discourse of these organizations is problematic in its construction of contraception as a tool of socioeconomic development. Contraception is painted as a tool for women to enter the labor market, which effectively places women’s value solely in the field of capital production and accumulation, rather than as a whole person. This narrative of contraception as a development tool fails to leave room for the construction of a woman as an active subject, whose choice to use contraception can alternatively be viewed as an act of resistance in a patriarchal society, in which fertility is traditionally controlled by men.

The discourse around the delivery of contraception in the developing world is predicated upon difference. This is the difference between the powerful and the powerless, from the western and the third world woman, from the one who possesses little blue pills to control her fertility and one who cannot. It is the difference between the developed world and the underdeveloped, between the secular and the religious, the liberated and the oppressed. This difference is necessary for these western organizations to establish themselves as the producers of discourse and knowledge and therefore as those who wield power. This difference is their power source.

Concurrently, the discourse relies upon assumptions of homogeneity. It expresses the belief that developing world women collectively possess the same issues, motives, needs and preferences, as do the communities they inhabit. They are assumed to share the universal desire of what western feminists already hold: liberation through reproductive control. What is not the same,
is rendered inferior. Audre Lorde speaks from the perspective of a woman who has been “forged in the crucible of difference” among other women of color, poor women, and women of the developing world. Lorde warns those women that “the master’s tools will never dismantle the master’s house.” Contraception as a tool of socioeconomic development is a master’s tool which sustains the master’s house of hegemonic colonial, white, western, capitalist patriarchy. As such, western organizations which were born in this house will never offer women in the developing world the tools to achieve their full active subject status in reproductive choice.
Notes:


14 Ibid, pg. 335


19 Ibid, p. 65.


Bill & Melinda Gates Foundation, “Family Planning.”


Population Council, “Strategic Priorities.”

World Health Organization, “Family Planning/Contraception.”


Cooper, “Productive, Relational and Everywhere?,” p. 441.


Ibid, p. 442.


Ibid, p. 89.

Ibid, p. 52.


Bill & Melinda Gates Foundation, “Family Planning.”


Marie Stopes International, “Contraception and Reproductive Health Services.”

PSI, “Contraception: PSI.”


Population Council, “Strategic Priorities.”

Cooper, “Productive, Relational and Everywhere?,” p. 447.


Al Riyami, Asya, Mustafa Afifi, and Ruth M Mabry, “Women’s Autonomy, Education and Employment in Oman.”


PSI, “Contraception: PSI.”

USAID, “Family Planning and Reproductive Health Program Overview”

World Health Organization, “Family Planning/Contraception.”


Ibid, p. 806.

Ibid, p. 806.


274 of 275
67 Ibid, p. 90.
68 Ibid, p. 93.
69 Ibid, p. 93-94
73 Briggs, “Reproducing Empire,” p. 82.
74 Ibid, p. 83.
75 Ibid, p. 103.
78 Crenshaw, “Demarginalizing the Intersection of Race and Sex,” p. 140.
79 Ibid, p. 139.
80 Briggs, “Reproducing Empire,” p. 82.
83 Briggs, “Reproducing Empire,” p. 129.
84 Ibid, p. 132.
85 Ibid, p. 130.
86 Mohanty, “Under Western Eyes,” p. 79.
87 Ibid, p. 79.
88 Leeb, “Power and Feminist Agency in Capitalism,” p. 3.
91 The “real” is a concept described by Jacques Lacan which refers to a “fault, a hole,” in the symbolic order and its signifier (Leeb 35). The “symbolic subject” is created through an identification performed by a signifier. The signifier exists in the domain of the “big Other,” which is the domain of language (Leeb 24). This meaning is established in two ways: the first is being linked to a chain of signifiers (a signifying chain), and the second is the symbolic value attached to each signifier through being opposed to another signifier. For example, “woman” is created in meaning through its opposite “man” (Leeb 27). However, the “real,” according to Lacan, resists symbolization. Due to this hole, the Lacanian symbolic order is not wholly dominating and does not hold exclusive power in determining people’s subjectivities.
92 Adorno explains that subjects cannot understand the object in its totality, and non-identical aspects are lost in acceptance of the whole concept. The non-identity is the “blind spot” in identity thinking; this parallels Lacan’s description of the “real” in the signifier. Therefore, the subject can never fully know a concept, meaning power structures are not all-powerful in object definition (Leeb 38).
93 According to Adorno, the thinking subject understands an object through placing it under a conceptual category; he refers to this as “identity thinking.”
95 Ibid, p. 22.
96 Ibid, p. 57.
100 Johnson-Odin, “Common Themes, Different Contexts,” p. 323.