SIGΣA

Journal of Political and International Studies

VOLUME XXXV

BYU

BRIGHAM YOUNG
UNIVERSITY
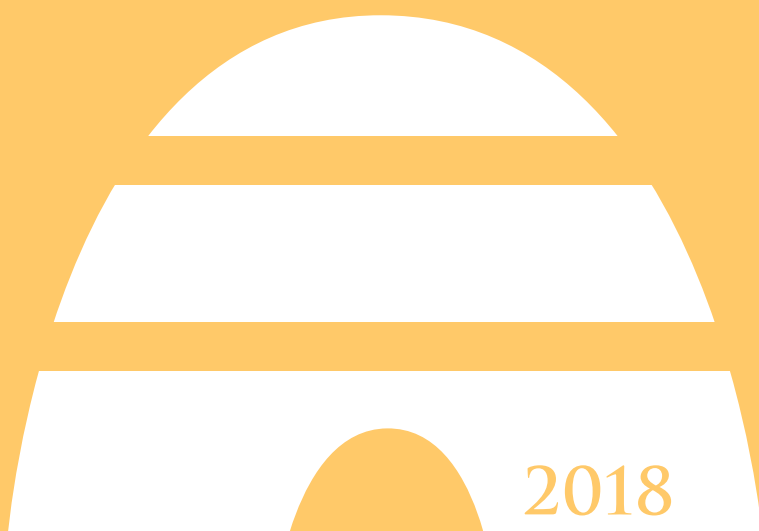
2018

SIGMA

JOURNAL OF POLITICAL AND INTERNATIONAL STUDIES

VOLUME XXXV · 2018

CONTENTS

A Letter From the Editor

Welcome to the 2018 edition of *Sigma.* This journal is the product of hours of intensive research, analysis, writing, reviewing, and rewriting by a dedicated team of student authors, faculty advisors, and student editors. It represents some of the finest academic research of Brigham Young University's undergraduate scholars in political science and international and area studies. As such, we invite you, our inquisitive reader, to closely examine and carefully consider each paper.

At a time when civil discourse is fracturing and growing increasingly *un*civil, *Sigma* stands as a beacon of hope and a clear message to all who read it: Young adults are interested and invested in answering the most pressing questions of our day, and they are equipped to do so with the tools of meticulous analytical research. In this volume, you will find new perspectives on timely issues, including the various—and often subtle—influences of gender in society, the repercussions of recent nationalist movements such as Brexit, and the nuances of engaging with those who have different views or experiences from our own. It is our hope that, as you peruse, you will ponder these findings and uncover additional insights and implications.

I express immense gratitude to the diligent authors, editorial staff, and professors who have generously given their time and expertise to bring these perspectives to light. I also thank you, our reader, for granting your consideration and intellectual engagement to this new body of knowledge. I hope your study of *Sigma* yields fresh understandings that will meaningfully contribute to the global conversation.

Sincerely,

Samantha Escamilla
Editor-in-Chief

Popular: The Monopoly of Force and Iraq's Popular Mobilization Units

Travis Birch

Introduction/Research Question

A monopoly on violence is often considered one of the key characteristics of a state, if not the defining feature. The existence of government-supported militias worldwide contradicts this view, leaving the question of why a state would choose to outsource force to external groups. In Iraq, for example, the mostly-Shiite militias that make up *al-Hashd al-Sha'abi*, translated as the Popular Mobilization, operate with the support of the state, which, until their legal incorporation into the Iraqi military structure in 2016, contradicted a constitutional prohibition against "the formation of military militia outside the framework of the armed forces" (Washington Post 2015). Despite their current status as part of the state security apparatus, questions about the allegiances of the militias, many of whom have expressed loyalty to Iran's Ayatollah Ali Khamenei, show the risks to state sovereignty that come with outsourcing the use of force.

The exact nature of the Hashd's relationship to the Iraqi state has changed frequently during the lifetime of the component militias and of the Popular Mobilization Units (PMUs) as an organization. Currently, the PMUs are legally affiliated with the state as a distinct branch of the military under the Ministry of Defense. According to a 2016 law, they are under the control of the prime minister, and as such, the designation of the Hashd as a nongovernment entity may seem incorrect (Iraqi Parliament 2016). There are several reasons why the militias make for a good case study in this context.

First, PMUs only recently reached their current level of integration into the state after years of independent operations—the Jaysh al-Mahdi, for example, fought against the Iraqi state and U.S. forces in the early 2000s before its tenuous membership in the PMUs. Second, there is reason to believe the Hashd's pledges of loyalty to the state may not be entirely sincere. Although some Iraqi political and religious figures have pushed

for the PMUs to be dissolved and their members incorporated into existing military units, the militias have preserved their leadership and organizational structures (Mansour and Jaba 2017). As a result, the component militias of the Hashd have maintained independence to pursue their political objectives, some of which appear to be more in line with Tehran than Baghdad (Ibid.). Third, although the 2016 law forbids Hashd leaders from running for political office, several are expected to seek the position of prime minister in next year's elections by transitioning out of militia leadership and into linked political parties. These issues highlight the risks pro-government militias pose to the state: They may wield their violent power against the state, they may pursue objectives that are in opposition to state policy, and their power may be leveraged to wrest political influence from state leaders. For these reasons, the Iraqi government's relationship with the Hashd is helpful in understanding governments' motivations for supporting paramilitaries.

Given these problems, why did the Iraqi state choose to support the PMUs, when doing so meant weakening its control over a critical state function? Baghdad has chosen to support the militias because of the weakness of the state military in the face of the Islamic state, sectarianism, and the influence of Iran.

Literature Review

According to Max Weber (1965), a monopoly on the legitimate use of force is the defining characteristic of a state, which uses violence to maintain its dominance over sub-state interests. The existence of pro-government militias, which Carey and Mitchell (2017) define as "armed groups that have a link to government but exist outside the regular security apparatus, and have some level of organization," defies the notion of the monopoly on force and poses a difficult question for scholars. Many have argued that militias are a cause, consequence, or indicator of failed states, but the presence of some type of pro-government militia in approximately one-third of the world's states defies easy conclusions (Bates 2008; Klare 2004).

A growing body of scholarship refutes the notion that paramilitaries are correlated with state weakness, arguing that states may authorize groups that are not part of the traditional military structure. Carey, Colaresi, and Mitchell (2016) use the "logic of delegation" to understand the relationship: Just as a state might delegate its authority to private contractors to perform administrative tasks, it may delegate the use of force to militias. Phelps (2014) agrees that the state can bestow legitimacy upon non-state forces, such as security contractors.

Still, there are obvious risks to the state from allowing vigilante violence, including the possibility that militias could become competitors to the state and even replace it. The attributes of pro-government militias and the nature of their relationship to the state varies widely. At one extreme are groups like Hezbollah, which acts as a quasi-state in Lebanon even as it participates in Lebanese political institutions. At the other end of the continuum are local defense militias, which are organized ad hoc for the protection of communities. Powerful militias threaten the state and pose a danger to the citizenry. Koren (2014) claims that the existence of such groups is predictive of mass killings,

while Ahram (2014) posits that nongovernment militias can be either an incentive or a deterrent to genocides or other atrocities. Carey and Mitchell (2017) have developed a typology of pro-government militias based on their relationship to the state and their relationship to society, adding that certain types of militias are more or less dangerous to the state or the populace. The nature of the Hashd, which includes groups with differing aims and identities, makes it difficult to classify, but most units would be classified as ethnoreligious militias, which pose a high risk to the state and a very high risk to civilians.

Given the severity of those risks and the high incidence of pro-government militias, we can infer the existence of strong benefits that militias offer states. Perhaps the most obvious benefit is that loyal paramilitaries outside of the military structure can protect leaders from coups (Carey, Colaresi, and Mitchell 2016). Another benefit is pro-government militias may give states the option to use violence in otherwise unacceptable ways while maintaining plausible deniability (Carey, Colaresi, and Mitchell 2015). Carey and Mitchell (2016) add that pro-government militias are cheap force multipliers, they provide local knowledge, and they grant the state legitimacy among the group from which the militia is drawn.

Method and Data

In this paper, I will undertake a qualitative case study using Iraq's Popular Mobilization Units to demonstrate benefits that motivate a state to permit and support the operations of pro-government militias. I derived possible reasons from the scholarly literature and from expert analysis, then used primary documents, including Iraqi laws and statements from influential figures, and secondary sources, including journalistic reports and expert analysis, to evaluate them.

Historical Background

The stage for a Shiite paramilitary mobilization was set by the invasion and occupation of Iraq by a U.S.-led coalition in 2003. The coalition's overthrow of Saddam Hussein ushered in a period of violent chaos, compounded by the actions of the Coalition Provisional Authority (CPA), the group charged with building a democratic Iraqi state. After the expulsion of Baath party members from the government, its second order was the complete dissolution of the Iraqi security and intelligence apparatus, dominated by Sunnis under the Baathist regime.

Ambitious Shiite figures like Muqtada al-Sadr were poised to take advantage of the resulting security vacuum. Sadr had inherited a dormant Shiite movement developed by his father Muhammad Sadiq al-Sadr in the 1980s, and he moved quickly after the invasion to mobilize his supporters and launch the *Jaysh al-Mahdi* (JAM) to fight the American occupiers. The Sadrist movement mostly consisted of undereducated slum-dwellers, and the JAM, which was "nothing more than a lightly armed and disorganized rabble," quickly proved itself unable to wage an effective campaign (Rayburn 2014, p. 181). JAM leaders who were dissatisfied with the Army's performance and disaffected

with Sadr's "erratic . . . dictatorial and mercurial" leadership decided to break away in 2004 to form a number of small, elite units to lead the insurgency (Ibid., pp. 180–88).

The founders of these "Special Groups," as they became known, sought the help of Qassem Soleimani, head of the Quds Force, a unit of the of the Iranian Revolutionary Guard Corps, which has been described as a "cross between the CIA, Special Operations Command, and State Department." Tehran, which wanted to expel the Americans from Iraq and to develop influence in Baghdad, was happy to oblige, and the Quds Force began sending weaponry and military advisors to Iraq and training Special Groups' forces in Iran (Ibid., p. 185). Two Iraqi leaders emerged as the most influential of the Special Groups. Qais al-Khazali, an acolyte of Muhammad Sadiq al-Sadr until the latter's assassination in 1999, presided over the creation in 2006 of the *Asa'ib Ahl al-Haq* (AAH) militia, a group that considered itself allied with Iran but not under its direct control (Ibid., p. 187). Abu Mahdi al-Muhandis's *Kata'ib Hezbollah* (KH), on the other hand, was "controlled directly by the Quds Force and loyal to the Supreme Leader," Ayatollah Khamenei (Ibid.). Muhandis had a long history of Iran-sponsored paramilitary action, having participated in attacks against U.S. and French embassies in Kuwait and an assassination attempt against the emir of Kuwait. He was elected a member of parliament in 2005 before being ousted for his past militancy (Ibid., p. 200).

In addition to AAH, KH, and JAM, which continued to operate as the Promised Day Brigades after the breakoff of the Special Groups (Ibid., p. 197), the anarchic insurgency included a group known as the Badr Brigades. Badr, led by Hadi al-Amiri, was created in Iran in 1982 as the military wing of the Supreme Council for Islamic Revolution in Iraq (SCIRI), a Shiite political group opposed to Saddam Hussein (Lake 2015). While these four groups' primary foe was the U.S. military, they also became involved in sectarian warfare against Iraqi Sunnis. The bombing of the Askari mosque in Samarra in 2006 by al-Qaeda escalated the conflict into a full sectarian war, and Shiite militias participated in a campaign of ethnic cleansing that lead to the "balkanization of Baghdad." (Marr and al-Marashi 2014). They also fought amongst themselves: there was, for example, fighting between the JAM and Badr in 2007 and between the JAM and the AAH after the U.S. withdrawal in 2011 (Rayburn 2014, pp. 28–31, 205).

A decision in 2011 not to renew the agreement allowing U.S. forces to operate in Iraq, made by Prime Minister Nouri al-Maliki and U.S. President Barack Obama, moved the burden of national security from the U.S. military to the reconstructed Iraqi Security Forces (ISF) (Karon 2011). Unfortunately, the ISF proved unable to bear that burden when the Islamic State (IS) burst into Iraq from Syria and began capturing territory. In 2014, a force of 1,500 IS fighters approached Iraq's second-largest city, Mosul, where 30,000 ISF troops were garrisoned (Marr and al-Marashi 2014, pp. 289–90). Alarmingly, the ISF fled the city without a fight, leaving their uniforms and a large cache of weapons behind (Schmitt and Gordon 2014).

The rapid progress of the Islamic state united the Shiite militias, which viewed the Sunni extremist group's advance with alarm. Ayatollah Sistani, influential head of

the Iraqi Shiite hierarchy, issued a *fatwa*, or religious ruling, in 2014, saying: "Citizens who are able to bear arms and fight terrorists, defending their country and their people and their holy places, should volunteer and join the security forces to achieve this holy purpose" (BBC 2014). Instead of joining the ISF, though, Shiites responded by joining militias, both pre-existing (JAM and AAH) and new.[1] The Iraqi national government, led by Maliki, subsequently formed the Commission for the Popular Mobilization to administer the militias and provide funding for their operations. The PMUs began participating in the war against the Islamic state, fighting alongside the ISF and becoming de facto allies of the U.S. military they had formed to defeat.[2]

The PMUs' battlefield success was undermined, however, by accusations of sectarian abuses against Sunnis. According to Amnesty International,

> PMU militias have extrajudicially executed or otherwise unlawfully killed, tortured and abducted thousands of men and boys. Victims were picked up from their homes, workplaces, camps for internally displaced persons, checkpoints or other public places. Some were later found shot dead. Thousands more are still missing, weeks, months and years after they were abducted. Amnesty International has documented such violations by PMU militias in and around Baghdad, and in the Anbar, Salah al-Din, Diyala and Kirkuk governorates. (2017)

Nevertheless, support for the militias in the highest levels of the Iraqi government has continued, with Prime Minister Haider al-Abadi, who replaced Maliki in 2014, calling the Hashd "the hope of country and the region" (Chmaytelli 2017). In late 2016, parliament passed a law moving the PMUs under the purview of the Ministry of Defense, effectively making them a discrete branch of the military, though the component militias continued to operate according to their own wills (Iraqi Parliament 2016).

Military Weakness

The most obvious reason for the Iraqi government's support for the PMUs is the weakness of the state military. The government was compelled to support the militias because their own military was incapable of policing the use of violence, particularly by the Islamic state, and because a lack of confidence in the ISF made it unlikely that Baghdad could have regained the monopoly on force through a conventional military buildup.

First, the Iraqi state military was unable to police the use of violence, the result of a series of decisions stretching back to 2003. After disbanding the ISF, the CPA had promised to "create in the near future a New Iraqi Corps, as the first step in forming a national self-defense capability for a free Iraq." The U.S. proceeded to spend billions of dollars recruiting, outfitting, and training a new military to shore up the authority of the state it was attempting to form. But, according to the International Crisis Group (2010), the rapidity with which Washington and Baghdad were forced to rebuild the ISF, as they rushed to fill the security vacuum created by the CPA decision, left a "legacy of expediency." Troops were recruited "without sufficient

regard to background or qualifications," and units were created without due regard for institutional structure, which resulted in political partisanship and fragmented ethnic loyalties.

The fall of Mosul to the Islamic state in 2014, three years after the U.S. withdrawal, proved the impotence of the ISF, but the Iraqi state's inability to govern the use of force was already evident before the extremist group's rise to power. For example, several of the militias now integrated into the PMUs were active years before the advent of the Islamic state. In 2006, in response to Baghdad's inability to "counter the insurgent groups that were . . . effectively challenging the hegemony of the state," the U.S. worked to form a Sunni anti-insurgency movement known as the *Sahwat*, a collection of tribal units that could be considered pro-government militias (though their support was later shown to be exclusively from the U.S and not from Baghdad) (Stansfield 2016, pp. 198–202). Of course, the U.S. military, rather than the ISF, was functioning as the protector and enforcer of the Iraqi state until the expiration of the Status of Forces Agreement in 2011 mandated the withdrawal of U.S. forces, marking the end of the ISF's gestation period (Ibid., pp. 210–11). It faced the Islamic state, then, as a fully formed military, providing the best example of its ineffectiveness.

Second, the Iraqi military's failure to defeat ISIS was a blow to its prestige and created recruitment problems, eliminating the possibility of recapturing the monopoly on force through a buildup of regular forces. Sistani's 2014 fatwa explicitly called for Iraqis to "volunteer and join the security forces" (BBC 2014). Later confirmed to refer to established, legal military units rather than militias (Sistani 2014), the statement was understood to represent his endorsement of the Shia militia movement. Mansour and Jafar, citing an interview with a civil servant in the Iraqi Ministry of Interior, claim that this misunderstanding was orchestrated and enabled by public distrust of the ISF: "Using the fatwa's message, Maliki and his allies pursued a wide-ranging campaign to recruit volunteers through hundreds of centers and offices. His recruitment was predicated on a smear campaign against the very Iraqi army that they had created—the same army that the fatwa had supposedly demanded the volunteers join" (Mansour and Jabar 2017, p. 7). The rapid expansion of the Hashd, in contrast to the contraction of the ISF, which in 2014 went from seventeen to five–seven brigades (Galbraith 2015), implies Hashd recruits' preference for the militias over the ISF. While this probably has something to do with sectarianism, which will be discussed below, the fact that Maliki's strategy focused on disparaging the ISF hints that the military's embarrassing defeat had diminished its reputation and made such a strategy effective.

The ISF weakness and Iraq's inability to strengthen it made state support for the Hashd a foregone conclusion. Not only would they have been unable to prevent the mobilization's formation if they had wanted, the state's survival was at risk in the absence of a capable military, and Baghdad had to take whatever help it could get. As will be shown below, however, there were other factors at play that encouraged the government to support the PMUs.

Sectarianism

Sectarian tension between Shiites and Sunnis is another reason Baghdad has chosen to support the Hashd. The Iraqi government, which fears its Sunni constituents but faces externally imposed quotas mandating sectarian balance in the military, supports the majority-Shiite PMUs as a means of ensuring Shiite dominance and security. Also it is a political move meant to maintain the support of a Shiite base among which the Hashd is extremely popular.

The origins of ethno-sectarianism in Iraq can be traced to the formation of the post–WWI order in the Middle East and the creation of the Iraqi state by Britain that combined Kurdish, Sunni, and Shiite-majority areas and established a Sunni monarchy. Subsequently, sectarian competition largely was kept under control by authoritarian rulers until the U.S. overthrew the Sunni establishment in its efforts to establish a liberal democracy in Iraq. The CPA, determined to block any group from dominating Iraqi politics, oversaw the creation of a new Iraqi constitution that codified an ethno-sectarian power-sharing system. Article 9 of the constitution, for example, dictated that "the Iraqi Armed Forces and Security Services," which had been predominantly Sunni under Hussein, would "be composed of the components of the Iraqi people with due consideration given to its balance and its similarity without discrimination or exclusion" (Washington Post 2015).

Such a system was understandably unpopular with Shiite politicians, as it prevented the majority group from dominating the country's politics. In particular, Nouri al-Maliki, a Shiite who became prime minister in 2006, "found the state's large bureaucracy inefficient, given its mandate under a sectarian quota system (*muhasasa ta'iya*), which included members from all major Kurdish, Shia, and Sunni political parties. Having loyal Shia militias, rather than the shaky cross-ethnic makeup of the Iraqi army, seemed a much more reliable way to secure a tighter command and control structure" (Mansour and Jabar 2017, p. 6), leading Maliki to support the creation of the PMU Commission in 2014.

Memories of Baathist brutality fed Shiite mistrust of Sunnis, which intensified with the rise in sectarian violence in 2006 and the emergence of the Sunni Islamic state as a dominant force in 2014. The perceived association between Sunnis and the Baath party, which continued long after the party was disbanded in 2003, and the general resentment of Sunnis by some Shiite politicians is evident in Maliki's rejection of the Sahwat brigades after the U.S. withdrawal in 2010. In response to escalating violence in 2006, coalition leaders had worked diligently to secure the cooperation of Sunni Arab tribesmen in western Iraq against insurgent groups like al-Qaeda, and by 2010, those efforts had paid off. In 2008, however, the Maliki government demanded that the U.S. turn over control of the Sahwat militias with the promise that the fighters would be given positions in the ISF (Oppel Jr. 2008). Instead, Baghdad disbanded the brigades and used the membership records to "round up and target" Sunni leaders. Stansfield attributes the decision to turn against the militias to a mixture of anti-Sunni sentiment and fears of a Ba'athist revival:

There was a deep-rooted fear of the Ba'th and the return of elements of Saddam's regime. There was also, of course, an even deeper angst concerning the expansion of Sunni *jihadists* and their visceral hatred of the adherents of Shi'ism. But, in effect, these fears merged into one. It would be difficult to exaggerate the level of concern that was felt among circles of Shi'i politicians in government of the possibility of the Ba'th, which was not becoming conflated with 'Sunnis,' returning to power. Even among Kurds, the leaderships in Erbil and Suleimani saw the hands of the Ba'th at every political juncture and insurgent assault (Stansfield 2016, p. 202).

The disparity between Baghdad's reaction to the Sunni Sahwat, which was to wrest control and destroy it, and its reaction to the Hashd, which was to support and incorporate it, betrays the sectarian motivations of Maliki and other Shiite leaders. Fears of Ba'athism became an avatar for anti-Sunni sectarianism, and a similar phenomenon occurred with the expansion of the Islamic State, evident in stories of human rights abuses committed by Shiite militias against local Sunnis, whom they accused of supporting the Islamic State (Human Rights Watch 2017).

Sectarianism drives the state's support for the Hashd from below as well—Shiite politicians are compelled to support the PMUs because of the militias' popularity among their voter base. According to Mansour and Jabar, 99 percent of respondents to an August 2015 poll supported the use of the Hashd to fight the Islamic State. Up to 75 percent of men between eighteen and thirty years of age in Shia-majority areas had enlisted to fight in the PMUs by spring 2016, though many are inactive because of funding restrictions (Mansour and Jabar 2017, pp. 10–11). The popular interpretation of Sistani's 2014 ruling inspired feelings of religious duty among Iraqi Shiites and gave the militias an air of religious legitimacy. One Hashd fighter interviewed by PBS in 2016 remarked: "I came in response to the fatwa, the doctrine to defend my country, my sacred places. My wish in life is the end of Da'esh [a synonym for the Islamic State] in Iraq. . . . The fatwa opened the door for us, so we volunteered with the Hashd militia. Ours is a belief and a will. We came because of our belief, not for a salary or anything else. We came because of our belief and our principles" (PBS NewsHour 2016). According to Renaud Mansour:

> Inside Iraq, the popular perception of the PMU among Iraqis—and particularly the Shia—is far more favorable than the group's reputation among Western politicians, analysts, and human rights groups.... Many Iraqis, and particularly the Shia, are convinced that had it not been for the PMU, ISIS would have taken over more territory and seriously threatened Baghdad. . . . The PMU occupies a new space in the Iraqi Shia imagination. As one fighter told the author, "you can criticize any politician or even religious cleric, but you cannot speak against the PMU and its martyrs." . . . Iraqi society is now full of popular songs, commercials, and banners, which display the leaders and martyrs of the different military groups. (2017)

Shiite politicians are constrained by the PMUs' high levels of participation and popularity among their base, especially with next year's election looming over them.

Support for the PMUs among Shiite political leaders has been motivated by sectarianism, including a desire to circumvent U.S.-imposed power-sharing rules and a desire to please Shiite constituents who strongly support the militias.

Iranian Influence

Another reason for the state's support of the Popular Mobilization Units is the influence of Iran, which has close connections with key politicians and militia leaders. Tehran's regional interests, coupled with its influence over key figures in the state and its willingness to provide support for Hashd, encouraged Iraq to support the militias.

According to scholars and commentators, Iran's interests in Iraq are best understood in the context of a greater Middle Eastern "cold war" between Iran and its allies on one side and Saudi Arabia and its allies on the other (Gause III 2014). More than a sectarian conflict, it is a battle for geopolitical influence. Iran, for example, supports Hezbollah in Lebanon, Bashar al-Assad in Syria, and the Houthi rebels in Yemen. Iran's interests in the Levant have been characterized as the creation of a political, ideological, and physical "axis of resistance" to bolster its position against Saudi Arabia (and Israel), including the creation of two corridors from Iran to the Mediterranean (Mohseni and Kalout 2017). These land routes would be protected by Iranian proxy forces rather than Iranian troops. Of course, the defeat of the anti-Shiite Islamic state, whose border-crossing advance posed a threat to Iran, was also of grave importance to Tehran.

It is within this context that Iran has supported the growth of al-Hashd al-Sha'abi in Iraq. According to a December 2014 *Washington Post* article, "Iran [had] sent more than 1,000 military advisers to Iraq, as well as elite units, and [had] conducted airstrikes and spent more than $1 billion on military aid" in response to Islamic State gains that summer, a claim supported by Badr leader Hadi al-Amiri (Ryan and Morris 2014; Morris 2014). AAH, which is fighting in Syria as well as in Iraq, is said to receive $1.5–2 million per month from Tehran, though the PMUs generally have received a greater proportion of their funding from Baghdad, as they have developed a closer relationship with the state (Stanford University 2017a; Eisenstadt and Knights 2017). Iranian leadership in Iraq continues to be managed by Qassem Suleimani, who maintains close relationships with Hashd militia leaders (Chulov 2014; Chulov 2010). According to Mansour and Jabar, the PMUs play several roles in Iran's larger Iraqi and regional strategy. Some of the smaller pro-Iran militias are viewed by Tehran as border protection units, while larger ones are being groomed to enter or expand their role in Iraqi politics as friends of Iran.

> Their military resources—including heavy armor, drones, and military advisers—all come from Tehran. Their cash and political legitimacy come from Baghdad. These paramilitaries are either full-fledged political parties or in the process of establishing political representation in the lead-up to Iraq's planned 2018 provincial and parliamentary elections. (Mansour and Jabar 2017, p. 13)

The extent of Iranian influence in Middle Eastern conflicts may be generally overstated, but the case for Iranian power in Iraqi politics is strong, with abundant anecdotes

of Tehran's intrusion in Iraqi affairs (Kendall 2017). The circumstances of Nouri al-Maliki's appointment to a second term as prime minister offer one such case. In the 2011 election, Maliki's *Dawlat al-Qanun* coalition was defeated by the *Iraqiya* party, led by Ayad Allawi, after Maliki refused to join forces with the Iraqi National Alliance. The surprise upset was followed by a vote recount and nearly five months of negotiations that resulted in the Erbil Agreement, a compromise that gave Maliki the premiership in exchange for some concessions. The reasoning behind the decision to let Maliki remain in power is contested—Stansfield says that he had "managed to accrue significant power in the period running up to the election and . . . used them to maintain his premiership" (Stansfield 2016, pp. 207–9). Stansfield portrays Iran as merely acquiescent:

> The Iranian government, too, had little interest in encouraging Iraqi politicians to respect the results of the elections. Not only was Iraqiyya seen as a haven for Sunni voters; it was also seen as the new manifestation of the Ba'th. Furthermore, Allawi was not a man who Iran, even through the IRGC, could influence. Throughout his period heading a leading opposition movement to Saddam, the Iraqi National Accord (*Al-Wifaq al-Watani al-Iraq*), Allawi had been closer to Western governments and intelligence agencies. This, combined with this secularism and non-sectarian appeal, meant that the Iranian kingmakers would follow their U.S. counterparts and simply let Maliki bully his way back to the premiership. (Ibid., p. 209)

American official Ali Khedery, on the other hand, claims that Iran ordered the outcome. Speaking of debates among U.S. leadership over which candidate to support for the position of prime minister, he writes:

> Our debates mattered little, however, because the most powerful man in Iraq and the Middle East, Gen. Qassim Soleimani, the head of the Quds Force unit of Iran's Revolutionary Guard Corps, was about to resolve the crisis for us. Within days of Biden's visit to Baghdad, Soleimani summoned Iraq's leaders to Tehran. Beholden to him after decades of receiving Iran's cash and support, the Iraqis recognized that U.S. influence in Iraq was waning as Iranian influence was surging. The Americans will leave you one day, but we will always remain your neighbors, Soleimani said, according to a former Iraqi official briefed on the meeting. After admonishing the feuding Iraqis to work together, Soleimani dictated the outcome on behalf of Iran's supreme leader: Maliki would remain premier; Jalal Talabani, a legendary Kurdish guerilla with decades-long ties to Iran, would remain president; and, most important, the American military would be made to leave at the end of 2011. Those Iraqi leaders who cooperated, Soleimani said, would continue to benefit from Iran's political cover and cash payments, but those who defied the will of the Islamic Republic would suffer the most dire of consequences. (Khedery 2014)

Marr and al-Marashi (p. 274) agree that Iran influenced the outcome by compelling Sadr to support Maliki.

Other instances of Iranian interference abound—a 2017 New York Times article, for example, claims that Iraqi parliamentarian Hoshyar Zebari was ousted from his role as finance minister at Iran's insistence, after Tehran found him too friendly to the United States,[3] but Iran's backing of Maliki has particular salience for the PMUs (Arango 2017; Patoon 2017). Mansour and Jabar claim that Tehran's support for Maliki was conditioned on his support for the Hashd, citing a pro-Iranian MP as saying that "[Iran] did not have any faith in the Iraqi army and groomed armed groups in almost all border provinces as a strategic asset to protect its national security" Mansour and Jabar 2017, p. 28). They argue that noticeable change occurred in Maliki's stance on non-state militias before and after 2010. He had, after all, attacked the Shiite Jaysh al-Mahdi, sending, in the words of a Shiite MP, "the message to all the militias including the Awakening that their days are numbered," only to staunchly support the PMUs six years later (Oppel Jr. 2008). This policy shift is consistent with the theory that the prime minister owed his allegiance to Iran after they intervened on his behalf in 2010. Speaking more generally, it indicates that Baghdad's support for the Hashd was a result of Tehran's interference in Iraqi politics. Iraqi leaders feared that they would lose Iranian support if they did not support the militias and therefore lose their positions.

Counter Arguments

Several other theories for state support of the Hashd exist that I find less compelling than those I have listed above. They focus first on the PMUs as political leverage and second on the state's desire to order atrocities while maintaining plausible deniability.

Political Leverage

The political leverage theory attributes state support for the PMUs to the interests of individual politicians, who see the militias as an opportunity to increase their power and protect their position. Mansour and Jabar write:

> Iraq is in the midst of an internal Shia political struggle over control of the state between former prime minister Nouri al-Maliki, who wants to return to power; current Prime Minister Haider al-Abadi, who is trying to maintain the power of the state; and cleric Muqtada al-Sadr, who is bent on ensuring that the Maliki faction does not return to power. A crucial factor that will help determine who gains an advantage in this struggle will be whether the PMF paramilitaries are integrated into the state's existing security apparatus and used to reinforce the country's political status quo, or if instead these paramilitary groups are retained as a separate parallel and independent military force that can be used to reshape Iraq's current political and security landscape. (Mansour and Jabar 2017, p. 4)

Maliki was and may still be the politician most able to take advantage of the political weight of the Hashd. Known as the "godfather" of the militias for his role in forming the PMU Commission in 2014, his premiership was marked by authoritarian efforts to consolidate control over all aspects of Iraqi security (Ibid. 25; Dodge 2012, pp. 124–30). If the Hashd militias were loyal to any one Iraqi leader, it would be Maliki, and taking advan-

tage of the PMUs would not be out of character for him. Perhaps the aforementioned change in Maliki's policy preference toward support for the militias in 2010 was not simply a product of Iranian pressure but an attempt to shore up his political position after the electoral loss made clear his weakness.

However, there is little evidence that he tried to use the units for that purpose when he lost the premiership in 2014. He had fared well in the elections, but the Shiite political establishment was unwilling to support him for a third term. Instead, they selected Haider al-Abadi as PM, infuriating Maliki, who adamantly opposed the decision (Stansfield 2016, pp. 238–40, 244). If he had intended to use his influence over the Hashd or the ISF to hold on to power, this would have been the moment, and observers feared he would do just that. Instead, although Maliki did "[deploy] security forces to strategic points across the capital," he "told security forces not to intervene in the political crisis and to protect the nation. . . . He said that members of the security forces had pulled back from the front lines of Iraq's conflict zones upon hearing he would be replaced but that he had ordered them to return to the fight" (Morris and Sly 2014; Arango, Reuben, and Gordon 2014). Days later, Maliki peacefully relinquished the premiership.

It is possible that Maliki intended to turn the PMUs and ISF against the opposition or that he expected to intimidate Abadi's supporters. He did "issue veiled threats," saying that he "could not be held responsible for those who might wish to fight on his behalf," but the apparent lack of effort to leverage military force to hold his position weakens the argument that support for the Hashd was intended as a kind of "coup-proofing." Another variant of the theory says that Maliki "will try and use the legitimacy of the PMU to regain power, himself, or via his protégés" in upcoming elections (Mansour 2017a). However, the reported electoral hopes of PMU leaders indicate they have accumulated their own political influence, effectively sidelining Maliki. Maliki's actions in the upcoming year should shed more light on this theory.

Plausible Deniability for Abuses

As mentioned above, scholars have found a strong correlation between the incidence of pro-government militias and human-rights abuses, leading to claims that governments choose to support paramilitaries out of a desire to order atrocities while keeping distance between themselves and the act (Carey, Colaresi, and Mitchell 2015). The Hashd has been accused of numerous human rights abuses, particularly in Sunni areas liberated from the Islamic state, where they are said to have detained and tortured innocent civilians whom they believed to be IS supporters (Frontline 2017).

These abuses, though tragic, only have relevance to the research question if it can be shown that the government desired the abuses, that they were unwilling to execute them through conventional military action, and that they supported the militias as a means of accomplishing them without direct responsibility. Given the sectarian tensions mentioned above and the way that Sunnism, Ba'athism, and support for the Islamic state became conflated in the eyes of the Shiite majority, it is possible

that the state supported violence against Sunnis as Islamic state supporters, whether or not the relationship truly existed. If the state wanted to distance itself from the Hashd's crimes, however, it would not have taken steps to incorporate them into the state military apparatus, which they have repeatedly done, and indeed the Hashd has deflected allegations of abuses by playing up its connection with the state (Al-Hashed 2017). Furthermore, Abadi has acknowledged the PMUs misdeeds rather than ignore or disassociate from them (Amnesty International 2017, p. 12). Although the human rights abuses are significant as such, it is unlikely that they cause government support for the militias.

Conclusion

Although Iraq has made changes to its relationship with the Hashd, the independence, political aspirations, and divided loyalties of the militia leaders continue to pose risks to Baghdad, threatening to sacrifice state sovereignty, alienate Sunni and Kurdish minorities, and inject violence into the state's politics (Sattar 2016). Despite the risks associated with delegating the use of force to these non-state militias, Iraq has supported the Popular Mobilization Units because of the weakness of their own military, the looming influence of Iran, and sectarian affiliation. With Iraq's military successes against the Islamic state, the apparent *raison d'être* of the Hashd, and the next year's elections, in which several PMU leaders are expected to participate, the true strength of the militias may become more apparent (Gurbuz 2017).

NOTES

1. For a list of the most significant militias of the Hashd, see Aymen Jawad Al-Tamimi, "Hashd Brigade Numbers Index," *Pundicity* blog, October 31, 2017, accessed November 6, 2017, www.aymennjawad.org/2017/10/Hashd-brigade-numbers-index or "Security and Politics Mix Inseparably in Iraqi Politics," *Inside Iraqi Politics* 111, July 20, 2015, accessed December 14, 2017, www.insideiraqipolitics.com/Files/Inside%20Iraq%20No111.pdf.

2. It is worth mentioning that the Hashd also contains non-Shiite militias, and that the motivations and loyalties of the component militias vary. I have included information about some of the most influential groups here, as they pose the greatest threat to the state and are therefore most relevant to the research question, but I have purposefully excluded many for simplicity.

3. The same article hints that Iran influenced the selection of former-Quds Force commander and Badr militia fighter Qasim al-Araji as to head the Ministry of Interior, which oversees the Federal Police and, until the month before his appointment, the PMUs.

REFERENCES

Ahram, Ariel I. 2014. "The Role of State-Sponsored Militias in Genocide." *Terrorism and Political Violence* 26, no. 3 (Spring/Summer): 488-503.

Al-Hashed. 2017. "The Popular Mobilization Responds to the Report of the Organization Amnesty International [Arabic]." Last modified January 6. Accessed December 4, 2017. http://al-hashed.net/2017/01/06/ردالحشد-الشعبي-على-تقرير-منظمة-العفو/.

Al-Tamimi, Aymen Jawad. 2017. "Hashd Brigade Numbers Index." Pundicity: Informed Opinion and Review. Last modified October 31. Accessed November 6, 2017. www.aymennjawad.org/2017/10/Hashd-brigade-numbers-index.

Amnesty International. 2017. "Iraq: Turning a Blind Eye." Amnesty International. Accessed November 13, 2017. www.amnestyusa.org/files/iraq_report_turning_a_blind_eye.pdf.

Arango, Tim. 2017. "Iran Dominates in Iraq After U.S. 'Handed the Country Over.'" *New York Times*, July 15. Accessed November 11, 2017. www.nytimes.com/2017/07/15/world/middleeast/iran-iraq-iranian-power.html.

Arango, Tim, Alissa J. Reuben, and Michael R. Gordon. 2014. "Iraqis Nominate Maliki Successor, Causing Standoff." *New York Times*, August 11. Accessed November 13, 2017. www.nytimes.com/2014/08/12/world/middleeast/iraq.html.

Bates, Robert. 2008. *When Things Fell Apart: State Failure in Late-Century Africa*. Cambridge: Cambridge University Press.

BBC. 2014. "Iraq Conflict: Shia Cleric Sistani Issues Call to Arms." *BBC*, June 13. Accessed November 5, 2017. www.bbc.com/news/world-middle-east-27834462.

Bengali, Shashank. 2010. "WikiLeaks: Maliki Filled Iraqi Security Services with Shiites." *McClatchy DC*, December 3. Accessed December 3, 2017. www.mcclatchydc.com/news/special-reports/article24602926.html.

Carey, Sabine C. and Neil J. Mitchell. 2017. "The Monopoly of Violence and the Puzzling Survival of Pro-Government Militias." Working Paper. Accessed November 8, 2017. https://static1.squarespace.com/static/54d49799e4b0b97e05e90186/t/56a5e5831a5203313c39dcc8/1453712794730/Carey+and+Mitchell+ARPS+online.pdf.

———. 2016. "Pro-Government Militias and Conflict." *Oxford Research Encyclopedias: Politics* (October). Accessed November 7, 2017. http://politics.oxfordre.com/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-33.

Carey, Sabine C., Michael P. Colaresi, and Neil J. Mitchell. 2016. "Risk Mitigation, Regime Security, and Militias: Beyond Coup-proofing." *International Studies Quarterly* 60, no. 1 (March): 59–72.

———. 2015. "Government, Informal Links to Militias, and Accountability." *Journal of Conflict Resolution* 59, no. 5 (August): 850–76.

Chmaytelli, Maher. 2017. "Abadi Defends Role of Iranian-Backed Paramiltaries at Meeting with Tillerson." *Reuters*, October 23. Accessed December 15, 2017. www.reuters.com/article/us-mideast-crisis-iraq/abadi-defends-role-of-iranian-backed-paramiltaries-at-meeting-with-tillerson-idUSKBN1CS26X.

Chulov, Martin. 2014. "Qassem Suleimani: Can This Man Bring About the Downfall of ISIS?" *The Guardian.*, December 6. Accessed November 12, 2017. www.theguardian.com/theobserver/2014/dec/07/qassem-suleimani-middle-east-mastermind-islamic-state.

———. 2010. "Qassem Suleimani: The Iranian General 'Secretly Running' Iraq." *The Guardian*, July 28. Accessed November 12, 2017. www.theguardian.com/world/2011/jul/28/qassem-suleimani-iran-iraq-influence.

Dehghanpisheh, Babak. 2017. "The Iraqi Militia Helping Iran Carve a Road to Damascus." *Reuters*, September 22. Accessed November 12, 2017. www.reuters.com/investigates/special-report/mideast-crisis-iraq-nujaba/.

Dodge, Toby. 2012. *Iraq: From War to a New Authoritarianism*. New York: Routledge.

Eisenstadt, Michael and Michael Knights. 2017. "Mini-Hizballahs, Revolutionary Guard Knock-Offs, and the Future of Iran's Militant Proxies in Iraq." *War on the Rocks*, May 9. Accessed November 12, 2017. https://warontherocks.com/2017/05/mini-hizballahs-revolutionary-guard-knock-offs-and-the-future-of-irans-militant-proxies-in-iraq/.

Frontline. 2017. "Iraq Uncovered." *Frontline,* Season 34, Episode 16. March 21. Accessed November 1, 2017. www.pbs.org/wgbh/frontline/film/iraq-uncovered/.

Galbraith, Peter W. 2015. "Should We Give Up on the Iraqi Army?" *Daily Beast*, March 4. Accessed November 10, 2017. www.thedailybeast.com/should-we-give-up-on-the-iraqi-army.

Gause III, F. Gregory. 2014. "Beyond Sectarianism: The New Middle East Cold War." Brookings Doha Center Analysis Paper, Number 11 (July). Accessed November 12, 2017. www.brookings.edu/wp-content/uploads/2016/06/English-PDF-1.pdf.

Gurbuz, Mustafa. 2017. "The Popular Mobilization Forces and Iraq's Next Elections." *Arab Center Washington DC,* September 6. Accessed December 4, 2017. http://arabcenterdc.org/policy_analyses/the-popular-mobilization-forces-and-iraqs-next-elections/.

Human Rights Watch. 2017. "Iraq: Men Fleeing Mosul Held in Secret." Human Rights Watch. Last modified February 12. Accessed November 12, 2017. www.hrw.org/news/2017/02/02/iraq-men-fleeing-mosul-held-secret.

Karon, Tony. 2011. "Iraq's Government, Not Obama, Called Time on the U.S. Troop Presence." *Time*, October 21. Accessed December 4, 2017. http://world.time.com/2011/10/21/iraq-not-obama-called-time-on-the-u-s-troop-presence/.

Kendall, Elizabeth. 2017. "Iran's Fingerprints in Yemen: Real or Imagined?" *Atlantic Council* (October). Accessed November 10, 2017. http://pushback.atlanticcouncil.org/papers/irans-fingerprints-in-yemen-real-or-imagined/.

Khedery, Ali. 2014. "Why We Stuck with Maliki—and Lost Iraq." *Washington Post*, July 3. Accessed November 12, 2017. www.washingtonpost.com/opinions/why-we-stuck-with-maliki-and-lost-iraq/2014/07/03/0dd6a8a4-f7ec-11e3-a606-946fd632f9f1_story.html?utm_term=.658aa0dca51f.

Klare, Michael T. 2004. "The Deadly Connection: Paramilitary Bands, Small Arms Diffusion, and State Failure." In *When States Fail: Causes and Consequences*, edited by Robert I. Rotberg. Princeton: Princeton University Press.

Koren, Ore. 2017. "Means to an End: Pro-Government Militias as a Predictive Indicator of Strategic Mass Killing." *Conflict Management and Peace Science* 34, no. 4 (September): 461–484.

Lake, Eli. 2015. "Iran's Militias Are Taking Over Iraq's Army." *Bloomberg*, February 3. Accessed December 10, 2017. www.bloomberg.com/view/articles/2015-02-03/exclusive-iran-s-militias-are-taking-over-iraq-s-army.

International Crisis Group. 2010. "Loose Ends: Iraq's Security Forces between U.S. Drawdown and Withdrawal." International Crisis Group. Last modified October 26. Accessed December 2, 2017. www.crisisgroup.org/middle-east-north-africa/gulf-and-arabian-peninsula/iraq/loose-ends-iraq-s-security-forces-between-us-drawdown-and-withdrawal.

Iraqi Parliament. 2016. "Law of the Popular Mobilization Units [Arabic]." Iraqi Parliament Site. Last modified November 26. Accessed November 10, 2017. http://ar.parliament.iq/2016/11/26/قانون-هيئة-الحشد-الشعبي/.

Mansour, Renad. 2017a. "After Mosul, Will Iraq's Paramilitaries Set the State's Agenda?" The Century Foundation. Last modified January 27. Accessed December 4, 2017. https://tcf.org/content/report/mosul-will-iraqs-paramilitaries-set-states-agenda/.

———. 2017b. "Iraq is not Iran's Puppet." *New York Times*, November 2. Accessed November 12, 2017. www.nytimes.com/2017/11/02/opinion/iraq-iran-haider-al-abadi.html?emc=edit_tnt_20171105&nlid=71169855&tntemail0=y&_r=0.

Mansour, Renad and Faleh A. Jabar. 2017. "The Popular Mobilization Forces and Iraq's Future." *Carnegie Middle East Center* (April).

Marr, Phoebe, and Ibrahim al-Marashi. 2014. *The Modern History of Iraq.* 4th ed. Boulder, CO: Westview Press.

Mohseni, Payam and Hussein Kalout. 2017. "Iran's Axis of Resistance Rises." *Foreign Affairs*, January 24. Accessed November 12, 2017. www.foreignaffairs.com/articles/iran/2017-01-24/irans-axis-resistance-rises.

Morris, Loveday. 2014. "Appointment of Iraq's New Interior Minister Opens Door to Militia and Iranian Influence." *Washington Post*, October 18. Accessed November 12, 2017. www.washingtonpost.com/world/appointment-of-iraqs-new-interior-minister-opens-door-to-militia-and-iranian-influence/2014/10/18/f6f2a347-d38c-4743-902a-254a169ca274_story.html.

Morris, Loveday, and Liz Sly. 2014. "Iran Endorses Haider al-Abadi as Iraq's New Prime Minister, Spurning Nouri al-Maliki." *Washington Post,* August 12. Accessed November 13, 2017. www.washingtonpost.com/world/middle_east/iran-endorses-haider-al-abadi-as-iraqs-new-prime-minister-spurning-nouri-al-maliki/2014/08/12/0487f02c-222c-11e4-8593-da634b334390_story.html.

New York Times. 2014. "Arms Windfall for Insurgents as Iraq City Falls." *New York Times*, June 10. Accessed December 15, 2017. www.nytimes.com/2014/06/11/world/middleeast/mosul-iraq-militants-seize-us-weapons.html.

Oppel Jr., Richard A. 2008. "Iraq Takes Aim at U.S.-Tied Sunni Groups' Leaders." *New York Times*, August 21. Accessed November 12, 2017. www.nytimes.com/2008/08/22/world/middleeast/22sunni.html.

Patton, Callum. 2017. "Iraq's Powerful Interior Minister, Now Leading Fight Against ISIS, Was Held Twice by U.S. Forces After 2003." *Newsweek*, May 8. Accessed November 12, 2017. www.newsweek.com/trained-iran-and-arrested-smuggling-explosives-iraqs-interior-minister-was-us-596161.

PBS NewsHour. 2016. "The Controversial Force Joining the Fight Against ISIS." *PBS.* Last modified October 5. Accessed November 12, 2017. www.pbs.org/newshour/show/controversial-force-joining-fight-isis.

Phelps, Martha Lizabeth. 2014. "Doppelgangers of the State: Private Security and Transferable Legitimacy." *Politics and Policy* 42, no. 6 (December): 824–849.

Rayburn, Joel. 2014. *Iraq After America: Strongmen, Sectarians, and Resistance.* Stanford, CA: Hoover Institution Press.

Ryan, Missy and Loveday Morris. 2014. "The U.S. and Iran Are Aligned in Iraq Against the Islamic State – For Now." *Washington Post*, December 17. Accessed November 12, 2017. www.washingtonpost.com/world/national-security/the-us-and-iran-are-aligned-in-iraq-against-the-islamic-state--for-now/2014/12/27/353a748c-8d0d-11e4-a085-34e9b9f09a58_story.html.

Sattar, Omar. 2016. "How Iraq's PMU Law is Disrupting National Unity Efforts." *Al-Monitor*, December 14. Accessed December 15, 2017. www.al-monitor.com/pulse/originals/2016/12/national-settlement-pmu-law-iraq.html.

Schmitt, Eric and Michael R. Gordon. 2014. "The Iraqi Army Was Crumbling Long Before Its Collapse, U.S. Officials Say." *New York Times*, June 12. Accessed December 15, 2017. www.nytimes.com/2014/06/13/world/middleeast/american-intelligence-officials-said-iraqi-military-had-been-in-decline.html.

Sistani, Ayatollah. 2014. "Statement Regarding the Current Situation in Iraq [Arabic]." Ayatollah Sistani's Official Site. Last modified June 14. Accessed November 2, 2017. www.sistani.org/arabic/in-news/24912/.

Stanford University. 2017a. "Asa'ib Ahl al-Haq." Mapping Militant Organizations. Stanford University. Last modified March 24. Accessed November 11, 2017. http://web.stanford.edu/group/mappingmilitants/cgi-bin/groups/view/143.

Stansfield, Gareth. 2016. *Iraq: People, History, Politics*. 2nd ed. Malden, MA: Polity.

———. 2017b. "Mahdi Army." Mapping Militant Organizations. Stanford University. Last modified November 26. Accessed November 11, 2017. http://web.stanford.edu/group/mappingmilitants/cgi-bin/groups/view/57.

Uticensis Risk Services. 2015. "Security and Politics Mix Inseparably in Iraqi Politics." *Inside Iraqi Politics*, no. 111 (July). Accessed December 14, 2017. www.insideiraqipolitics.com/Files/Inside%20Iraq%20No111.pdf.

Washington Post. 2015. "Full Text of Iraqi Constitution." *Washington Post*, October 12. Accessed November 12, 2017. www.washingtonpost.com/wp-dyn/content/article/2005/10/12/AR2005101201450.html.

Weber, Max. 1965. *Politics as a Vocation*. Minneapolis, MN: Fortress Press.

Yaari, Ehud. 2017. "Iran's Ambitions in the Levant." *Foreign Affairs*, May 1. Accessed November 12, 2017. www.foreignaffairs.com/articles/iran/2017-05-01/irans-ambitions-levant.

Religious Targets: Where 2016 Presidential Candidates Used Religious Rhetoric

Rachel Day and Adam Johnson[1]

Introduction

The dynamics of elections have drastically changed over time. Elections in the past have involved everything from smoke-filled-room nominations to riots. As the dynamics of elections have changed, so has the strategy to win elections. Media now makes candidates' words widely available.

Virtually every word a candidate says in public or writes on social media will be heard or read by the people. A candidate's word choice matters now more than ever. As a result, candidates must be precise with the words they use. Political scientists have examined the strategy behind candidate language (Hart 2006; Chapp 2013). Others have written handbooks on how to effectively write speeches for them (Moffitt 1999; Trent 2000). Researchers observe that a key component of any speech is creating an atmosphere where audience members feel attached to the speaker. This attachment needs to build trust and create a lasting candidate likeability.

Therefore, this study examines how religious rhetoric is used by presidential candidates to connect with Christian audiences specifically. Because roughly 78 percent of Americans have faith in a religion, it is disappointing that religion and politics have not been studied together more (Pew Research Center 2016). While a wealth of research has been done on behalf of campaign talk and on the importance of using religious rhetoric, not much research currently connects the two (Hart 2006).

This paper seeks to combine the literature from campaign talk and religious rhetoric in order to examine current strategies that employ religious campaign talk in certain geographical areas. Rather than examining the *how* or the *what* of a speech, we are looking to see trends in *where* Christian religious rhetoric is more likely to be used.

Literature Review

One common concept in speech writing is that of *stock speeches* (also known as *module speeches*). Candidates employ these speech templates of their main message components while travelling around the country in order to maintain message purity. The candidates then tailor them for every audience along the campaign trail (Trent 2000). This forms a foundation for our theory regarding regional variance, which candidates apply in their speeches while maintaining message purity.

We are interested in assessing whether presidential candidates add general Christian rhetoric to their stock speeches in areas of higher Christian adherence rates. We examine this theory by using a list of keywords that imply Christian beliefs. The best way to describe this keyword idea comes from "Political Keywords: Using Language That Uses Us" by Hart (2006, p. 247):

> Such a scenario puts special demands on keywords. Because they are only *words* they must compete for attention in a world of sight and sound. . . . But because they are *keywords*, because they are (1) repeated with special frequency in (2) high- profile environments by (3) persons of considerable stature and, especially, because (4) they are easily overlooked (being only words after all), they can slowly, teasingly, affect our hearts and minds. Keywords are at their most powerful when they are taken for granted.

Choosing appropriate keywords is vital in speech design because they can resonate with an individual for months after the speech. Moffitt (1999, p. 158) describes this designing as a "deductive process," which consists of creating messages thematically and then choosing the ones that best fit the audience. From here, the speech authors then attract specific audiences by choosing keywords that will resonate within them. To test if candidates try to resonate with voters through specific language, we will analyze words common in Christian rhetoric, such as "God" and "Jesus." The importance of invoking spiritual feelings cannot be underestimated when studying religion and campaign strategy. Religious rhetoric, when used by a candidate, can positively influence an individual (Albertson 2010, p. 127). Albertson continues:

> [O]verall, this study demonstrates that a politicians' use of religious rhetoric led to attitude change at an implicit level. . . . It is particularly impressive that religious language shifted implicit attitudes for such well known figures, and that attitudes shifted even for people who would prefer less religious expression in politics.

Albertson's study provoked us to question whether presidential candidates implement religious vocabulary to appeal to a variety of regional audiences. Therefore, we conclude that politicians may effectively target specific audiences by using religious rhetoric in their speeches.

Although many candidates may invoke religious terms, McLaughlin and Wise state that "the effectiveness of religious cues depends on a citizen's level of religios-

ity" (2014). Therefore, we seek to improve upon existing literature to discover if candidates use religious rhetoric in areas of higher Christian adherence rates.

H1: As the religious adherence of a geographic region increases, the candidates will use more religious rhetoric when they address the people in the region.

The underlying assumption of this hypothesis is that some areas are perceived to be more Christian than others. Consequently, a campaign attempts to capitalize on this by including familiar Christian terminology in their speeches when in areas highly populated with Christians.

We decided to use general Christian rhetoric for our keywords instead of rhetoric specific to any particular religious tradition. According to Chapp, "The data indicate that, when candidates use religious language, voters respond in predictable and politically powerful ways" (2013). It is essential for a candidate speaking to a regional audience to connect with as many individuals as possible. Using general rather than specific Christian rhetoric is a better measure because people notice overt clues more than subtle ones (which are typically tradition-specific) (Karpowitz 2016). The use of language is not an appeal to specific religions in regional speeches but an appeal to all Christians. Chapp continues, "[L]anguage intended to excite religious predispositions in a nonsectarian manner affects candidate evaluations for those most committed to religion" (2013). Thus, for presidential candidates to influence the greatest number of religious voters, they are better off using general Christian rhetoric instead of tradition-specific rhetoric only.

Chapp also writes, "Candidates adopt different emotive styles to resonate with the identities of the audience members and to accommodate existing partisan patterns of voting behavior" (2013). Studies show that the right wing has a higher proportion of Christians than does the left wing (Maniam 2016), so we questioned if party could affect regional Christian targeting. Previous research concludes that the politically right will mobilize the religious base more so than the politically left (Monson and Oliphant 2007). Our study examines the truth in these findings, and based on this literature, we are led to the following hypothesis:

H2: As the religious adherence of a geographic region increases, the Republican candidates will use more religious rhetoric when they address audiences than will their Democratic counterparts.

Our null hypothesis states that no relationship exists between the uses of religious rhetoric in presidential campaign speeches in different regions of the U.S., even when accounting for party. It is possible that certain candidates are more likely to use religious rhetoric than others simply due to their own personal religious behavior and beliefs. However, we believe our analysis includes enough variables to come to a preliminary conclusion and open the door for future research on the topic.

Data

Much of the data comes from the specific language used in presidential campaign speeches. The speeches used were given by Democratic and Republican candidates during the 2016 presidential race. Included in this broad "speeches" category are press releases and statements made by the candidates during the election season. The database was created by hand for this project (Quinton 2017). The source web site for the speeches does not contain every speech given by the candidates.

This limited collection disproportionally underrepresents private events, which lack sufficient press coverage. However, because these events are intended for fundraising, the candidates' language may be different than it would otherwise be in more public venues. Since we are focusing on campaign language, the lack of private speeches does not affect our theory. However, some private event speeches may be in the database unknowingly.

Additionally, we dropped a few candidates from the analysis. These candidates include Democrats Martin O'Malley, Lincoln Chafee, and Jim Webb, and Republicans Carly Fiorina and George Pataki. Dropping them from the analysis appeared inconsequential, because we lacked confidence to include them due to their small number of speeches combined with their short presidential campaigns. Overall, the source web site does contain thousands of documented speeches, making the dataset reliable for this research. With the available data, we measure the dependent variable by the number of times a candidate uses a religious keyword in his or her remarks (making our dependent variable continuous).

To make this dependent variable possible to analyze, we first needed to operationalize terms. Initially, we searched for a list of Christian rhetorical terms created by other scholars but to no avail. We then searched web sites to find common words used among various dictionaries. These sources contained lists specific to common Christian rhetoric. The following list contains the most common religious keywords and constitutes the terms we used when analyzing the campaign texts: baptism, Bible, born again, Christian, Christianity, faith, God, gospel, heaven, hell, Jesus, pray, prayer, prophet, prophecy, redemption, religion, sacred, Satan, scripture, testify, worship, yoked.

The Census Bureau provided the state boundary shapefile at the 5m level (Geographic Products Branch 2012). We manually entered our statistical and locational data for mapmaking. The title of each speech provided the location data needed.

For purposes of basic analyses, we included non-geographic texts to determine candidates' religious rhetoric patterns in speeches intended for a more universal audience. To see the difference, we compare text based on geographic locations to those intended for a broader media audience.

To better analyze geographically based religious rhetoric, we included the religious adherence of the geographic region (state). We measure a state's religiosity based on the rate of religious adherence per 1,000 people, using data provided by the American Reli-

gion Data Archives (ARDA). All states' religiosity index were coded as the difference of the state's religious adherence and the national mean's religious adherence, going either positive or negative (for example, if the national average was 400 per 1,000, then a state with a religious adherence of 300 per 1,000 would have a score of -100 in our data). For the non-geographic texts (such as statements made from the campaign office that were intended for everyone and not just a specific region), we coded their religious adherence rate with a 0 (representing the national mean), assuming that the remarks were nationally intended. Although the measurement includes those of non-Christian religions, the U.S. is a predominantly Christian nation (Gallup reported that roughly 75 percent of Americans identify with Christianity in 2015 (Newport 2015)). Thus, with the unit of analysis at the state level, the percentage of those of other faiths is proportionally low and does not affect our analysis.

To measure the seriousness of a candidate's campaign, we created a few versions to measure viability. The measure indicates whether a candidate's ability to win the election influences his or her willingness to use Christian religious rhetoric. The apparent weakness here is that we cannot measure how viable a candidate thinks he or she is, but this is the best we can do with what is available. The Green Papers provided a popular vote percentage from the primaries, and we ranked each candidate, beginning at 1, by who had received the most votes.[2]

We included standard control variables to account for social differences that may influence the candidates' language. These candidate characteristics include age, education level (some college, bachelor's degree, and graduate degree, on a continuous scale), and the binary variables of gender (1 being male), political party (1 being Republican), and whether Christian or not (1 being Christian). We wanted to add a control for religious activeness, but we did not have information to gather this data, so we added the religious views of the candidate instead. Originally, it was unclear if the religious views of the candidates themselves should be included, but we decided to include this information due to the direction of Campbell, Green, and Layman, who state, "Here we focus on candidate religion, not because it is more important than other social characteristics to partisan voting, but because it is especially illustrative" (2011). This quote means that there is a possibility that a candidate's religious affiliation may affect his or her campaigning and thus must be included in the data.

Results

Table 1 shows the percentage of texts with at least one religious keyword for every candidate, as well as the percentage of religious keywords used in all (including nongeographic, which applies to both percentages) texts. If a candidate used at least one religious keyword in a speech (our list of religious keywords is found on the previous page), it was coded on a binary scale as a 1. We then used the software R to determine the percentage of texts with at least one religious keyword. We performed this analysis for each candidate. To analyze the percentage of religious keywords used in all texts, we ran a different analysis in R to extract the number of keywords used throughout all the speeches and divided

it by the total word count of all the speeches. Again, we repeated the process for each candidate. Those results are found in Table 1. (All tables and figures are in the Appendix.) The data show that Republicans use more religious rhetoric than Democrats overall. Out of all the texts from the Democratic Party, only 4 percent contained a religious keyword, compared to 6 percent for the Republican Party.

Of the Republican candidates, Ted Cruz came in first for both the percentage of texts containing at least one religious keyword and for the percentage of total religious rhetoric used. In other words, Ted Cruz used more religious rhetoric than any other candidate in the race. Donald Trump ranked last among the Republicans for percentage of texts with at least one religious keyword. Out of all his texts, only 3.3 percent contained a religious keyword. Rick Santorum came in second to last with 7.3 percent of his text containing at least one religious keyword (still a much higher percentage than Trump's). Yet Trump ranked seventh out of fourteen among the Republicans for the percentage of religious keywords used in total. From this preliminary analysis, it appears that Trump strategically used religious rhetoric, using religious rhetoric in highly concentrated instances. While the number of his speeches containing at least one keyword remained low, he used more religious keywords than half of his competitors.

On the other side of the aisle, Hillary Clinton had a much larger percentage of texts that used a religious keyword (4.5) compared to Bernie Sanders (2.5). Yet, it is interesting to note that Bernie Sanders and Hillary Clinton are basically identical with the percentage of total religious keywords spoken (Clinton at 0.21 percent and Sanders at 0.22 percent). Similar to Trump, Sanders might have strategically used religious rhetoric based on his audience. It appears that both candidates might have reserved the use of religious rhetoric for audiences with greater religiosity (see Table 1 Percentage of Religious Rhetoric during the Primaries).

Table 2 tests the usage of religious rhetoric by Donald Trump and Hillary Clinton during the general election. We hypothesized that Clinton would use more religious rhetoric during the general election than she did during the primaries to appeal more to soft-leaning Republicans or Independent voters with religious affiliation. We predicted the opposite from Trump, as the Republican nominee, because he would want to appeal to soft-leaning Democrats and Independent voters who tend to be more secularist. Surprisingly, the exact opposite occurred. Clinton used less religious rhetoric than she did during the primaries. The percent of her texts containing at least one religious keyword dropped from the primaries as well. Trump included more texts during the general election with at least one religious keyword than he did during the primaries, but the percentage of total religious keywords he used decreased from .22 percent during the primaries to .05 percent during the general election. From these results, we cannot draw any firm conclusions because there are many unknowns; nonetheless, this table seems to have some exciting numbers (see Table 2 Percent of Religious Rhetoric during the General Election).

Column 2 of Table 1 shows that Trump used relatively few religious words, yet in the general election he received 58 percent of the Protestant or other Christian vote and 52 percent of the Catholic vote (Kent 2016). During the primaries, "two-thirds of regular churchgoing Republicans were not supporting Trump for the GOP nomination even in April" (Smith 2016). This study does not determine the effects that religious rhetoric has on voting, but it is interesting to see the actual abundance of a candidate's religious rhetoric and compare that to the preferences of churchgoers. There are many reasons why churchgoers might be swayed to a candidate, but that is a topic for another paper.

After reviewing these numbers, we performed some qualitative analysis of the speeches and found examples of the religious rhetoric used. Even with this qualitative measure, it appears that candidates use religious rhetoric with more religious audiences. For example, when Bernie Sanders visited Salt Lake City, Utah, a state known for its LDS population, he gave a speech addressing the issues of immigration, in which he used a fair amount of religious rhetoric. In one remark he said, "We have seen too many wars, too much killing, too much suffering. And let us all together—people of good faith—do everything we can to finally, finally bring peace and stability to that region" (2016). In another speech, with a higher than normal religious audience at Liberty University, Sanders said, "I am not a theologian or an expert on the Bible or a Catholic, but I agree with Pope Francis when he says: "The current financial crisis . . . originated in a profound human crisis: the denial of the primacy of the human person! We have created new idols. The worship of the ancient golden calf has returned in a new and ruthless guise in the idolatry of money and the dictatorship of an impersonal economy lacking a truly human purpose" (2016). In two speeches of religiously strong audiences, Sanders uses religious rhetoric to appeal to the people—even though he is not a Christian himself. Most interestingly, he quoted the pope. It could be that he was attempting to establish credibility by aligning his views with those of the pope because Catholics view the pope as an authority figure. Just as Catholics follow the pope, research suggests that Latter-day Saints are more willing to change their views to follow what their prophet says (Campbell, Monson, and Green; 2014). If Catholics view the pope in the same manner that the average Latter-day Saint views the prophet of their church, then Sanders might be trying to demonstrate his agreeability with their religious leaders.

Clinton provided us with other interesting cases of attempting to connect with religious audiences. In a speech in North Carolina and two speeches in Ohio, she emphasized her passion to help children "live up to their God-given potential" (2016). In another speech given at the University of North Carolina–Greensboro, she talked about how her family and her Methodist faith have taught her the importance of helping children live up to their God-given potential (2016). We also found at least two speeches Clinton gave in Iowa where she used similar religious rhetoric. It appears she focused on the family (especially children), and relates love for family back to devotion to God. We do not know the exact effects this religious language has on voters, but, noting studies from Albertson

and from other scholars, we would not be surprised to see voters be more inclined to support a candidate who cares for children and God.

From the Republicans, we will analyze the two main front-runners: Ted Cruz and Donald Trump. Comparing them provides insight into two candidates who used religious rhetoric to very different degrees.

Cruz opened his presidential bid with a speech given at Liberty University. In that speech, he used the word "God" nine times, the word "Jesus" three times, and the word "faith" three times. In the speech, Cruz used religion to talk about his upbringing and how religion shaped his life. He commonly used such phrases as "our God-given rights" and "God bless" in many other speeches as well. From this brief qualitative analysis, it seems he might be using religion to establish his background and to convey to others why he holds the ideological beliefs he does.

Trump used religious rhetoric in a totally different manner than Ted Cruz. When Trump declared his candidacy, he only used one religious keyword. He said, "I will be the greatest jobs president that God ever created. I tell you that" (2015). Cruz used religious language in his opening speech to build his background and his viability, whereas Trump used religious rhetoric in his opening speech to say that God had created him superior to others. In one speech at a rally in North Carolina, Trump said the following: "Let me quote the same passage from the Bible I read on Saturday, from 1 John 4: 'No one has ever seen God; but, if we love one another, God lives in us and his love is made complete in us.' . . . Imagine what we our country could accomplish if we started working together as One People, under One God, saluting One Flag" (2016). In a separate speech in North Carolina, Trump talked about what we could accomplish if we started working as one people under God (2016).

After running a simple test to measure percentages and after analyzing speeches qualitatively, we decided to follow-up with quantitative measures. We ran three OLS models, one for all observations, one for just the primaries, and one for just the general election. We also ran a probit model based on a binary variable of whether or not a speech contained at least one religious keyword. Our OLS model came back with statistically significant results. We used the percentage of religious keywords as the dependent variable, with the key independent variable as the difference of religious adherence rates for the states from the national mean. We have multiple observations for various states, so we decided to run the regression by clustering the states together. We broke down the analysis even further by analyzing the percentage of religious rhetoric during both the primaries and the general election (Model 1), the primaries only (Model 2), and the general election only (Model 3).

We can see in Table 3 that there is statistical significance for the use of religious rhetoric on religious adherence. Although statistically significant, the substantive significance proves difficult to measure. For one, this paper is not concerned with the effect of religious rhetoric on the voters; it is merely concerned with testing the usage of religious rhetoric according to geography. Second, it should not be of any great sur-

prise that the coefficients are small; if anything, it should be a surprise that we found significance in all three models. Considering that Cruz used the most religious rhetoric out of all the candidates, and only 1.64 percent of his words were religious key words, we expected to receive small numbers from our regression. We are working with a dataset we assembled ourselves and with a dependent variable that is small in nature (a percentage of religious keywords the candidates used). Additionally, although Model 1 has an increase of just 0.000372 percent per one unit of religious adherence, the case of Utah is a multiplier of 306.

We also analyzed the effect in battleground states and found it to be statistically significant. We defined battleground states by using measures from FiveThirtyEight as competitive states that could have gone either way in the election (FiveThirtyEight 2016). In total, there were fourteen battleground states (see Appendix for list of states). These states were coded as a binary variable, with 1 indicating a battleground state.

A predictive measure for our OLS model produces nothing truly significant at the 95 percent confidence interval. It was close, but the confidence intervals overlap slightly.

Although, the means of the religious rhetoric do increase with the religious adherence rate. It is still a mystery how the confidence intervals for the religious adherence rates of -200 and -150 fall below zero (as explained on page 5, a state's religious adherence is the difference between a state's religious adherence rate per 1,000 and the national average religious adherence rate per 1,000). It is impossible for a candidate to use negative religious talk (see Table 3 Predicted Percent of Religious Rhetoric (OLS model), as well as Figure 1 Predictive Margins for Model 1 of Table 3).

We then decided to run a probit test, using a binary variable of whether or not a text had at least one religious keyword. If a speech contained any religious keyword, then the variable was coded as a 1, and if there were no religious keywords, the variable was coded as a 0. The numbers are still small, but again, they are statistically significant in indicating that states with a religious adherence rate above the national mean are more likely to use religious rhetoric than states with adherence rates below the mean. Table 4 show the probit model with a predictive margins plot (Probit Model).

The probit model is statistically significant, but it is hard to interpret the significance of a probit model. In order to better analyze the model, we ran a predictive margins command (see Figure 2 Predictive Margins for Table 4). This model shows a statistical significance between states with a negative religious rate and states with a positive religious rate from the baseline. There is not a strong statistical significance between states that have religious adherence above the national mean, but this model, just like the OLS model, supports our hypothesis that candidates will use more religious rhetoric in states with higher religious adherence rates.

While small, the statistically significant percentages are exciting in demonstrating the difference between the states with a negative adherence rate below the baseline and states with a positive adherence rate above the baseline. The fact that there is a statistically significant difference between the states with negative and positive

adherence rates supports our hypothesis that candidates will use more religious rhetoric in more religious states.

Conclusion

Our results suggest that candidates are predicted to use more Christian religious rhetoric as the religiosity of the state increases (see Figure 3 Statistical Significance in the Geographic Dispersion of Religious Rhetoric Use by Presidential Candidates for a visual representation of the results). We also found that party matters. Republican candidates are more likely to use religious rhetoric than Democratic candidates. The substantive significance is much harder to assess than the statistical significance. Further study on the effect of Christian religious rhetoric on voters, especially compared to tradition-specific targeting, would help to interpret the substantive significance.

The 2016 presidential election was not a common election. An outsider to Washington, Donald Trump won the election, contrary to many predictions. He also stood out by using less religious rhetoric than his competitors. This could imply that the stereotypically religious right no longer places as much value on the role of religion in politics. For this reason, further research should be performed on the effect of religious rhetoric on citizens' voting habits. Research in the past on the effect of religious rhetoric (Albertson 2010; Djupe and Calfano 2014) primarily deals with individuals' perceptions of public office-holders who use religious rhetoric. The research does not gauge the effect of religious rhetoric in an election. Thus, our research contributes by suggesting that candidates target religious individuals in predominantly religious areas. Does this strategy actually help the candidates win the presidency? Questions like this one could be answered by further researching and analyzing content of the speeches given by presidential candidates.

NOTES
 1. We would like to thank Professor Gubler, Devon Tenney, and James from the statistics lab for helping us with some of our questions about R. Their help was instrumental for our project.
 2. We created a few versions of the viability measure and determined that this one was the simplest, but it also provided similar results as the others. Due to multi-collinearity, we had to choose only one version. The simplest measure consisted of ranking all the candidates of each party by the end date of his or her campaign We assumed that the closer the campaign was to the general election, the more viable the candidate was. This consisted of ranking Donald Trump and Hillary Clinton as 1, then proceeding up in value until the first candidate who exited the primaries. A similar method of ranking records the date each candidate announced his or her candidacy in the primaries. A second variable takes the percentage of delegate votes each candidate received from their respective national conventions. Many candidates were not viable at this point, and consequently received zero percent. However, nearly every candidate was on the ballot (in at least some states) and received part of the popular vote. We included all of the previously mentioned candidates that we had dropped in these measures to include their effect in the race on other candidates.

APPENDIX

Table 1: Percentage of Religious Rhetoric During the Primaries

Candidates	Percentage of Texts with at least one Religous Keyword	Candidates	Percentage of Religious Keywords
Democratic Candidates	4.0	Democratic Candidates	0.49
Hillary Clinton	(1) 4.5	Hillary Clinton	(1) 0.21
Bernie Sanders	(2) 2.5	Bernie Sanders	(2) 0.20
Republican Candidates	6.0	Republican Candidates	1.71
Ted Cruz	(1) 33.5	Ted Cruz	(1) 1.64
Mike Huckabee	(2) 31.7	Marco Rubio	(2) 0.45
Ben Carson	(3) 31.0	Ben Carson	(3) 0.41
Lindsey Graham	(4) 28.1	Mike Huckabee	(4) 0.38
Rand Paul	(5) 18.5	Jeb Bush	(5) 0.35
Rick Perry	(6) 17.0	Lindsey Graham	(6) 0.24
Bobby Jindal	(7) 16.4	Donald Trump	(7) 0.22
Chris Christie	(8) 16.1	John Kasich	(8) 0.21
Jeb Bush	(9) 13.4	Bobby Jindal	(9) 0.14
Scott Walker	(10) 13.0	Rick Santorum	(10) 0.09
Marco Rubio	(11) 9.2	Rand Paul	(11) 0.05
John Kasich	(12) 7.63	Rick Perry	(12) 0.05
Rick Santorum	(13) 7.3	Scott Walker	(13) 0.05
Donald Trump	(14) 3.3	Chris Christie	(14) 0.04

*Candidates listed in descending order of religious rhetoric use. A definition of these percentages can be found on page 7.

Table 2: Percentage of Religious Rhetoric During the General Election

Candidate	Percentage of texts with at least one religious keyword	Percentage of religious keywords used in all texts	Number of Texts (n)
Hillary Clinton	3.3	0.17	1734
Donald Trump	6.5	0.05	764

Table 3: Predicted Percentage of Religious Rhetoric (OLS model)

Variables	Model 1 percentage	Model 2 Percentage	Model 3 Percentage
State Religious	3.72e-04*	7.84e-04**	2.71e-04*
Adherence	(1.96e-04)	(3.68e-04)	(1.52e-04)
Popular Vote	-3.52e-03*	-3.47e-03*	-0.0281**
Ranking	(2.01e-03)	(2.00e-03)	(0.0135)
Battle Ground	0.0388*	0.103**	-0.0117
State	(0.0226)	(0.0395)	(0.0293)
State	3.17e-03 (2.12e-03)	7.28e-03 (4.42e-03)	-8.51e-04 (1.19e-03)
Candidate Gender	0.0348*** (9.90e-03)	0.0492*** (0.0107)	---
Candidate Education Level	0.0340*** (4.66e-05)	0.0492*** 0.0339***	---
Popular Vote	1.81e-04	3.04e-04	-5.79e-03**
Ranking	(1.77e-04)	(1.88e-04)	(2.36e-03)
Party	0.0745*** (8.11e-03)	0.0786*** (5.39e-03)	0.0564* (0.0331)
Constant	-0.0995 (0.0590)	-0.193* (0.100)	0.502*** (0.155)
Observations	4,084	3,859	225
R-Squared	0.019	0.021	0.023

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Model 1 is for all observations. Model 2 is for the primaries. Model 3 is for the general election.

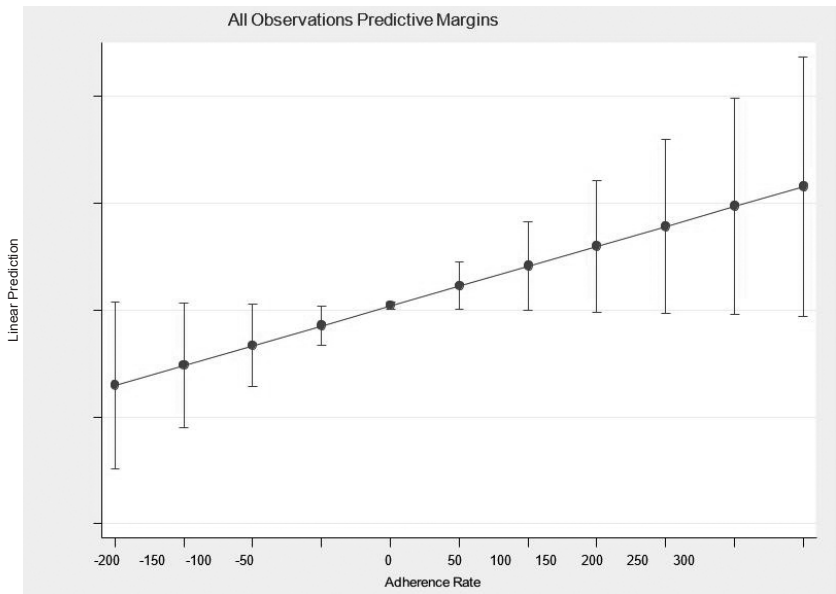**Figure 1: Predictive Margins for Model 1 of Table 3**

Figure 2: State Religious

Adherence	Margins	P-Value	Confidence Intervals	
-200	0.029	0.509	-0.048	0.108
-150	0.049	0.101	-0.01	0.106
-100	0.067	0.001	0.028	0.105
-50	0.085	0	0.067	0.104
0	0.104	0	0.1	0.107
50	0.123	0	0.101	0.144
100	0.141	0	0.099	0.183
150	0.16	0	0.098	0.221
200	0.178	0	0.097	0.26
250	0.197	0	0.096	0.298
300	0.215	0.001	0.094	0.337

Table 4: Probit Mode

VARIABLES	Texts with at least one religious keyword
State Religious Adherence	0.00545** (0.00217)
Population Rating	0.00930 (0.0114)
Battle Ground State	2.122*** (0.198)
State	-0.0193 (0.0139)
Candidate Gender	-0.323** (0.144)
Candidate Education Level	0.149*** (0.0419)
Popular Vote Rating	(0.00313 (0.00213)
Party	0.518*** (0.0953)
Constant	-1.162*** (0.352)
Observations	4,088

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Pseudo – R2 0.0880
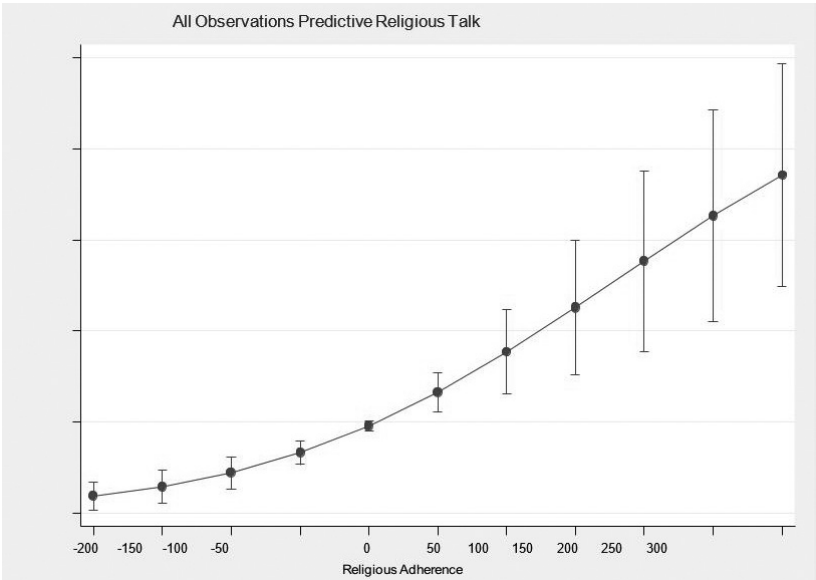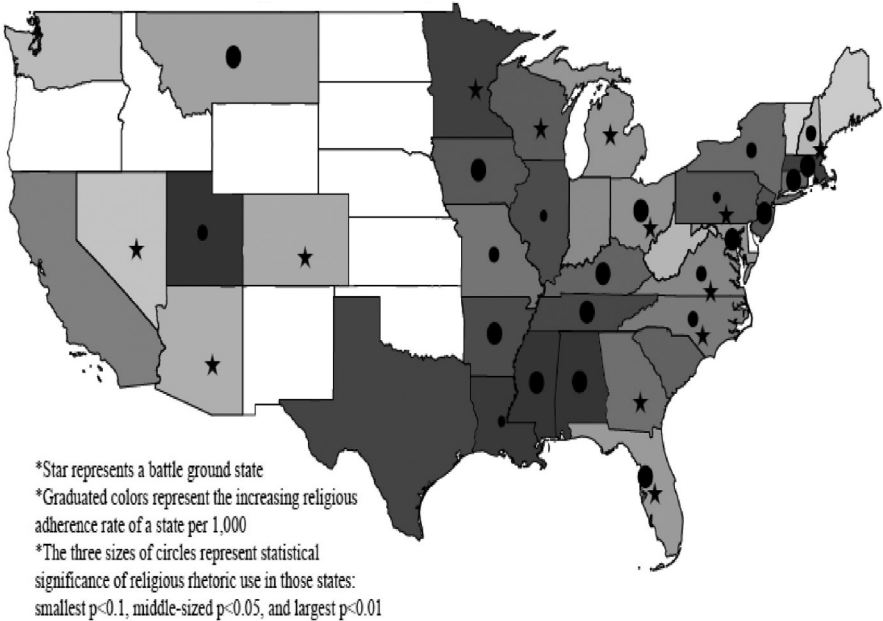
Figure 3: Predictive Margins for Table 4

All Observations Predictive Religious Talk

Table 5: *Percentage on the Y- axis Religious

Adherence	Margins P-Value		Confidence Intervals	
-200	0.034	0.021	0.006	0.07
-150	0.058	0.002	0.022	0.094
-100	0.089	0	0.053	0.124
-50	0.133	0	0.106	0.158
0	0.192	0	0.18	0.202
50	0.266	0	0.223	0.308
100	0.355	0	0.261	0.447
150	0.453	0	0.303	0.6
200	0.554	0	0.359	0.752
250	0.653	0	0.42	0.885
300	0.742	0	0.497	0.986

*Data of the margins plot in a table

Figure 4: Statistical Significance in the Geographic Dispersion of Religious Rhetoric Use by President Candidates

*Star represents a battle ground state
*Graduated colors represent the increasing religious adherence rate of a state per 1,000
*The three sizes of circles represent statistical significance of religious rhetoric use in those states: smallest p<0.1, middle-sized p<0.05, and largest p<0.01

This is the predictive margins plot for our second OLS model, which is analyzing texts only from the primaries. As one can see, the religious adherence rates for -200, -150, and -100 all have confidence intervals below zero. This does not make sense since it is impossible for a candidate to use negative religious rhetoric.
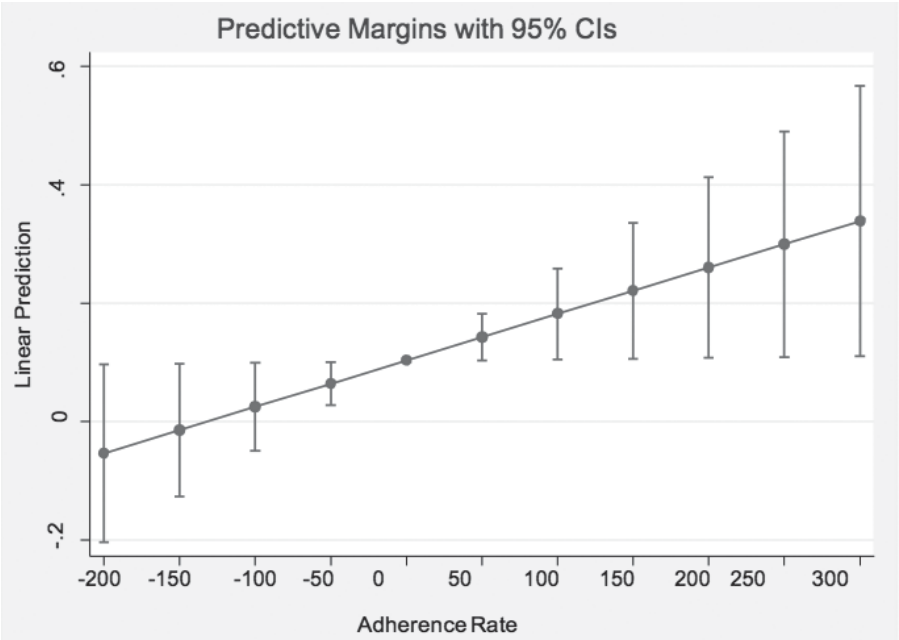
Figure 5

Table 6

Religious Adherence	Margins	P-Value	Confidence Intervals	
-200	-0.053	0.471	-0.203	0.097
-150	-0.014	0.797	-0.126	0.098
-100	0.025	0.496	-0.049	0.099
-50	0.064	0.001	0.028	0.101
0	0.103	0	0.101	0.107
50	0.143	0	0.103	0.182
100	0.182	0	0.105	0.259
150	0.221	0.001	0.106	0.336
200	0.26	0.002	0.107	0.413
250	0.23	0.003	0.109	0.49
300	0.339	0.005	0.11	0.567

*Data of the margins plot in a table

The battleground states include Florida, Pennsylvania, Michigan, North Carolina, Virginia, Colorado, Ohio, Wisconsin, Minnesota, Nevada, Arizona, New Mexico, New Hampshire, and Georgia.

REFERENCES

Albertson, Bethany L. 2011. "Religious Appeals and Implicit Attitudes." *Political Psychology* 32, no. 1 (February): 109–30.

Campbell, David E., John C. Green, and Geoffrey C. Layman. 2011. "The Party Faithful: Partisan Images, Candidate Religion, and the Electoral Impact of Party Identification." *American Journal of Political Science* 55, no. 1 (January): 42–58.

Chapp, Christopher B. 2012. *Religious Rhetoric and American Politics: The Endurance of Civil Religion in Electoral Campaigns.* Ithaca, NY: Cornell University Press.

Clinton, Hillary. 2016. "Remarks at Goodyear Hall and Theater in Akron, Ohio." The American Presidency Project, Gerhard Peters and John T. Woolley, October 3. Accessed December 7, 2017. www.presidency.ucsb.edu/ws/?pid=119153.

Clinton, Hillary. 2016. "Remarks at Martin Luther King, Jr., Plaza in Toledo, Ohio," The American Presidency Project, Gerhard Peters and John T. Woolley, October 3. Accessed December 7, 2017. www.presidency.ucsb.edu/ws/?pid=119154.

Clinton, Hillary. 2016. "Remarks at Wake Technical Community College in Raleigh, North Carolina." The American Presidency Project, Gerhard Peters and John T. Woolley, September 27. Accessed December 7, 2017. www.presidency.ucsb.edu/ws/?pid=119163.

Clinton, Hillary. 2016. "Remarks at the University of North Carolina in Greensboro." The American Presidency Project, Gerhard Peters and John T. Woolley, September 15. Accessed December 7, 2017. www.presidency.ucsb.edu/ws/?pid=119158.

Cruz, Ted. 2015. "Remarks Announcing Candidacy for President at Liberty University in Lynchburg, Virginia." The American Presidency Project, Gerhard Peters and John T. Woolley, March 23. Accessed December 7, 2017. www.presidency.ucsb.edu/ws/index.php?pid=109774.

Djupe, Paul A. and Brian R. Calfano. 2013. *God Talk: Experimenting with the Religious Causes of Public Opinion.* Philadelphia: Temple University Press.

FiveThirtyEight. 2016. "Who Will Win the Presidency?" 2016 Election Forecast, November 8. Accessed April 21, 2017. https://projects.fivethirtyeight.com/2016-election-forecast/.

Geographic Products Branch. 2012. "Cartographic Boundary Shapefiles - States." United States Census Bureau, September 1. Last modified April 6, 2017. Accessed April 7, 2017. www.census.gov/geo/maps-data/data/cbf/cbf_state.html.

Hart, Roderick P., Sharon E. Jarvis, William P. Jennings, and Deborah Smith-Howell. 2006. *Political Keywords: Using Language that Uses Us.* New York: Oxford University Press.

Karpowitz, Christopher F., J. Quin Monson, and Kelly D. Patterson. 2016. "Who's In and Who's Out: The Politics of Religious Norms." *Politics and Religions* 9, no. 3 (September): 508–36.

Kent, David. 2016. "Presidential Vote by Religious Affiliation and Race." Pew Research Center, November 9. www.pewresearch.org/fact-tank/2016/11/09/how-the- faithfulvoted-a-preliminary-2016-analysis/ft_16-11-09_relig_exitpoll_religrace/.

McLaughlin, Bryan and David Wise. 2014. "Cueing God: Religious Cues and Voter Support." *Politics and Religion* 7, no. 2 (June): 366–94.

Moffitt, Mary Anne. 1999. *Campaign Strategies and Message Design: A Practitioner's Guide from Start to Finish.* Westport, CT: Praeger.

Monson, J. Quin and J. Baxter Oliphant. 2007. "Microtargeting and the Instrumental Mobilization of Religious Conservatives." In *A Matter of Faith: Religion in the 2004 Presidential Election*, edited by David E. Campbell, 95-119. Washington, D.C.: Brookings Institution Press.

Newport, Frank. 2015. "Percentage of Christians in U.S. Drifting Down, but Still High." Gallup, December 24. http://news.gallup.com/poll/187955/percentage-christians-drifting-down-high.aspx.

Pew Research Center. 2016. "The Changing Composition of the Political Parties." In *The Parties on the Eve of the 2016 Election: Two Coalitions, Moving Further Apart.* September 13. Accessed March 1, 2017. www.people-press.org/2016/09/13/1-the-changing-composition-of-the-political-parties/.

Quinton, Adam. Database of Presidential Campaign Speeches. Created for and provided to the students of this project. Licensed. https://github.com/zvakanaka/candidate-speeches.

Sanders, Bernie. 2015. "Remarks at the Liberty University Convocation in Lynchburg, Virginia." *The American Presidency Project*, Gerhard Peters and John T. Woolley, September 14. www.presidency.ucsb.edu/ws/?pid=114488.

Sanders, Bernie. 2016. "Remarks on Policy in the Middle East in Salt Lake City, Utah." *The American Presidency Project*, Gerhard Peters and John T. Woolley, March 21. www.presidency.ucsb.edu/ws/?pid=116694.

Smith, Gregory A. 2016. "Churchgoing Republicans, Once Skeptical of Trump, Now Support Him." Pew Research Center, July 21. www.pewresearch.org/fact-tank/2016/07/21/churchgoing-republicans-once-skeptical-of-trump-now-support-him/.

The Association of Religion Data Archives. Accessed February 03, 2017. www.thearda.com.

The Green Papers. 2016. *Election 2016: Presidential Primaries, Caucuses, and Conventions.* Last modified March 4, 2018. Accessed March 7, 2018. www.thegreenpapers.com/P16/.

Trent, Judith S. and Robert V. Friedenberg. 2000. *Political Campaign Communication: Principles and Practices.* 4th ed. Westport, CT: Praeger.

Trump, Donald J. 2015. "Remarks Announcing Candidacy for President in New York City." *The American Presidency Project*, Gerhard Peters and John T. Woolley, June 16. www.presidency.ucsb.edu/ws/?pid=110306.

Trump, Donald J. 2016. "Remarks at a Rally at the Greenville Convention Center in Greenville, North Carolina." *The American Presidency Project*, Gerhard Peters and John T. Woolley, September 6. www.presidency.ucsb.edu/ws/?pid=119197.

Trump, Donald J. 2016. "Remarks at High Point University in High Point, North Carolina." *The American Presidency Project*, Gerhard Peters and John T. Woolley, September 20. www.presidency.ucsb.edu/ws/index.php?pid=119192.

Woolley, John and Gerhard Peters. 2017. "Election of 2016." The American Presidency Project. Last modified February 16. Accessed February 3, 2017. www.presidency.ucsb.edu/showelection.php?year=2016.

Congress and the Benefits of Social Media: A Low-Cost Investment in Constituency Approval and Electoral Success

Mandi Eatough

Media in the U.S. has experienced rapid changes since the introduction of the Internet to mainstream society. Since the 1990s, the population of the world that is active on the Internet has grown from millions to billions. Worldwide, people began to want more information, and they wanted to have access to the information instantly. While newspapers, news-based radio stations, and television news providers have tried to keep up with the continuously evolving media outlets, a different media form has emerged to fill the gap (Shirkey 2011). Social media web sites such as Facebook, Twitter, and YouTube have become simple ways for individuals to obtain news with minimal effort on their part. As social media has integrated into the lives of the American public, it has also integrated into American politics.

As early as the 2008 election cycle, the presence of social media in politics has been considered an important part of political news. Local and federal legislators and legislative candidates alike have taken to social media as a way of creating a pseudo-personal relationship with their constituents or voters. While commonly acknowledged that many legislators have jumped on the social media bandwagon, what has not been examined is the benefits legislators or candidates may receive from social media presence and activity. This research aims to answer what benefits senators and Senate candidates see from investment in social media, as well as what differences are seen in these benefits based on gender. This research also takes into account existing theories about media coverage and gender for legislators and legislative candidates in the U.S. that suggest men often benefit more from personalized media messaging.

Using pre-existing data from the Cooperative Congressional Election Study (CCES) and original data collected about social media sites belonging to sitting senators, as well as Senate candidates, this research will examine a three-part hypothesis.

First, sitting senators who have higher levels of social media presence and activity will experience higher rates of name-recognition and approval among their constituents than members with lower levels of social media presence. Second, Senate candidates who have higher levels of social media presence and activity will see an increase in vote share, meaning they will be more likely to be re-elected. Finally, this paper will examine differences in the benefits of social media presence and activity for both senators and Senate candidates based on gender, with the expectation that male candidates will receive greater benefits from social media investment.

This paper aims to provide sufficient background about the theories behind the research performed, as well as detailed information about the studies and their results. It will begin with an overview of the existing literature regarding the behaviors of legislators on social media and the existing theories of how social media influences politics in the U.S., with a specific focus on differences in senator and Senate candidate gender. Next will be a detailed outline of the research question, as well as the hypotheses. The paper will then provide a detailed explanation of how the data from 2012, 2014, and 2016 CCES datasets and the collected original data have combined to create a model to address my hypotheses. Finally, the models and results will be presented with concluding comments about the implications of the results and suggestions for further research on this topic.

Theoretical Background
Social Media and Politics

Since the introduction of social media in the early 2000s, political scientists have identified that the media platform has political uses. The study of the influence of social media on legislators and constituents alike has led to several theories about how social media is political. Several studies have identified social media as an important way for politicians to interact with their constituencies while in Washington (Watall et al. 2010; Gainous and Wagner 2014; Fitch and Goldschmidt 2014), while others have focused on the political usage of social media by constituents (Hand and Ching 2011; Shirkey 2011; Carlisle and Patton 2013). Each study has come to the same conclusion: Social media matters in politics.

Social media is the most likely outlet for constituents of any age to invest themselves in the political process (Carlisle and Patton 2013). In recent years, the accessibility of social media for most citizens in the U.S. has increased while more "traditional" news methods (newspapers, news-based radio, news television outlets) have become less accessible. Social media has easily become more convenient than more traditional news outlets for many citizens as computers, smartphones, and wireless hotspots have increased in popularity throughout the country. In addition to convenience, social media is primarily available at no cost to users. While it is free to sign up for an account on most social media sites (this is true for Facebook, Twitter, and YouTube), newspapers and television often cost money to access some or all of their content. This makes social media a monetarily advantageous news outlet for both legislators,

who may disseminate news from their office for incredibly low cost, and constituents, who may receive news from their legislators at little to no additional cost.

It is important to note that while social media is the most convenient news outlet regardless of age, many political scientists have identified that younger citizens in the U.S. (those born after 1980) are the most likely to be influenced by politicians on social media (Samuels 2011; Kahne and Middaugh 2012; Norval 2012). Younger citizens use social media as a news outlet for both political and non-political information and are much less likely to seek out news from more traditional news outlets than older citizens (Kahne and Middaugh 2012). This same demographic of young citizens uses social media to maintain their own political behavior as well as influence the behavior of others. Social media is used by younger citizens to organize protests and to engage in political discussions with those they have connected with on social media outlets. Political interactions on Facebook have been shown to be as effective for political socialization among the younger generation as political interactions that occur in person (Norval 2012). Interactions online are considered just as valuable by these citizens as those they have in person, making social media a particularly advantageous outlet for politicians to connect with younger constituents.

The Costs and Benefits of Social Media for Legislators

Social media is unique as a media form in the fact that, at a basic level, it is available to everyone at little or no additional cost. As discussed previously, social media has become increasingly popular in the U.S., particularly when compared to "traditional" forms of media predominantly because of differences in cost. Given the low cost levels associated with using social media for both legislators and constituents, it is expected that legislators believe that investment in this media platform will be beneficial. What is less clear is where these benefits actually are seen for legislators. Existing research has suggested two main theoretical benefits of digital connections made between legislators and their constituents who connect with them on social media: 1) potential benefits that legislators who are currently in office receive from using social media to interact directly with their constituency and 2) potential benefits of social media usage in elections.

For sitting senators, media use is an important tool to connect with their constituencies. While it has not been examined on the congressional level, a study by Hand and Ching showed that the presence of government-centric Facebook pages provides a way for citizens to directly engage with their local political leaders (2011). Their research showed that when local government officials maintained Facebook pages that were regularly updated with information about local politics, there was an increase in the approval of local legislative bodies. Several studies have suggested that the more time a legislator spends interacting directly with their constituents, the more likely they are to be approved of by those same constituents (Fenno 1975; Yiannakis 1982; DiGrazia et al. 2013; Fitch and Goldschmidt 2014; Sulkin, Testa, and Usry 2015). If we assume that in the modern world individuals value interactions through social

media at least partially as much as they value in-person interactions, there should be a direct correlation between the social media presence of a legislator and approval of the legislator similarly to the correlation seen on the local government level.

However, many political scientists suggest that the greatest benefit politicians can gain from using social media is by using it as a tool during an election. In recent years, our social and political lives have become increasingly based on personal attributes as they are presented online (Bennett 2012; Jacobs and Spierings 2014). Elections are not exclusively about who will do the best job but who resonates the most with the voters. As previously discussed, social media is one of the easiest ways for a legislator to interact with their constituents. Existing research suggests that social media activity by legislators creates a sense of transparency and accountability between legislators and their constituents that more traditional news outlets do not (Fitch and Goldschmidt 2014). This leads to a consistently high level of social media presence among legislators, particularly when leading up to elections, so we should expect a similar effect on candidates running for these offices.

In a study about the 2008 presidential election cycle and social media, it showed that Facebook users interact with more posts about politics during an election cycle than at any other time (Carlisle and Patton 2013). This suggests that politicians are likely to get the most interaction with constituents on social media during an election. Many political scientists agree that political campaigns worldwide have been moving toward Internet-based media outlets (Williamson 2010; Obar et al. 2012; Jacobs and Spierings 2014). Social media is providing a new way for legislators to organize and fund-raise for their campaigns (Siegel 2012). Candidates no longer have to focus their media efforts exclusively on who can get the most time with local and national media outlets, because social media is always available to them.

A comparative study of voting in 2010 elections in the Netherlands guides a significant portion of this research (Jacobs and Spierings 2014). Using information about presence and activity on Twitter for 493 candidates, their study evaluated the impact of differences in presence and activity on the platform on vote share. They found that when a candidate has a high number of Twitter followers and actively uses the social media platform, they are likely to see an increase in vote share. Their work emphasizes the importance of active use of the platform, indicating that candidates who maintained a profile on Twitter but did not regularly post saw no increase in vote share based on their number of followers. This study gives an indication of what expectations we should have when considering similar elements of social media presence and activity in U.S. congressional elections.

It is uncontested among political scientists that social media is having an influence on politics in the United States. What has remained unclear is the extent to which investment in social media as a political tool leads to direct benefits for legislators and candidates. As outlined above, current theory regarding social media and politics suggests there are two main areas where legislators may see direct benefits from

social media: increased levels of recognition and favorability among constituents for sitting legislators and increased favorability in elections for candidates. This research uses an original observational measure of the social media presence and activity of senators and Senate candidates to further examine the theorized benefits of social media for legislators in the U.S. beyond what has been studied previously.

The Gendered Expectations of the Effect of Social Media in Politics

This research also takes into account existing theories about media coverage and gender for legislators and legislative candidates in the United States. Historically, male and female senators have differed in the way they have chosen to communicate with their constituents (Alexander and Andersen 1993; Kahn 1992; Kahn 1994; Fridkin and Kenney 2014). Generally, senators tend to play toward stereotypical strengths based on gender when interacting with their constituents, and the media tends to follow this lead. Male senators are often covered for their legislative accomplishments on high saliency issues particularly those that are related to economic or international issues. Comparatively, female senators are often covered for work on "women's issues"—issues where women traditionally bear more of the associated costs with the legislation, such as abortion or childcare (Rosenwasser and Dean 1989).

Given the historical differences of media messaging among U.S. senators based on gender, we may expect that senators present themselves in similar ways on social media. However, recent studies on differences in social media activity of candidates in the U.S. suggest candidates engage in distinct ways with voters on social media based on their gender in a way that differs from traditional media coverage (McGregor et al. 2016; Meeks 2017). These studies have found that while both male and female candidates seem to be using social media at similar levels during campaigns, the content of their engagement and the potential benefit of this engagement differs based on gender. They find that while male candidates benefit from social media personalization (maintaining personality through their social media pages by discussing hobbies, family, or non-campaign related topics on their pages), female candidates are often disadvantaged by this behavior. While this helps men to be seen as personable, women are more likely to be seen as unfocused in their careers. This means that, instead of increased favorability among voters, this personalization forms a perception of incompetency among female legislators and candidates.

To win an election, a candidate must be perceived as both likable and competent by voters. This, of course, requires a balance for all candidates, but women face specific challenges in maintaining this balance. Research on gender representation has suggested there is a "double-bind" that prevents women from winning elected offices. The expected behaviors of women in society often are in direct conflict with the expected behaviors of elected leaders creating a scenario in which women often are perceived as likable but incompetent or competent but unlikable by voters (Gimenez et al. 2017).

Given these outlined differences in political media engagement based on gender, it is reasonable to expect differences in the benefits that senators and Senate

candidates received based on their gender. Since social media has been identified as the platform for candidate personalization by a significant amount of the existing research, we should expect that male candidates, and potentially male senators, will be more likely to experience the theorized benefits of investment in social media than their female counterparts.

Research Question and Hypotheses

This research aims to answer the following questions:

Do senators and/or Senate candidates see benefits in constituency recognition/ approval or vote share when they increase their investment in social media?

Assuming there are benefits to social media investment, do male and female senators and Senate candidate experience these benefits equally?

This project uses a combination of existing survey data and original observational data to evaluate changes in these potential benefits based on social media presence and activity as well as potential differences in these benefits based on gender. The hypothesized outcomes are as follows:

H1: Sitting senators with higher levels of social media presence and activity will experience greater constituency recognition and approval.

1. The effect of social media activity on constituency recognition and approval will be greater than the effect of social media presence.
2. Male senators will experience higher levels of constituency recognition and approval based on social media investment than their female counterparts.

H2: Senate candidates with higher levels of social media presence and activity will receive a greater amount of the vote share.

1. The effect of social media activity on vote share will be greater than the effect of social media presence.
2. Male Senate candidates will receive a greater increase in vote share based on social media investment than their female counterparts.

H3: Controlling for partisanship and other demographic factors, citizens in the 2016 election cycle with higher levels of political social media engagement will be more likely to vote for candidates with higher levels of social media presence and activity.

1. The effect of social media activity on vote choice will be greater than the effect of social media presence.
2. Male Senate candidates will be favored more highly for social media investment by citizens with high levels of political social media engagement than their female counterparts.

Data and Methodology

This study uses a combination of existing survey data and original observational data. The existing survey datasets this study uses are from the 2012, 2014, and 2016 Cooperative Congressional Election Study (CCES).[1] Each CCES dataset includes information for over 50,000 individuals throughout the United States. This research specifically uses information about constituency approval and name recollection for the senators of each respondent, as well as their senator vote choice, when applicable, from all three datasets. This research also uses additional data from the 2016 CCES about media engagement and political social media behaviors of respondents. For a complete list of CCES questions used from each year see Appendix.

The original data collected about the social media presence and activity of senators and Senate candidates was compiled using The Wayback Machine, a digital archive of web sites created by the Internet Archive. This tool allows archived web sites to be viewed as they were presented on specific dates that provides a way to accurately measure the previous social media presence and activity of both senators and Senate candidates across several years. In the case of senators, social media presence and activity were measured based on archives of social media pages in January of the first year of each congressional session. In the case of Senate candidates, social media presence and activity were measured based on archives of social media pages before, but as close as possible,[2] to Election Day of each election year. The social media pages for Senate candidates were only measured based on pre-Election Day dates because of anticipated changes in follower counts and social media activity that are associated with winning or losing elections. Information regarding social media presence and activity were collected based on the archived pages of Facebook, Twitter, and YouTube profiles for senators of the 112th–114th Congresses, as well as candidates in the 2012, 2014, and 2016 senatorial elections.

Social media presence is defined in this study based solely on whether or not the senator or candidate has a profile on a social media platform. If a senator or candidate does not have a profile on a platform, they do not have a presence on that platform. Differences in social media presence among senators and candidates have decreased significantly over time. In 2012, several senators and candidates had a singular social media presence and a few had no social media presence. Comparatively, in 2016 every senator had a social media presence on each platform with the single exception of Senator Risch from Idaho who did not join Facebook until mid-2017. A similar change was seen among Senate candidates in 2016 with nearly all candidates having a presence on Facebook and Twitter and over half of the candidates with a presence on YouTube. Percentages of social media presence by year for both senators and Senate candidates is shown in Table 1 (p. 50). Social media activity is defined by how frequently a legislator uses a social media platform to interact with their constituents. The information collected on activity varies slightly based on what information can be collected about the profile from each social media site.

For each platform that a senator or Senate candidate maintained a social media presence on, information was collected about how long they had maintained that presence, as well as how many followers they had on that site. The way this was determined varies slightly between the three sites. Below is a brief explanation of the differences in data collection between Facebook, Twitter, and YouTube (A breakdown of what data was collected from each site is found in Table 2 on p. 55):

Facebook: The amount of time a senator or candidate maintained a presence on Facebook was coded as the year of the first post available on their Facebook page.

Twitter: The amount of time a senator or candidate maintained a presence on Twitter was coded as the year they joined as provided by Twitter. Follower count was coded as the rounded value provided on the profile by Twitter (generally rounded to the nearest thousand). The number of Tweets was coded as the exact value provided on each profile by Twitter.

YouTube: The amount of time a senator or candidate maintained a presence on YouTube was coded as the year they joined as provided by YouTube. Follower count was coded as the exact follower value provided on the page by YouTube. Number of views is coded as the exact value provided on the information page for each YouTube account.

Using a combination of these two data sources this research will conduct two studies. The first is an analysis of how the social media presence and activity of senators influences their name recognition and favorability among their constituents. This analysis will be conducted using logistic regression models of name recognition and favorability for members of the 112th, 113th, and 114th Congresses. The second study is an analysis of how the social media presence and activity of Senate candidates influences their vote share. This analysis will be conducted using OLS regression models of total vote share, as well as a mixture of OLS and logistic regression models of individual constituency vote choice and overall election results in the 2012, 2014, and 2016 senatorial elections.

Analyses and Results

Analysis 1: The Effect of the Social Media Presence and Activity of Senators on Constituency Recognition and Approval

Existing theory suggests that Facebook, Twitter, and YouTube are the most influential social media platforms in terms of political impact. Senators, more so than Senate candidates as discussed in the following section, are overwhelmingly present on all three of these platforms and have been since at least 2012 (see Figure 1). Twitter has been identified as a particularly important platform for political figures, with most research about the influence of social media on politics focusing on Twitter behavior. The data collected for this research allowed for a comparison of the effect of presence and activity on each social media platform. Given the information collectible from these platforms (see Table

2), the best measure for a comparison across all three platforms was a measure of constituency engagement (see Figure 2).
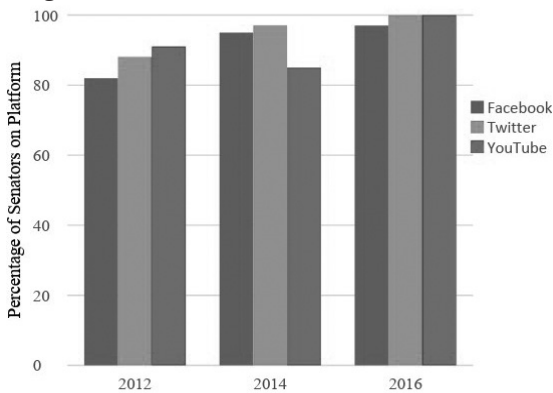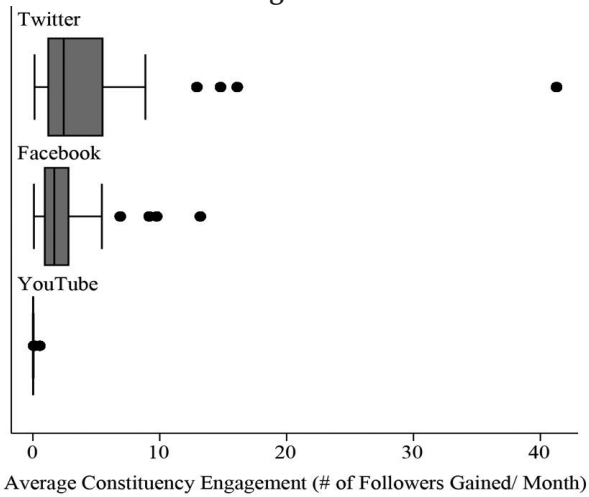
Figure 1: Social Media Presence of Senators

Figure 2

Average Constituency Engagement (# of Followers Gained/ Month)

This measure was created by dividing the number of followers on the platform by the number of months the senator had been on the platform. While imperfect, this gives a rough estimate of the level of constituency engagement that occurred on each platform for each senator. Twitter shows the highest level of constituency engagement across nearly all senators. Given the high level of constituency engagement on Twitter, this analysis, as well as the following analysis of Senate candidates, will focus on the social media activity of senators on Twitter.[3] The following logistic regression model was used to evaluate the effects of both social media presence and activity on senator name recall, party identification, and approval:

Pr (dependent variable=1 | X=1,2…)

> *=F(primary independent variable (social media presence or activity) + respondent social media use*

> *+ respondent political social media use + senator partisanship + respondent partisanship*

> *+ partisanship match + senator gender + respondent gender + gender match + respondent race*

> *+ respondent age + respondent income + state fixed effects)*

Using this model, six logistic regressions were conducted to evaluate the effects of both social media presence and activity on the recognition and approval of senators by their constituents. First, analyzing the social media presence of senators, logistic models were used to evaluate the effect of social media presence on constituency recognition and approval. Based on existing theory, it is hypothesized that an increase in social media presence should lead to increases in recognition and approval among constituents. The results of these models (see Appendix 1, Table 3) were mostly statistically and substantively insignificant. In one case, the predicted probability of a constituent being able to recall the name of their senators is statistically significant at the 99 percent confidence level, with a predicted negative effect on name recall as the social media presence of a senator increases. The predicted probability of a constituent correctly recalling the name of a senator based on the senator's social media presence is shown in Figure 3. When comparing a senator with no social media presence, to a senator who is present on all three social media platforms, it is predicted that there will be a ten percentage point decrease in the probability of a constituent recalling their senator's name. It is important to recognize a limitation of this model in the discussion of the results, because, in the 114th Congress,[4] the only variation in social media presence is between senators with and without Facebook as all senators were present on both Twitter and YouTube. Given the high social media presence of senators across all three years, it is unlikely that there will be particularly large differences in recognition or approval based on social media presence, and much of the variation is based in predicted values rather than observations, giving much larger standard errors.

Figure 3: Effect of Senator Social Media Presence on Name Recall

Predicted Probability of Senator Recognition (2016)
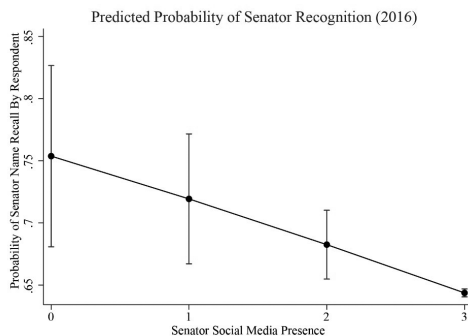
While senator social media presence showed little effect on constituency recognition and approval, logistic models of senator social media activity showed consistent positive and statistically significant effects on these factors (see Table 1). The predicted probabilities of a constituent recalling the senator's name, correctly identifying the senator's political party affiliation, and approving of the senator based on the Twitter activity of their senator are shown in Figure 4. The effect of senator Twitter activity is the smallest on the probability of constituent approval of their senator, with about a one percentage point increase if a senator were to go from not tweeting at all to tweeting about 200 times per month. However, the effect of senator Twitter activity has a much larger effect on the probability of a constituent recalling the name of their senator or correctly identifying the partisanship of their senator.

Figure 4: Effect of Senator Social Media Activity on Recognition and Approval

Predicted Effects of Senator Social Media Activity
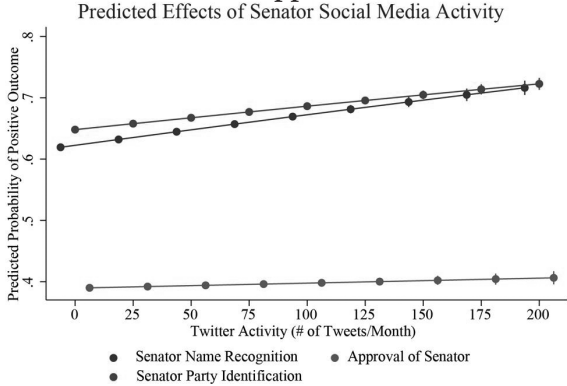
In the case of name recall, if a senator were to go from not tweeting at all to tweeting about 200 times per month, their constituents would have a ten percentage point increase in the probability of a constituent remembering their name. In the case of partisan identification, the same shift in Twitter activity would have an eight percentage point increase in the probability of a correct partisan identification by a constituent. Since these are logistic models there is inconsistency in the amount of change predicted in the probability of a positive outcome across increases in senator Twitter activity. However, in the case of senator name recall, it takes an increase of about twenty tweets per month by the senator for an approximate one percent increase in the probability of a positive outcome. In the case of senator party identification, it takes an increase of about twenty-five tweets per month to see an approximate one percent increase in the probability of a positive outcome. While these predicted effects are not particularly large, the results suggest that increased Twitter activity by senators does positively influence the probability of constituents having greater recognition of them. It is also important to note that these effects are seen while controlling for factors that have a much larger effect on constituent recognition of senators, including matching partisanship between constituents and senators.

Table 1: Logistic Regressions of the Effect of Senator Social Media Activity

| VARIABLES	Senator Name Recall	Senator Party Identification	Senator Approval
Senator Twitter Activity	0.00264***	0.00227***	0.000445**
	(0.000222)	(0.000204)	(0.000190)
Res. Social Media Use	-0.418***	-0.309***	-0.118***
	(0.0229)	(0.0207)	(0.0203)
Res. Political Social Media Use	0.306***	0.246***	0.0811***
	(0.00717)	(0.00624)	(0.00555)
Senator Republican	0.370***	0.352***	-0.281***
	(0.0190)	(0.0169)	(0.0162)
Respondent Republican	-0.0538**	-0.0933***	-0.555***
	(0.0225)	(0.0193)	(0.0185)
Partisan Match	-	2.042***	2.247***
		(0.0244)	(0.0171)
Senator Male	-0.213***	-0.290***	-0.166***
	(0.0214)	(0.0188)	(0.0179)
Respondent Male	0.629***	0.517***	0.184***
	(0.0203)	(0.0179)	(0.0169)
Gender Match	0.0450**	0.0684***	0.0755***
	(0.0201)	(0.0177)	(0.0166)
Respondent White	0.129***	0.195***	0.150***
	(0.0192)	(0.0173)	(0.0175)
Respondent Age	-0.0409***	-0.0420***	-0.0129***
	(0.000577)	(0.000520)	(0.000472)
Respondent Income	0.115***	0.108***	0.0311***
	(0.00292)	(0.00260)	(0.00212)
State Fixed Effects?	YES	YES	YES
Constant	79.92***	81.79***	24.27***
	(1.136)	(1.022)	(0.929)
Observations	73,292	101,298	101,298

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Analysis 2: The Effect of the Social Media Presence and Activity of Senate Candidates on Voter Support

As discussed in the first analysis, Facebook, Twitter, and YouTube are the most influential social media platforms in terms of political impact. The variation in social

media presence on these platforms is much larger among Senate candidates than it is among senators as shown in Figure 5 (Senator social media presence distribution from Analysis 1 is shown in Figure 1). As discussed in analysis 1, the constituency engagement measures for Senate candidates look essentially identical to those of senators, indicating once again that Twitter activity is likely the best platform for social media activity to have an influence on constituent behavior.

Figure 5: Social Media Presence of Senate Candidates over Time
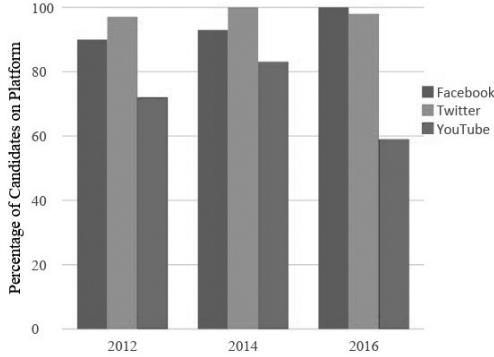
The analysis of the effect of Senate candidate social media presence and activity has been conducted in two parts: first, an analysis of the effect of Senate candidate social media presence and activity on election results, measured with OLS regressions of the effects on vote share and logistic.

Predicted Candidate Vote Share=β_0 + β_1 primary independent variable (social media presence or activity)

+β_2 candidate partisanship + β_3 candidate gender + β_4 state fixed effects + δ

Pr(candidate winning election=1 |X=1,2…)

=F(primary independent variable (social media presence or activity) + candidate partisanship

+ candidate gender + state fixed effects)

Second, an analysis of the effect of Senate candidate social media presence and activity on individual constituency vote choice using the following logistic regression model:

Pr(vote choice=1 |X=1,2…)

=F(primary independent variable (social media presence or activity) + respondent social media use

*+ respondent political social media use + candidate partisanship + respondent partisanship*

+ partisanship match + candidate gender + respondent gender + gender match + respondent race

+ respondent age + respondent income + state fixed effects)

These analyses allow for a consideration of the effect of social media presence and activity on senator election outcomes in terms of both the overall election result as well as individual vote choice. The evaluation of both election results and individual vote choice allows for the consideration of benefits in terms of general electoral success. With this evaluation, we can also consider which voters are most responsive to social media activity, giving us a greater understanding of where any differences in electoral success may be coming from based on social media presence or activity.

Data was collected about the social media presence and activity of Senate candidates in the 2012, 2014, and 2016 congressional elections. However, an effect of social media on vote choice or electoral outcomes does not appear in analyses until the 2016 congressional elections. Figure six shows a comparison of the relationship between Senate candidate social media activity and expected vote share in 2012, 2014, and 2016 Senate elections based on the results of the OLS model outlined for each year. These results clearly indicate that the benefit of social media activity on vote share for Senate candidates is unique to the 2016 Senate elections.[5] Given the unique importance of social media in the 2016 election, this analysis will focus on the effects of Senate candidate social media presence and activity on election outcomes and individual vote choice in 2016.

Figure 6: The Effect of Social Media Activity of Senate Candidates over Time
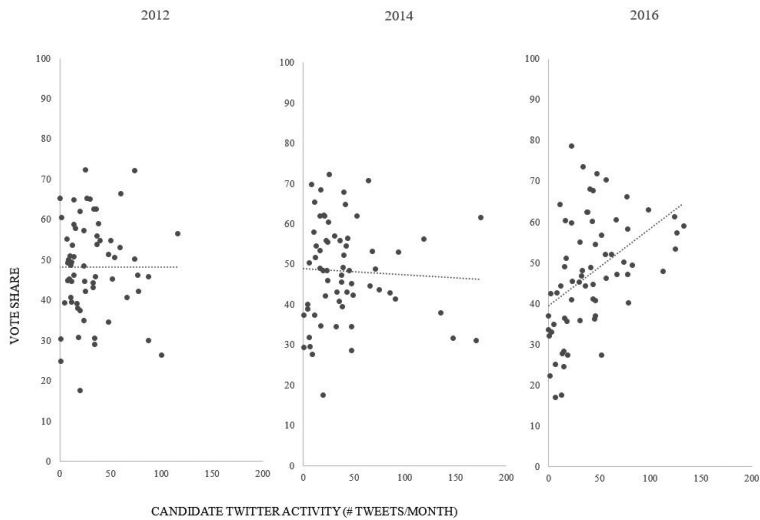
The Effect of Senate Candidate Social Media Presence and Activity on Election Outcomes

As outlined earlier in this section, the effect of Senate candidate social media presence and activity on election outcomes was measured using OLS regressions of these factors on candidate vote share and logistic regressions of these factors on the final electoral outcome (see Appendix Table 4). In each of these models, increases in social media

investment by Senate candidates leads to statistically and substantively significant positive changes in election outcomes. The predicted effects of social media investment by Senate candidates on both vote share and the result of their electoral race are shown in Figure 7 and Figure 8.
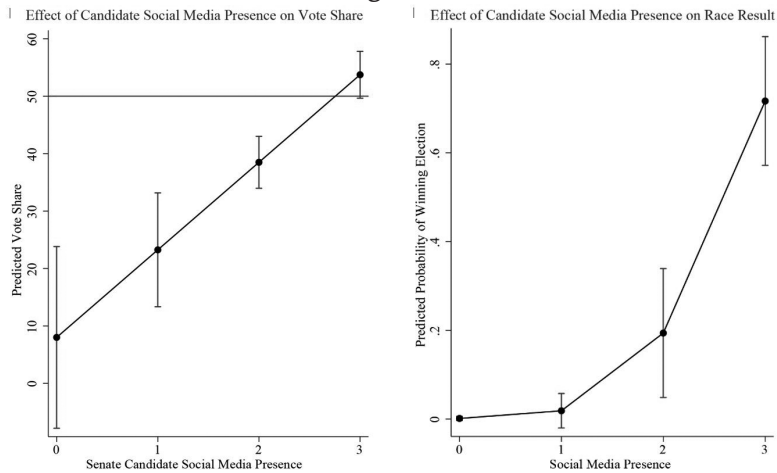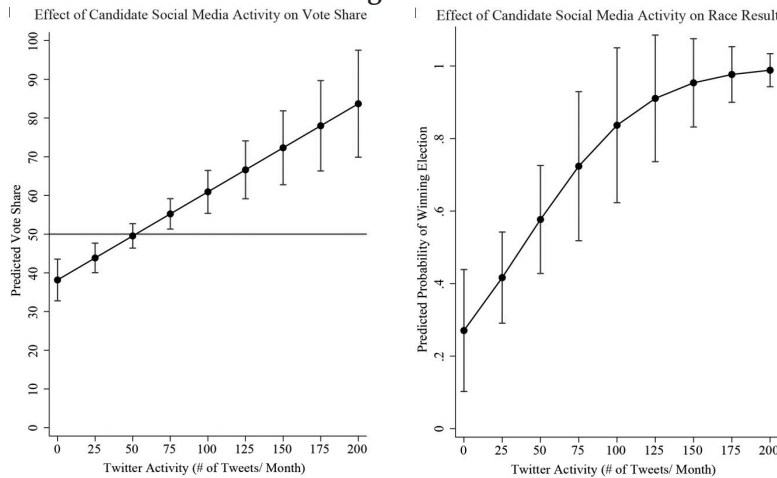
Figure 7

Effect of Candidate Social Media Presence on Vote Share | Effect of Candidate Social Media Presence on Race Result

Figure 8

Effect of Candidate Social Media Activity on Vote Share | Effect of Candidate Social Media Activity on Race Result

If a Senate candidate were to go from not being present on any social media platforms to being present on Facebook, Twitter, and YouTube, all else remaining the same, the predicted vote share would increase from just under 10 percent to just over 50 percent. Even if a Senate candidate were to increase their social media presence by a single platform, they would experience a vote share increase of around 15 percent. Similarly, in the consideration of race outcome, if a Senate candidate were to go from having

no social media presence to being present on all three platforms, their predicted probability of winning their election would increase by over sixty percentage points from essentially no chance of winning initially.

Evaluating Senate candidate social media activity (in the form of Twitter activity as outlined previously), we again see significant positive increases in vote share and probability of winning an election as the candidate increases their investment in social media presence. For each five-tweet increase per month by a Senate candidate, the candidate experiences a one percentage point increase in vote share. This means that if a candidate were to increase their Twitter activity by at least one tweet per day, they would receive approximately a six percentage point increase in vote share. Senate candidates also see significant benefits in their probability of winning an election as their Twitter activity increases. Candidates who tweeted at least fifty Tweets per month on average were pushed above the .5 probability of winning their election. In comparison, if a candidate were to not tweet at all, they have a .30 probability of winning their election. For candidates with the highest levels of Twitter activity, around 200 tweets per month, their predicted probability of winning their election is .98.

The Effect of Senate Candidate Social Media Presence and Activity on Individual Vote Choice

The effect of Senate candidate social media presence and activity on individual vote choice was measured using logistic regressions as outlined previously (see Table 2). In both models, the effect of social media investment by a Senate candidate leads to statistically and substantively significant positive increases in the probability of an individual voting for the candidate. The predicted probability of individual vote choice based on a Senate candidate's social media investment is shown in Figure 9.

Figure 9

Effect of Candidate Social Media Presence on Vote Choice | Effect of Candidate Twitter Activity on Vote Choice
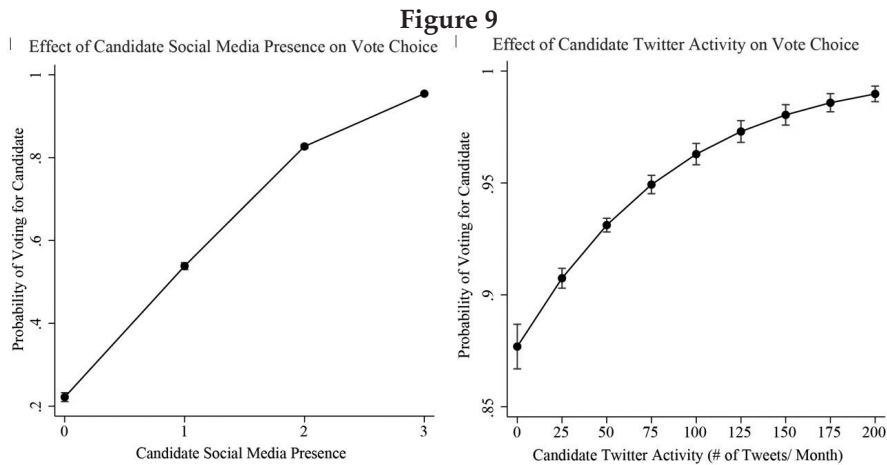
Senate candidates with no social media presence have a predicted probability of gaining an individual vote, all else including partisanship held equal, of just above .2. If a Senate candidate has a social media presence on at least one social media platform,

Table 2: Logistic Regressions of the Effect of Candidate Social Media Presence and Activity on Individual Vote Choice

VARIABLES	(1) Vote Choice	(2) Vote Choice
Candidate Social Media Presence	1.540*** (0.0184)	-
Candidate Twitter Activity	-	0.0133*** (0.00109)
Res. Social Media Use	0.0934* (0.0504)	0.101 (0.0615)
Res. Political Social Media Use	0.0120 (0.0143)	0.0207 (0.0181)
Candidate Republican	-	-
Respondent Republican	0.453*** (0.0497)	0.425*** (0.0535)
Partisan Match	0.856*** (0.0428)	1.311*** (0.0584)
Candidate Male	1.043*** (0.0415)	0.520*** (0.0483)
Respondent Male	-0.0232 (0.0416)	-0.0219 (0.0467)
Gender Match	0.150*** (0.0407)	0.173*** (0.0453)
Respondent White	0.435*** (0.0425)	0.286*** (0.0537)
Respondent Age	-0.00826*** (0.00122)	-0.0107*** (0.00151)
Respondent Income	-0.00223 (0.00533)	0.0289*** (0.00705)
State Fixed Effects?	YES	YES
Constant	13.38*** (2.394)	21.66*** (2.967)
Observations	33,577	29,278

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

their predicted probability of an individual voting for them increases to .53, giving them a higher probability of receiving the vote than not. Senate candidates who are present on all three social media platforms have a predicted probability of an individual voting for them of .95.

While the predicted effects of Senate candidate social media presence are huge in changing the probability of an individual voting for a candidate, the effect of Twitter activity simply increases an already high probability of vote choice for all candidates present on Twitter. Candidates who are on Twitter but do not tweet at all have a predicted probability of an individual voting for them of .87. Compared to candidates with the highest level of Twitter activity (candidates tweeting around 200 times per month who

have a predicted probability of an individual voting for them of .98), candidates with low levels of Twitter activity are at a slight disadvantage. However, while the effect of Senate candidate Twitter activity causes a notable positive increase in the probability of an individual voting for the candidate, the substantive effect is much smaller than the substantive effect of social media presence on individual vote choice, going against the hypothesis that social media activity would have a larger impact on vote choice than social media presence.

Finally, the third hypothesis of this research suggests that, when controlling for partisanship and other demographic factors, citizens in the 2016 election cycle with higher levels of political social media engagement will be more likely to vote for candidates with higher levels of social media presence and activity. In only one case, when considering the impact of social media presence on vote choice, is the effect of social media use by a constituent statistically significant on vote choice. The political social media presence of constituents shows no effect on vote choice.

Figure 10
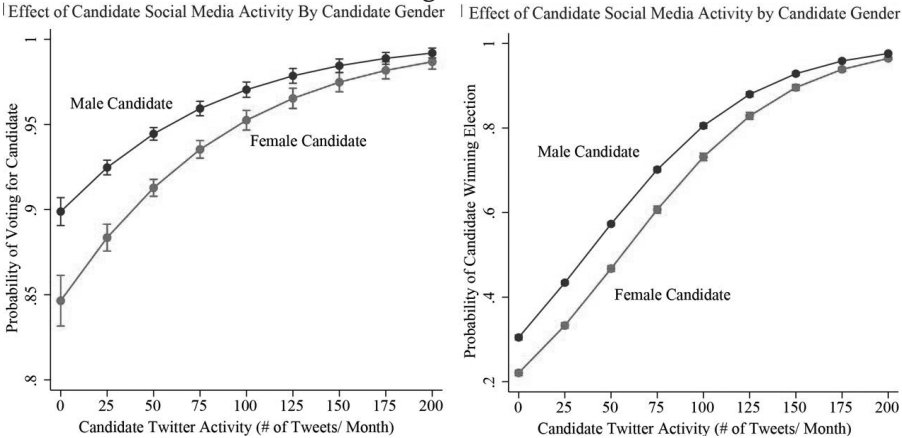
The Effects of Gender

Theories about the social media behaviors of politicians and candidates suggest that there may be gendered differences in the benefits these individuals receive from social media investment. Many theoretical arguments for this gendered difference rely on the idea that social media is inherently more personal than other forms of media and that male politicians and candidates are more likely to benefit from their constituents or voters seeing their "personal side." This may be because women in these positions are more likely to be seen as unreliable or unprofessional for sharing personal or emotional information, particularly when it is about their families. Given this theoretical argument, it is expected that men and women will see differing benefits from social media investment, particularly when it comes to social media activity.

In terms of gendered differences among sitting senators, female senators are slightly more likely than male senators to be recognized and approved of by their constituents regardless of senator social media engagement. This is particularly true among female constituents of female senators. However, the major difference in benefit of social media investment is seen among Senate candidates. The predicted effects of candidate social media activity on the probability of an individual voting for a candidate, as well as a candidate winning their election based on gender, are shown in Figure 10. The predicted effects indicate there is a gender gap in the benefit Senate candidates receive from social media activity that slowly decreases as candidates approach higher levels of social media activity. This suggests female Senate candidates can close the gender gap and benefit from social media engagement with constituents by raising their level of social media activity compared to their male counterparts.

In addition to these gendered candidate effects, this research also indicates there are gender differences among constituents in terms of both senator recognition and approval. Men are more likely in all cases to recall the name and correctly identify the partisanship of their senator than women are. The men are also slightly more likely than women to approve of their senator. There is no significant difference in individual vote choice based on the gender of the respondent, except for cases where the candidate and the respondent are of the same gender. In these cases, there is a significant and strong increase in the likelihood of an individual voting for a candidate.

Discussion

Through several analyses, this paper has shown that senators and Senate candidates experience benefits in terms of constituency recognition and approval and electoral outcomes from increased investment in social media. While not all of the theoretical hypotheses of this research were confirmed, the analyses show in what conditions senators and Senate candidates are most likely to experience benefits from both social media presence and activity.

The first analysis showed that the benefits sitting senators receive from increased social media engagement are entirely based on social media activity rather than social media presence. While the increases in constituency recognition and approval from increased levels of Twitter activity for senators in the 114th Congress are small, they are consistently positive and statistically significant. When compared to members of the 112th and 113th Congresses, where all results were null, these findings indicate an increase in the benefits senators see from increased social media activity over time. This may be attributable to the increase in social media use over time (with the 112th Congress beginning in 2011 and this study focusing on the 114th Congress in 2016) by legislators and constituents alike.

The second analysis showed that Senate candidates receive benefits in terms of both electoral outcomes and individual constituent vote choice as they increase both their social media presence and activity. Most notably, it was shown that for each 5-tweet

increase per month by a Senate candidate, the candidate experiences a one percentage point increase in vote share.

This research also considered the potential for benefits that senators and Senate candidates experience from social media investment to be gendered. Among sitting senators, female senators are slightly more likely than male senators to be recognized and approved of by their constituents regardless of senator social media engagement. However, among Senate candidates a distinct gender gap in the benefit received from social media activity is seen, with the gap decreasing as female candidates increase their Twitter activity when compared to their male counterparts.

It is important to recognize the limitations of the analyses conducted within this research. Beginning with problems of data collection and availability, there are important questions about whether the data collected is truly representative of the social media investment by senators and Senate candidates. This research uses Twitter activity as its predominant measure of social media activity. This measure is based on data collected from each individual Twitter profile as outlined in the Data and Methodology section of this paper. However, there are important measurement errors likely present within this research because of the non-permanency of social media behaviors. Individual actions on social media, as well as entire social media profiles, can be deleted. This creates a situation in which the certainty of the behaviors of these individuals on social media is fairly low. It is incredibly likely that the measure of average Twitter activity (the average number of tweets a senator or candidate posts per month) is not representative of all their Twitter activity, as they have possibly deleted several posts, particularly if they have been on the platform for several years. However, the current limitations of data availability about political uses of social media in the U.S. make the data collected for this research about senators and Senate candidates across several years the best data currently available to examine the questions considered in this project.

There are also limitations on the measurement of constituency behavior because of the nature of the CCES dataset. The questions asked in 2016 about social media use as well as specific political actions on social media asked respondents whether they had done the action within the last twenty-four hours. This small period of time specified in the questions likely truncates the sample of both social media users and respondent's levels of political social media engagement. Future research on this topic could be improved by conducting a survey meant to measure important factors of social media engagements by constituents, as well as their recognition and approval of their senators and their vote choice in elections.

Conclusions

This research has shown that while varied in magnitude, senators and Senate candidates experience benefits in terms of constituency recognition and approval and electoral outcomes from increased investment in social media. This research also showed that benefits senators and Senate candidates experience from social media investment are greater for men than women. These results have significant implications on both

the understanding of political engagement of constituents and the political strategies of legislators and candidates. Given that constituents are becoming increasingly more likely to use social media as their main form of news acquisition, understanding the way this news acquisition may involve direct engagement with their legislators and the impact this may have on political engagement and efficacy is crucial. It is also important for legislators and candidates to understand where they may best strategically invest in social media engagement to increase their constituency recognition and approval as well as their probability of electoral success.

Further work on the benefits legislators and candidates receive from social media investment should consider both the findings and the limitations of this research. Additional exploration of the social media activity of legislators and candidates should be conducted. Potential research could examine both the frequency of engagement on the part of the political figure and the engagement on the part of those within their networks towards the political figure. This would give substantially more information about the true social media activity of these individuals. It would also be beneficial to examine the content of social media engagement to understand potential differences in the benefit of social media investment based on the types of content legislators and candidates are engaging with on their social media platforms.

Given the shown dramatic increase in the importance of social media in the 2016 congressional elections when compared to these elections in both 2012 and 2014, it is clear that political uses of social media are becoming more influential. If social media continues to increase in importance for both legislators and candidates, it will be crucial to understand the benefits that they may receive from investing in social media as well as what types of social media investment are most likely to give these benefits.

NOTES

1. The CCES survey data from each year includes weighted variables to allow for a more nationally representative sample (Vavreck, 2008). Given that this research is looking specifically for differences between individuals within state-level constituencies rather than a nationally representative sample these weights are not used in this study.
2. The majority of Senate candidate's social media pages are archived within one week before Election Day in each election. For a small portion of candidates, the closest archive was slightly earlier with the furthest measurement being seventeen days prior to the election.
3. Preliminary analyses were conducted using social media activity on both Facebook and YouTube, all results showed no effect from activity on these platforms. This is also true for the models of Senate candidate social media activity.
4. Analyses were conducted for the 112th and 113th Congress as well. They showed similar results in terms of both statistical and substantive significance.
5. Analyses of the effect of Senate candidate social media activity and presence were conducted for the 2012 and 2014 congressional elections were conducted. These regressions showed predominately null results with no statistical or substantive significance.

APPENDIX

Table 3: Senator and Senate Candidate Social Media Presence by Year

	Senators			Senate Candidates		
	Facebook	Twitter	YouTube	Facebook	Twitter	YouTube
2012	82%	88%	91%	90%	97%	72%
2014	95%	97%	85%	93%	100%	83%
2016	97%	100%	100%	100%	98%	59%

Table 4: Social Media Data Collection by Platform

	Facebook	Twitter	YouTube
Presence	X	X	X
Years Present	X	X	X
Number of Followers	-	X	X
Number of Posts	-	X	-
Number of Views	-	-	X

Table 5: Logistic Regressions of the Effect of Senator Social Media Presence

VARIABLES	(1) Senator Name Recall	(2) Senator Party Identi- fication	(3) Senator Approval
Senator Social Media Presence	-0.209*** (0.0786)	-0.0776 (0.0730)	-0.0797 (0.0666)
Res. Social Media Use	-0.421*** (0.0229)	-0.311*** (0.0207)	-0.118*** (0.0203)
Res. Political Social Media Use	0.305*** (0.00716)	0.246*** (0.00623)	0.0812*** (0.00555)
Senator Republican	0.371*** (0.0191)	0.359*** (0.0170)	-0.282*** (0.0163)
Respondent Republican	-0.0696*** (0.0224)	-0.102*** (0.0193)	-0.556*** (0.0185)
Partisan Match	-	2.042*** (0.0243)	2.247*** (0.0171)
Senator Male	-0.235*** (0.0214)	-0.308*** (0.0188)	-0.170*** (0.0178)
Respondent Male	0.629*** (0.0202)	0.518*** (0.0179)	0.184*** (0.0169)
Gender Match	0.0447** (0.0200)	0.0677*** (0.0176)	0.0755*** (0.0166)
Respondent White	0.121*** (0.0192)	0.188*** (0.0173)	0.149*** (0.0175)
Respondent Age	-0.0407*** (0.000576)	-0.0418*** (0.000519)	-0.0129*** (0.000472)
Respondent Income	0.116*** (0.00292)	0.110*** (0.00260)	0.0313*** (0.00212)
State Fixed Effects?	YES	YES	YES
Constant	80.32*** (1.162)	81.79*** (1.047)	24.47*** (0.951)
Observations	73,292	101,298	101,298

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Regression of the Effect of Candidate Social Media Presence and Activity on Vote Share and Race Results

VARIABLES	(1) Vote Share	(2) Race Won	(3) Vote Share	(4) Race Won
Candidate Social Media Presence	15.24*** (3.041)	2.612*** (0.682)	-	-
Candidate Twitter Activity	-	-	0.228*** (0.0445)	0.0287*** (0.0110)
Candidate Male	0.945 (3.120)	-0.144 (0.762)	5.025 (3.501)	0.453 (0.702)
Candidate Republican	4.359 (3.222)	1.424** (0.672)	5.290 (3.382)	1.179** (0.586)
Candidate Incumbency	18.88*** (3.291)	-	20.33*** (3.430)	-
State Fixed Effects?	YES	YES	YES	YES
Constant	5.199 (7.829)	-7.365*** (1.965)	31.71*** (3.817)	-2.013** (0.805)
Observations	63	63	62	62
R-squared	0.352		0.260	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

APPENDIX 2

CCES Variables and Coding

Case ID (All Years)

coded: continuously

Abbreviation of State of Respondent (All Years)

coded: string

Birth Year of Respondent (All Years)

coded: continuously

Gender of Respondent (All Years)

coded: binary (1=Male, 2=Female)

Education of Respondent (All Years)

coded:

1= No High School

2= High School Graduate

3= Some College

4= Associates Degree

5= Bachelor's Degree

6=Post-Grad

8/9=Skipped/Not Asked

Race of Respondent (All Years)

coded:

1= White

2=Black

3=Hispanic

4=Asian

5=Native American

6=Mixed

7=Other

8=Middle Eastern

98/99=Skipped/Not Asked

Party of Respondent (All Years)

Coded: Democratic/Republican/Independent

Party Recall/Name Recall Senator 1 (All Years)

Coded:

1= Never Heard of

2=Republican

3=Democrat

4=Other/Independent

5=Not sure

8/9=Skipped/Not Asked

recoded: 0=Never Heard Of (1), 1= Any Choice (2,3,4,5)

Choices 8/9 omitted.

Party Recall/Name Recall Senator 2 (All Years)

Coded:

1= Never Heard of

2=Republican

3=Democrat

4=Other/Independent

5=Not sure

8/9=Skipped/Not Asked

recoded: 0=Never Heard Of (1), 1= Any Choice (2,3,4,5)

Choices 8/9 omitted.

Approve of Senator 1 (All Years)

Coded:

1= Strongly Approve

2= Approve

3=Disapprove

4=Strongly Disapprove

5=Never Heard Of

8/9=Skipped/Not Asked

Choices 5,8,&9 omitted.

Approve of Senator 2 (All Years)

Coded:

1= Never Heard of

2=Republican

 3=Disapprove

4=Strongly Disapprove

5=Never Heard Of

8/9=Skipped/Not Asked

Choices 5,8,&9 omitted.

Party of Senator 1 (All Years)

Coded: Democratic/Republican/Independent

Party of Senator 2 (All Years)

Coded: Democratic/Republican/Independent

Gender of Senator 1 (All Years)

Coded: binary (1=Male, 2=Female)

Gender of Senator 2 (All Years)

Coded: binary (1=Male, 2=Female)

2016 Media Use Questions

Each of the following questions asks respondents if they have engaged with a specific media platform in the last 24 hours.

Blog

Coded: binary (1=Yes, 2=No)

Television

Coded: binary (1=Yes, 2=No)

Newspaper

Coded: binary (1=Yes, 2=No)

Radio

Coded: binary (1=Yes, 2=No)

Social Media

Coded: binary (1=Yes, 2=No)

None

Coded: binary (1=Yes, 2=No)

All coding for the 2016 media use questions were recoded to a 0/1 binary with 1 = Yes and 0 = No.

2016 Political Social Media Engagement Questions

Each of the following questions asks respondents if they have engaged with politics on social media in various ways in the last 24 hours. These questions were only asked of individuals who identified they had engaged with social media in the last 24 hours in the media use questions.

Posted a story, photo, video or link about politics.

Coded:

1 = Yes

2 = No

9 = Not Asked

Posted a comment about politics.

Coded:

1 = Yes

2 = No

9 = Not Asked

Read a story or watched a video about politics.

Coded:

1 = Yes

2 = No

9 = Not Asked

Followed a political event.

Coded:

1 = Yes

2 = No

9 = Not Asked

Forwarded a story, photo, video or link about politics to friends.

Coded:

1 = Yes

2 = No

9 = Not Asked

All coding for the 2016 media use questions was used to create a measurement of political social media engagement based on a sum of the number of ways an individual behaved politically on social media.

REFERENCES

Alexander, Deborah and Kristi Andersen. 1993. "Gender as a Factor in the Attribution of Leadership Traits." *Political Research Quarterly* 46, no. 3 (September): 527–45.

Bennett, W. Lance. 2012. "The Personalization of Politics: Political Identity, Social Media, and Changing Patterns of Participation." *Annals of the American Academy of Political and Social Science* 644 (October): 20–39.

Carlisle, Juliet E. and Robert C. Patton. 2013. "Is Social Media Changing How We Understand Political Engagement? An Analysis of Facebook and the 2008 Presidential Election." *Political Research Quarterly* 66, no. 4 (December): 883–95.

DiGrazia, Joseph, Karissa McKelvey, Johan Bollen, and Fabio Rojas. 2013. "More Tweets, More Votes: Social Media as a Quantitative Indicator of Political Behavior." PLoS ONE 8, no. 11 (November): e79449.

Fenno Richard. 1975. "If, as Ralph Nader Says, Congress Is 'The Broken Branch,' How Come We Love Our Congressmen So Much?" In Congress in Change: Evolution and Reform, edited by Ornstein Norman, 277–87. New York: Praeger.

Fitch, Bradford and Kathy Goldschmidt. 2014. "Social Media Innovations in Congress Add to Transparency and Accountability." *Roll Call*, May 20.

Fridkin, Kim L. and Patrick J. Kenney. 2014. *The Changing Face of Representation: The Gender of*

U.S. Senators and Constituent Communications. Ann Arbor, MI: University of Michigan Press.

Gainous, Jason F. and Kevin M. Wagner. 2014. *Tweeting to Power: The Social Media Revolution in American Politics*. New York: Oxford University Press.

Gimenez, Alejandra T., Christopher F. Karpowitz, J. Quin Monson, and Jessica R. Preece. 2017. *Selection Effects and Self-Presentation: How the Double Bind Strangles Women's Representation.* Working Paper.

Hand, Laura C. and Brandon D. Ching. 2011. "'You Have One Friend Request': An Exploration of Power and Citizen Engagement in Local Governments' Use of Social Media." *Administrative Theory and Praxis* 33, no. 3 (December): 362–82.

Jacobs, Kristof and Neils Spierings. 2014. "Getting Personal? The Impact of Social Media on Preferential Voting." *Political Behavior* 36: 215–34.

Kahn, Kim F. 1992. "Does Being Male Help? An Investigation of the Effects of Candidate Gender and Campaign Coverage on Evaluations of U.S. Senate Candidates." *Journal of Politics* 54, no. 2 (May): 497–517.

Kahn, Kim F. 1994. "Does Gender Make a Difference? An Experimental Examination of Sex Stereotypes and Press Patterns in Statewide Campaigns." *American Journal of Political Science* 38, no. 1 (February): 162–95.

Kahne, Joseph and Ellen Middaugh. 2012. "Digital Media Shapes Youth Participation in Politics." *Phi Delta Kappan* 94, no. 3 (November): 52–56.

McGregor, Shannon C., Regina G. Lawrence, and Arielle Cardona. 2016. "Personalization, Gender, and Social Media: Gubernatorial Candidates' Social Media Strategies." *Information, Communication and Society* 20, no. 2 (March): 264–83.

Meeks, Lindsey. 2017. "Getting Personal: Effects of Twitter Personalization on Candidate Evaluations." *Politics and Gender* 13, no. 1 (March): 1–25.

Norval, Aletta. 2012. "Don't Talk Back!—The Subjective Conditions of Critical Public Debate." *Political Theory* 40, no. 6 (December): 802–10.

Obar, Johnathan A., Paul Zube, and Clifford Lampe. 2012. "Advocacy 2.0: An Analysis of How Advocacy Groups in the United States Perceive and Use Social Media as Tools for Facilitating Civil Engagement and Collective Action." *Journal of Information Policy* 2: 1–25.

Rosenwasser, Shirley M. and Norma G. Dean. 1989. "Gender Role and Political Office: Effects of Perceived Masculinity/Femininity of Candidate and Political Office." *Psychology of Women Quarterly* 13, no. 1 (March): 77–85.

Samuels, Bob. 2011. "Facebook, Twitter, YouTube—and Democracy." *Academe* 97, no. 4 (June–July): 32–34.

Shirkey, Clay. 2011. "The Political Power of Social Media: Technology, the Public Sphere, and Political Change." *Foreign Affairs* 90, no. 1 (January–February): 28–41.

Siegel, Hanna. 2012. "Three's Not a Crowd: Technology and the Political Shakeup." *Kennedy School Review* 12 (April): 62–65.

Sulkin, Tracy, Paul Testa, and Kaye Usry. 2015. "What Gets Rewarded? Legislative Activity and Constituency Approval." Political Research Quarterly 68, no. 4 (December): 690–702.

Wattal, Sunil, David Schuff, Munir Mandviwalla, and Christine B. Williams. 2010. "Web 2.0 and Politics: The 2008 U.S. Presidential Election and an E-politics Research Agenda." *MIS Quarterly* 34, no. 4 (December): 669–88.

Williamson, Andy. 2010. "The Internet and the British Election: Politicians Get Their Clicks." *The World Today* 66, no. 2 (February): 14–16.

Yiannakis, Diana. 1982. "House Members' Communication Styles: Newsletters and Press Releases." *Journal of Politics* 44, no. 4 (November): 1,049–071.

Inter-Female Hostility: Attractiveness and Femininity vs. Likeability

Savannah Henshaw, Lydia Estes, and Lauren Olsen

Introduction

Gender discrimination in business and politics is not a new phenomenon. Stories about the gender gap, feminist movements, and the glass ceiling have peppered the news since the women's movement began, and, until recently, people have generally blamed men for this inequality. However, research suggests that women actually play a significant role in perpetuating the gender gap due to a concept known as female hostility (Loya, Cowan, and Walters 2006). Female hostility is a destructive comparison mentality women sometimes exhibit when they feel inferior to another woman in some way. Examples of this mentality include: "If she is pretty, I'm not;" "If she is successful, I am a failure;" and "If she is smart, I am dumb." By comparing themselves to each other in this way, women inadvertently foster a mindset of competition that destroys relationships among them. While many factors could influence this mentality, we focused specifically on physical appearance and its effect on female hostility. Our experiment attempted to answer the following question: what is the relationship between a woman's physical appearance and her perceived likeability among other women? Specifically, we investigated whether women were more likely to ascribe negative traits to a female candidate if she is perceived as attractive. Using a simple survey, we studied the behavior of 139 women and found that attractiveness is negatively related to likeability among female voters at a 95 percent confidence level.

Female Hostility and Attractiveness in Politics

Our research question and design was based on an intriguing field of research within political science that investigates why women are underrepresented in Congress. Research in this field has demonstrated interesting relationships between attrac-

tiveness (femininity), gender, and success in politics. For example, despite multiple feminist movements and millions of women seeking representation in Congress, only 20 percent of federal legislators are female (Center for American Women and Politics 2017). Although many women advocate for female voices to represent them in Congress, recent polls show that U.S. women prefer men in positions of leadership (Litwan 2017; Ferber, Huber, and Spitze 1979). The obvious contradiction between what women say and how they vote has played a significant role in the underrepresentation of women in Congress.

For decades, research has shown that women are more hostile to each other than they are to men (Bleske-Recheck and Lighthill 2010; Chelser 2009; Gaitskill 2006; Haas and Gregory 2005; Jack 2009; Simmons 2002; Tulshyan 2012). In fact, from an early age, women begin demonstrating aggression and hostility almost exclusively to their female peers (Chesler 2009; Simmons 002). Women judge more attractive females to be less trustworthy; they threaten the resources and opportunities of their peers and create a lack of trust in other women (Gatskill 2006). As a result, women are much harsher critics of their female colleagues and, in many cases, they endorse gender-stereotypes more enthusiastically than their male counterparts (Ellemers, Van den Heuvel, de Gilder, Maas, and Bonvini 2004). This comparison and competition among females encourages women to succeed by elevating themselves above other women in the workplace and perpetuating gender inequalities among their inferiors. This pattern of separation and subjugation among professional women is known as the *queen bee syndrome* (Ibid.). It causes women to perceive assertiveness in other women as negative and augments feelings of distrust in female leadership (Haas and Gregory 2005; Mathiason 2010). This environment of competition can cause an attractive woman to decline in social status and popularity among her female friends and colleagues (Haas and Gregory 2005; Loya, Cowan, and Walters 2006; Simmons 2002). Though ample research has investigated the way successful women perceive and react to other women, little has been done to evaluate how those women perceive and react to their successful female colleagues. In this paper, we hope to address this gap.

In addition to the queen bee syndrome, the relationship between political success and attractiveness plays an important role in our research. In politics today, many citizens are poorly informed about candidates' qualifications or views. Unfortunately, this often leaves appearance as a primary factor for election success (Atkinson, Enos, and Hill 2009; King and Leigh 2009; Rosar, Klein, and Becker 2007; Martin 2014). Not only are attractive politicians more likely to be elected and supported by their constituents, but handsome and pretty politicians get away with unethical behavior while their less attractive counterparts are punished (Stockemer, Prain, and Moscardelli 2016). The voter population has repeatedly given their support to candidates who are well dressed, carefully coiffed, and genetically blessed. This phenomenon is not manifested exclusively in politics; in nearly every aspect of work, success is positively correlated with how well a person conforms to cultural beauty standards (Hamermesh 2011).

There is ample evidence that positively links hostility among women to attractiveness and femininity as well as to attractiveness and electoral success. Social scientists have thoroughly investigated the relationship between physical appearance and election results, but there have been insufficient studies about the influence of gender on this trend. To fill this void, we sought to combine previous scholarship by connecting female hostility and attractiveness to electoral success.

Attractiveness and Likeability Defined

We hypothesized that as a female candidate's attractiveness increased, her perceived likeability among other women would decrease. We planned to test our theory by manipulating the attractiveness of a female candidate and using a survey to measure her perceived likeability among female voters. In order to test this accurately, we needed to have clear definitions of our variables. Our primary independent variable was attractiveness, which we defined as the physical trait in an individual that is visually appealing to others. We operationally defined attractiveness by how well a person conforms to traditional beauty standards through the use of makeup. Beauty standards change constantly and vary by country and culture; however, the use of cosmetics has consistently been associated with higher levels of perceived "attractiveness, femininity, and sexiness" across borders and nations (Narang 2013). Research has suggested that the use of makeup positively reinforces sex-role stereotypes and is used regularly around the world as a method for women to appear and feel more attractive and feminine (Cox and Glick 1982; Miller and Cox 1982). These studies linking makeup and female attractiveness or femininity are abundant and reliable. In addition, using makeup allowed us to present the same political candidate in both treatments to avoid unnecessary bias in our results; because beauty is subjective, using different models as our political candidate would have created an enormous bias within our subject pool. Although makeup application is not the only method of measuring attractiveness and femininity, it was the best measure of the variable for our purposes.

In this experiment, we defined our dependent variable, likeability, as positive reactions to an individual. We measured this variable operationally by asking survey participants to rank the candidate's likeability using a Likert scale of one to ten. We also tested phrases that national public opinion polls typically associate with positive or negative attributes of politicians, such as informed, trustworthy, and qualified (Botti 2014). Using these phrases, we focused the participants' attention on what they liked and disliked about the woman specifically as a politician. Each participant rated the candidate on a scale of one to ten for different positive and negative attributes. We then looked at how likeable they rated her, which we used as our dependent variable to perform our statistical analysis.

Experimental Design

Our experimental design included a treatment group, a placebo group, and a control group. Based on common statistical norms, we needed at least thirty participants

for each group in order to achieve statistical power, with ninety participants overall. To find these participants, we used a script to personally invite our Facebook friends to take a survey, which they could complete at their convenience (see Appendix—A). We invited people from various socioeconomic classes to participate to make the sample as representative as possible. Using this method, we gathered data from about 139 women ages eighteen and older. We acknowledge that the group may not have been an accurate representation of the general public, but it did give us a basic idea of how plausible our theory was.

At the beginning of the survey, we asked participants to give their consent to be part of our research and to answer a series of demographic questions (see Appendix—B). Though we were explicitly studying the effects of attractiveness or femininity on our outcome measure, we were also interested in how differences among several subgroups in our sample affected our results. The subgroups we evaluated included age, employment status, and political affiliation. The distribution of these subgroups is shown in Table 1. By analyzing this demographic information, we measured the influence of these subgroups on our results.

Table 1

Distribution of Demographics							
Number of participants in each demographic							
Age		**Employment Status**		**Household Income**		**Political Affiliation**	
18-28	81	Employed for Wages	48	$0–29,000	18	Republican	75
29-38	7	Self-Employed	11	$30,000–50,000	17	Democratic	11
39-48	16	Out of Work	2	$60,000–80,000	14	Independent	25
49-58	14	Homemaker	21	$80,000–100,000	17	Green	0
59-68	4	Student	42	$100,000–200,000	35	Other	1
69-78	4	Retired	2	$200,000+	27		
79-88	2	Unable to Work	0				

Following the demographic questions, each participant listened to a twenty-second clip of a speech by a fictional female politician running for Senate (see Appendix—C). To avoid partisan bias, we chose a mock stump speech written by two former professional speechwriters—one Republican and one Democrat (Swaim and Nussbaum 2016). We also chose a section of the speech that focused on neutral topics, such as improving America, without addressing controversial or partisan issues.

All participants listened to the same speech, but each group saw a different screen while doing so. The placebo group saw a blank page. The control group saw a picture of a female candidate with unstyled hair and no makeup (see Appendix—E). The treatment group saw a picture of the same woman, but with well-kept hair and professional makeup (see Appendix—F). As we explained in the definitions section, we selected makeup as our treatment because females without cosmetics are typically viewed as less attractive or less feminine. For our placebo, the blank screen

eliminated any bias based on levels of attractiveness and femininity. Using the blank screen forced the participants to focus exclusively on the audio clip from the politician rather than on any visual cues. In this way, we determined how the candidate's physical appearance, rather than the content of the speech, affected participants' opinions.

After listening to the speech, participants ranked how likeable the candidate was using a simple Likert scale from one to ten. We used their responses to this question as our dependent variable to gauge how the women participating in the survey viewed the potential candidate. Finally, the participants answered several open-ended, qualitative questions about the candidate. These questions gave us more insight about how participants felt toward the candidate and why (see Appendix—D).

Causes and Effects of Female Hostility

Based on this experimental design, we hypothesized that as the female candidate's attractiveness increased, her perceived likeability among other women would decrease. This hypothesis was based on the evidence that female competition causes women to feel threatened by attractive females (Haas and Gregory 2005). As previously mentioned, this pattern is referred to as the queen bee syndrome. If a woman perceives another woman as attractive and successful, she is more prone to judge that female candidate according to traditional gender stereotypes (Ellemers, Van den Heuvel, de Gilder, Maas, and Bonvini 2004). For example, women will perceive attractive women as unintelligent, unambitious, or unassertive (Huddy and Terkilsden 2016; Ellemers, Van den Heuvel, de Gilder, Maass, and Bonvini 2004). Once women have attributed these stereotypes to the candidate, they subconsciously judge her as less qualified and will not vote for her.

Statistical Analysis of Likeability

To reliably assess this hypothesis, we ran several statistical tests in order to determine the impact of each of the treatment conditions as well as other variables on the candidate's likeability. Our survey included questions regarding three other subgroups in our sample that could have influenced the perceived likeability of the candidate. The subgroups we measured were age, employment status, and political affiliation. We asked about age to determine if women in various age groups responded differently to successful women. For example, older participants closer to the candidate's age might experience stronger levels of female hostility, because they feel more threatened by her than younger participants do. We also investigated how political affiliation affected participants' opinions by asking participants to indicate which political party they identified with. Different political parties have different perceptions about feminism and gender equality, which could influence participants' responses. Finally, we evaluated differences in socioeconomic status by asking about employment status. Participants who are currently working might feel a more competitive response toward the candidate than a stay-at-home mom would, because they are competing for the

same opportunities and resources. However, in our analysis of these subgroups, we found no statistically significant differences between these groups. This suggests there is no relationship between age, employment, or political affiliation and likeability (See Appendix—G).

After evaluating the subgroups, we began our analysis by using a t-test to compare the makeup treatment to the no makeup control. The results are included in Table 2. The difference between the mean likeability score for both groups was .80, which had a p-value of .036. Thus, we can confirm at the 95 percent confidence level that there was a statistically significant difference between the two groups. Additionally, with a p-value of .018, we can confirm at the 95 percent confidence level that there was not only a difference, but the mean for the makeup group was statistically lower than the mean of the no makeup group. Based on this evidence, we can reject the null hypothesis that there was no difference between the two groups.

Table 2

Makeup vs. No-Makeup T-test		
Ho: Likability for Makeup Treatment (M)=Likability for No Makeup Control (NM)		
Treatment	Makeup	No Makeup	Difference
Observations	33	37	-
Mean	7.09	7.89	0.80
Standard Error	0.26	0.27	0.37
		T-statistic	2.14**
		P-value (Ha: M ≠ NM)	.036**
		P-value(Ha: M < NM)	.018**

To further investigate this relationship, we also compared the makeup group to the no makeup and placebo groups combined. The results of this t-test are included in Table 3. As it shows, there is a surprising difference between this test and the last test. When comparing the makeup group to the no makeup group, we found statistically significant evidence that the groups were different. However, when we added the placebo group, the difference was no longer significant. This suggests that the placebo group was more similar than the no makeup group was to the makeup group. Though we cannot confirm any causal reasons behind this discovery, further research could investigate them. It is possible that the women listening to the speech with no picture were imagining the speaker's physical appearance as they listened and expected that she would be wearing makeup. Additionally, perhaps not wearing makeup actually makes the candidate more likeable, rather than our original theory that wearing makeup makes the candidate less likeable. None of this speculation can be confirmed through our experiment, but it does provide opportunities for further research.

Table 3

| Makeup vs. Other Groups T-test | | | |
| Ho: Likability for Makeup Treatment (M)=Likability for Other Groups (O) | | | |
Treatment	Makeup	Other Groups	Difference
Observations	33	70	-
Mean	7.09	7.42	0.34
Standard Error	0.26	0.21	0.36
T-statistic			0.95
P-value (Ha: M $\neq$ O)			0.34
P-value(Ha: M $<$ O)			0.17

The information we have gathered from the two t-tests and the regression of our subgroups have significantly added to the current literature on female hostility. The candidate was statistically less likeable when she was wearing makeup, meaning there is a significant relationship between attractiveness and likeability. We can confidently reject the null hypothesis that the candidate was equally likeable whether she was wearing makeup or not. Though the t-test that included the placebo group was not statistically different from the makeup group, future research with a larger sample size and improved methods of testing could achieve greater statistical power to investigate this relationship more thoroughly.

Qualitative Analysis of Likeability

Our qualitative data, like our quantitative data, demonstrated that there is a correlation between attractiveness, femininity, and likeability. Analyzing participants' responses to our survey's open-ended questions helped us discern the rationale behind the participants' perceptions. While these responses are individual and do not represent the respondents in general, they are valuable, because they give insight into our quantitative results.

To evaluate these mechanisms, we asked participants to give advice to the candidate and to identify what they liked and disliked about her. We avoided asking questions regarding the candidate's political views in order to collect information regarding the participants' overall perception of the candidate. The responses to these questions helped us identify the relationship between hostility, femininity, and attractiveness by illustrating the factors that affected likeability. They also indicated whether participants subconsciously stereotyped the candidate with makeup according to the traditional gender stereotypes as discussed earlier.

To accurately distinguish the differences in the qualitative data, we started by identifying comments that appeared across all groups and exempted them from our analysis. Most of the responses generated by the placebo group focused on the disin-

genuous nature of the candidate's speech. In an attempt to avoid party biases, we chose a speech that was vague and nonpartisan. Due to the general nature of the speech, participants in the control group thought the candidate sounded "robotic and fake" and suggested that she include more specifics. Comparing these comments to the treatment and control groups, we found that they appeared consistently across all groups. Similarly, when asked what they liked about the candidate, participants in all groups mentioned her "friendly" nature and "kind" demeanor. These comments illustrated several limitations of our experimental design. If we were to repeat this study in the future, we would alter the instructions, speech, and questions to guide the survey participants to respond about the candidate herself, not about her political opinions. Because these remarks appeared frequently in all groups, we assumed these comments originated from other influences and not from the effects of the treatment. Accordingly, we ignored comments that appeared in responses across all of the groups and those regarding the content of the speech. We also recognize that one person's comment does not speak for the entire group. By focusing exclusively on the differences between the groups, we were able to identify comments that provide insight pertaining to specific individuals' experiences within each group.

We began our qualitative analysis by first evaluating responses that specifically mentioned physical appearance. As explained earlier, we hypothesized that attractive candidates would be more likely to be perceived as untrustworthy or incompetent due to intrasexual competition among females. We anticipated this competition would cause participants to feel hostility toward the candidate and attribute these negative feelings to the candidate's appearance. Therefore, if participants cited physical appearance as an influential factor in their opinion of the candidate, this would indicate a causal relationship between appearance and likeability. The most interesting comment that mentioned physical appearance was one woman's response about the candidate with makeup: "She doesn't really look the part of a politician. I'd probably take her more seriously if she was brunette." This comment strongly suggested that the participant viewed the candidate as less qualified because of her appearance. By implying that she judged the candidate more stereotypically because of her hair color, she illustrated the competitive mentality that results from female hostility.

Another response that mentioned physical appearance shed additional light on the connection between attractiveness, femininity, and likeability. The participant said, "You seem nice, but perhaps some light makeup will make you look more polished and less granola." Initially, this comment seemed to contradict our hypothesis—this woman noticed the candidate's appearance and suggested she change it, just like the woman who commented about the candidate's hair color. On the surface, both participants appeared to have the same reaction. However, as we compared the two responses, we noticed an intriguing contradiction. Both comments suggested the candidate change her appearance in some way. However, in the first comment, the participant implied that the can-

didate's appearance negatively impacted her opinion of the candidate's capability. The second comment, however, only suggested that she change her appearance to look more polished, not to seem more capable. Though the reason for the participant's response is unclear, it is interesting that she commented on physical appearance rather than the candidate's other characteristics.

Though the first two responses supported our hypothesis, another woman who saw the woman with makeup gave a different perspective. She said, "[The candidate] seemed inviting to talk to. I think her appearance in the picture helped illustrate that." This was an intriguing response because it demonstrated another possible aspect of female hostility that we did not anticipate: female hostility could affect women in varying degrees. A woman's susceptibility to female hostility could potentially be weaker or stronger depending on her confidence level, life circumstances, or personal goals. Further research could examine the link between a woman's personal confidence and how strongly she experiences female hostility.

The qualitative responses we analyzed demonstrated important insights concerning the way individual women experience female hostility. They also provoked questions we would like to investigate further regarding why different women experience hostility differently. Though these results are not universally applicable, they do provide important insights to our quantitative findings about the relationship between attractiveness and female hostility.

Limitations

Though we have carefully designed our experiment and analyzed the results, we do recognize that all research inherently has limitations. In our case, we experienced limitations due to the need for ambiguity about the subject of our research. We have attempted to mitigate potential errors in our research by identifying and addressing these limitations individually.

One potential limitation was that likeability depends on more than simple levels of attractiveness and femininity. Some participants could have rated the candidate based on factors besides appearance, such as political biases, tone of voice, or other influences controlled for through our randomization. Our experiment sought to measure implicit biases; as a result, participants could not know the intent of the research, because it would have influenced their responses. Consequently, our questions elicited some vague or irrelevant comments in our qualitative data. We expected participants to discuss the candidate's character and ascribe negative or positive attributes due to judgments about her physical appearance. However, some participants inferred that the survey intended to measure political affiliation and gave answers regarding the content of the speech instead of the candidate herself. If we could conduct this experiment again, we would explicitly state that we were not measuring political values to discourage participants from responding about the speech's content. This would encourage participants to base their responses on the candidate's likeability and personality instead of on her political affiliation or policies.

Similarly, the intrinsic limits of quantitative and qualitative data posed another potential problem. While quantitative data provides statistical information, it does not explain the causal mechanisms behind these numbers. Conversely, qualitative responses do reveal causal mechanisms behind individual participants' responses but do not necessarily represent the population as a whole. In spite of these limitations, using both types of responses helped us understand the relationship between attractiveness, femininity, and likeability. The quantitative data helped us evaluate the direction and magnitude of this relationship, and the qualitative data helped us identify the causal logic behind it. Our qualitative data allows us to infer a relationship, but we cannot definitively confirm it for the entire sample.

Finally, one of the greatest limitations of our research is our sample. Because of limited resources, we only administered the survey through our personal social networks. Our acquaintances are possibly skewed toward our community's political leanings and are not representative of women from a normally distributed socioeconomic and geographic population. Thus, if we were to conduct this experiment again, we would survey a larger, more random, and normally distributed sample. Although many limitations restricted our design, our experiment helped us obtain important data regarding female hostility.

Conclusion

In spite of our limitations, we found that our quantitative data reflected a statistically significant relationship between attractiveness and femininity and likeability. We found statistically significant evidence that the candidate without makeup was more likeable than the candidate with, and we rejected the null hypothesis. In addition, our qualitative responses told an important story regarding the relationship between female hostility, attractiveness and femininity, and electoral success. For example, the participants in the makeup group evaluated the candidate much more negatively and stereotypically than those in the control group. Various comments demonstrated the role of attractiveness and femininity in this relationship.

Though our results were compelling and informative, we have only begun to examine the full implications of female hostility. The shocking void in current research must be filled for women to be represented equally in politics. Other factors, such as confidence and intelligence, could influence female competition. Further questions we could investigate include the following: Would women who are competing directly against each other dislike their more attractive competitors? Does the reaction of women to other attractive women vary from culture to culture? How does personal confidence affect variation in female hostility?

Though more research is needed, our findings have certainly added to the current literature on female hostility and have given us insight into what more we can study. Based on our compelling results, efforts to balance the ratio of men to women in Congress need to change and refocus. Thus far, most of these movements have been focused on influencing male voters, but if a woman's attractiveness and femininity

truly does decrease the likelihood that other women will vote for her, female hostility may be the new key to shattering the glass ceiling.

APPENDIX

A. Script for inviting people to participate in the study

Hi ___,

I hope you're doing well. For one of my classes at BYU, we are researching elections and voter perceptions of candidates and we have to gather data using a survey. Would you mind taking a 5-minute survey to help me gather the data we need? I would appreciate your help so much.

Thank you in advance,

___

B. Consent

This study is being conducted by Lydia Estes, Lauren Olsen, and Savannah Henshaw as part of a political science course. We are studying what people look for in their political candidates. If you choose to proceed with this survey, we will introduce you to a potential candidate to gauge interest and future support for them in the upcoming state Senate elections. You will listen to a 30-second message and answer a few questions about how well they come across to the public. The survey should take no more than five minutes of your time. All information will be kept confidential and anonymous. Participation is entirely voluntary. You may refuse to participate altogether or withdraw from the survey at any time. If you give your consent, click next. We appreciate your time!

C. Script for Speech

You're not here for a lot of political rhetoric. You didn't show up to hear a lot of talk about how things could be different and better than what they are. You're here because you're done with politics. You're here because the promise of American democracy—the promise of opportunity for all— remains unfulfilled for too many.

If you paid any attention to this election, or the last one, or the one before that, you can be forgiven for thinking that our problems are insurmountable, that there's just nothing we agree on enough to get it done. But that's not true.

There's more that unites us than divides us. And while elections focus on the divide, I want to focus on what we share. Because—yes—there's sacred ground. But for each of us, I believe, there's also common ground. And that's the ground we need to cultivate.

We know we can make progress. But for us to do something, I need you to do something. I need you to vote. I need you to get your friends and neighbors to vote. I ask you to stand with me. Join me. And together we'll build the country we know we can be.

Because remember: This election isn't about any one candidate. It's about where we're headed as a nation, and I believe we're headed toward economic freedom, self-governance, strength abroad and prosperity at home. Thank you.

D. Survey Questions

1. What is your age?

   a. 18–28   b. 29–38   c. 39–48   d. 49–58   e. 59–68   f. 69–78   g. 79–88

2. What is your annual household income?

   *College students married and unmarried signify your parents' household income

   a. $0–29,000   b. $30,000–50,000   c. $60,000–80,000   d. $80,000–100,000

   e. $100,000–$200,000   f. $200,000+   g. prefer not to answer

3. Are you currently

   a. Employed for wages   b. Self-employed   c. Out of work and looking for work

   d. Out of work but not curently looking for work   e. A homemaker   f. A student

   g. Retired   h. Unable to work

4. What Political Party do you most identify with?

   a. Republican   b. Democratic   c. Independent   d. Green   e. Other   f. None

5. The candidate speaking below is running for Senator in your state. Please click the link below to hear a portion from her speech. In answering the questions that follow, please assume the candidate agrees with your political views and shares the same values.

   [Random assignment of picture one, picture two, or no picture]

   [audio clip]

6. Select how strongly you agree with each statement. *(on a scale of 1–10)*

   This person is friendly.

   This person is likeable.

   This person is warm.

   This person is approachable.

   I would ask this person for advice or for help.

   I would like this person as my representative.

   I would like to be friends with this person.

   I would trust this person.

   This person is similar to me.

   This person is knowledgeable.

   This person is entitled.

   This person is bossy.

   This person is annoying.

   This person seems uninformed about important issues.

   This person is arrogant.

   This person appears indecisive.

   This person appears to be easily swayed.

   This person presented their message poorly.

   This person is a poor leader.

   This person is unqualified to run for office

7. What three words best describe the potential candidate?

8. What advice do you have for the candidate if they were to run for office?

9. What political party do you think this candidate belongs to? Why?

10. What do you like about the candidate?

11. What do you dislike about the candidate?

E. Photo One	F. Photo Two

G. Subgroup Regression

Table 4

Effect of Education Level on Gun Control Support *Dependent Variable: Gun Control Support*	
VARIABLES	
Age 18 to 28	-0.162
	(0.930)
Age 29 to 38	-0.532
	(1.003)
Age 39 to 48	-2.525***
	(0.912)
Age 49 to 58	-0.978
	(1.216)
Age 69 to 78	-3.405**
	(1.282)
Homemaker	0.384
	(0.623)
Student	-0.738
	(0.447)
Retired	2.271**
	(1.123)
Republican	0.607
	(0.402)
Democratic	-0.965*
	(0.498)
Makeup Treatment	-0.528
	(0.388)
Constant	8.134***
	(0.906)
Observations	70
R-squared	0.336

REFERENCES

Atkinson, Matthew D., Ryan D. Enos, and Seth J. Hill. 2009. "Candidate Faces and Election Outcomes: Is the Face-Vote Correlation Caused by Candidate Selection?" *Quarterly Journal of Political Science* 4, no. 3 (June): 229–49.

Botti, David. 2014. "What Do Americans Want Most in a President?" (video). BBC News. Accessed November 10, 2017. www.bbc.com/news/av/magazine-27599064/what-traits-do-americans-want-most-in-a-president.

Bleske-Rechek, April and Melissa Lighthall. 2010. "Attractiveness and Rivalry in Women's Friendships with Women." *Human Nature* 21, no. 1 (March): 82–97.

Center for American Women and Politics. "Women in U.S. Congress 2017." Rutgers Eagleton Institute of Politics. Accessed October 16, 2017. www.cawp.rutgers.edu/women-us-congress-2017.

Chesler, Phyllis. 2009. *Woman's Inhumanity to Woman*. Chicago: Lawrence Hill Books.

Cox, Cathryn L. and William H. Glick. 1982. "Resume Evaluations and Cosmetics Use: When More is Not Better." *Sex Roles* 14, no. 1 (January): 51–58.

Derks, Belle, Colette Van Laar, and Naomi Ellemers. 2016. "The Queen Bee Phenomenon: Why Leaders Distance Themselves from Junior Women." *Leadership Quarterly* 27, no. 3 (June): 456–69.

Dion, Karen, Ellen Berscheid, and Elaine Walster. 1972. "What is Beautiful is Good." *Journal of Personality and Social Psychology* 24, no. 3 (December): 285–90.

Ellemers, Naomi, Henriette Van den Heuvel, Dick de Gilder, Anne Maass, and Alessandra Bonvini. 2004. "The Underrepresentation of Women in Science: Differential Commitment or the Queen Bee Syndrome?" *British Psychology Society* 43, no. 3 (September): 315–38.

Ely, Robin J. 1994. "The Effects of Organizational Demographics and Social Identity on Relationships among Professional Women." *Administrative Science Quarterly* 39, no. 2 (June): 203–08.

Ferber, Marianne, Joan Huber, and Glenna Spitze. 1979. "Preference for Men as Bosses and Professionals." *Social Forces* 58, no. 2 (December): 466–76.

Gaitskill, Mary. 2006. "Don't Hate Them Because They're Beautiful." *O, The Oprah Magazine*, April. Accessed October 16, 2017. http://eds.a.ebscohost.com/eds/detail/detail?vid=0&s id=ef1a3cc9-e563-4807-bb86-975d515fe6b1%40sessionmgr4009&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=20280180&db=tth.

Haas, Anne and Stanford W. Gregory, Jr. 2005. "The Impact of Physical Attractiveness on Women's Social Status and Interactional Power." *Sociological Forum* 20, no. 3 (September): 449–71.

Hammermesh, Daniel S. 2011. *Beauty Pays: Why Attractive People Are More Successful.* Princeton: Princeton University Press.

Huddy, Leonie and Nayda Terkilsden. 2016. "The Consequences of Gender Stereotypes for Women Candidates at Different Levels and Types of Office." *Political Research Quarterly* 46, no. 3 (July): 503–25.

Jack, Dana Crowley. 2009. *Behind the Mask: Destruction and Creativity in Women's Aggression.* Boston: Harvard University Press.

King, Amy and Andrew Leigh. 2009. "Beautiful Politicians." *Kyklos* 62, no. 4 (November): 579–93.

Litwan, Anne. 2017. "Six Reasons Why We Need More Women in Government." Anne Litwan and Associates. Last modified January 9. Accessed October 18, 2017. www.annelitwin.com/six-reasons-why-we-need-more-women-in-government/.

Loya, Bianca N., Gloria Cowan, and Christine Walters. 2006. "The Role of Social Comparison and Body Consciousness in Women's Hostility toward Women." *Sex Roles* 54, no. 7/8 (October): 575–83.

Martin, Annie. 2014. "Rob Lowe Talks Being 'Pretty' and Politics." UPI Entertainment. Last modified April 7. Accessed October 16, 2017. http://eds.a.ebscohost.com.erl.lib.byu.edu/

eds/detail/detail?vid=0&sid=5df20296-0fa4-4308-980a-4eb2ea08cb5d%40sessionmgr4010&bdata=JnNpdGU9ZWRzLWxpdmUmc2NvcGU9c2l0ZQ%3d%3d#AN=B92Y1601054818&db=n5h.

Miller, Lyn C. and Cathryn L. Cox. 1982. "For Appearance's Sake: Public Self-Consciousness and Makeup Use." *Personality and Social Psychology Bulletin* 8, no. 4 (December) 748–51.

Narang, Devina. 2013. "The Psychological Factors that Affect Makeup Usage and the Perception of Makeup in Different Situations." Paper presented at the National Conference Undergraduate Research 2013, La Crosse, WI, April 10–13. Accessed November 11, 2017. www.ncurproceedings.org/ojs/index.php/NCUR2013/article/view/666.

Reysen, Stephen. 2005. "Construction of a New Scale: The Reysen Likeability Scale." *Social Behavior and Personality: An International Journal* 33, no. 2 (January): 201–08.

Rosar, Ulrich, Markus Klein, and Tilo Beckers. 2007. "The Frog Pond Beauty Contest: Physical Attractiveness and Electoral Success of the Constituency Candidates of the North Rhine-Westphalia State Election of 2005." *European Journal of Political Research* 47, no. 1 (January): 64–79.

Sanders, Linley. 2017. "What Happened? In Interview, Hillary Clinton Gives Sexist Argument for Why She Lost the 2016 Election." *Newsweek*, September 13. Accessed October 18, 2017. www.newsweek.com/hillary-clinton- sexist-losing- 2016-election- what-happened-book-664295.

Simmons, Rachel. 2002. *Odd Girl Out: The Hidden Culture of Aggression in Girls*. New York: First Mariner Books.

Stockemer, Daniel, Rodrigo Praino, and Vin Moscardelli. 2016. "The Good, the Bad and the Ugly: Do Attractive Politicians Get a 'Break' When They Are Involved in Scandals?" Lecture, ECPR General Conference Charles University, Prague, September 7. Accessed October 16, 2017. https://ecpr.eu/Events/PaperDetails.aspx?PaperID=30450&EventID=95.

Swaim, Barton and Jeff Nussbaum. 2016. "The Perfect Presidential Stump Speech." FiveThirtyEight. Last modified November 3. Accessed October 23, 2017. https://projects.fivethirtyeight.com/perfect-stump-speech/.

Tulshyan, Ruchika. 2012. "Why Women Are the Worst Kind of Bullies." *Forbes Magazine*, April 30. Accessed October 9, 2017. www.forbes.com/sites/worldviews/2012/04/30/why-women-are-the-worst-kind-of-bullies/#d5e04ae36b7d.

Who is My Neighbor? Identifying Epistemic Peers among Polarized Communities

Alex Hoagland

While philosophers have extensively discussed the problem of disagreement among epistemic peers, they have generally set aside the question of identifying these same peers. However, especially in the context of questions of politics and religion, the labeling of others' opinions as important or meaningless is particularly relevant. In such a situation of disagreement, where the correct[1] opinion is not immediately apparent yet has significant moral weight, whose opinions are worth hearing and under what conditions can we exclude other opinions?

Before beginning a thorough exploration of this topic, I will present some terminology used frequently by philosophers in this conversation.[2] A person is considered your *epistemic peer* in a subject or issue if, loosely speaking, that person is "as good as you at evaluating claims [related to the issue]" (Elga 2006). That is, in performing a complicated mathematical calculation, I would count as an epistemic peer someone who 1) had equivalent mathematical training as I have, 2) had about equivalent experience performing this type of calculation as I have, and 3) was about as likely as I was to make mistakes in the course of the calculation. My peer is someone whom, once I set aside my own reasoning, I think is just as likely as I am to get the "right" answer to a certain question. A person who is more likely to solve the problem correctly than I am would be my *epistemic superior*, while one who is less likely to do so is my *epistemic inferior*.

The classification of a person as an epistemic peer (or superior or inferior) is issue-specific. While I may be inclined, for example, to see a colleague as an epistemic peer in mathematical questions, this labeling does not necessarily extend to any other issue. I would not be under any obligation to classify my graduate advisor as an epistemic superior in politics or religion, even while I may consider her my superior in our field

of research. It is therefore possible—even likely—for a person to fill the roles of epistemic inferior, peer, and superior in relation to another person across a continuum of issues.

The question of epistemic disagreement, then, is as follows: given that I hold an opinion *p* about an issue and subsequently discover that someone I consider an epistemic peer holds the opinion ~*p*, how should I alter 1) my belief or 2) my confidence in that belief? There are several hypotheses that attempt to answer this question. For simplicity, I will discuss the two main camps. Those ascribing to the Equal Weight View (Ibid.) assert that you should give your peer's conclusion equal weight as your own—that is, "you should think that the two of you are equally likely to be correct" (Ibid., p. 13). On the other hand, those who deny this view assert that there are some instances in which it is appropriate, even upon discovering that your peer has adopted the opinion ~*p*, to "stick to your guns" and continue to maintain your confidence in your original supposition of *p* (Schoenfield 2014, p. 3).

Neither of these views, however, gives any indication regarding who qualifies as an epistemic peer. Therefore, I will take a few steps back in order to answer this question, and then see what this implies about disagreement. This is particularly relevant in issues that are non-mechanistic. Even when performing a complicated calculation, the conditions for considering another person my epistemic peer seem relatively obvious. On the other hand, when discussing the existence of God among a group, there are less well-defined conditions. In fact, Elga discusses a case similar to this in his exposition of the Equal Weights View:

> Consider Ann and Beth, two friends who stand at opposite ends of the political spectrum. Consider the claim that abortion is morally permissible. Does Ann consider Beth a peer with respect to this claim? That is: setting aside her own reasoning about the abortion claim, and Beth's contrary view about it, does Ann think Beth would be just as likely as her to get things right? (2006, p. 21).

Elga asserts that the answer to this question is "no," precisely because "by Ann's lights, Beth has reached wrong conclusions about most closely related questions," such as "whether human beings have souls, whether it is permissible to withhold treatment from terminally ill infants, and whether rights figure prominently in a correct ethical theory" (Ibid., pp. 21–22). Because the two differ on issues that are *related* to the abortion claim, but somehow separate from the claim itself, each is unlikely to consider the other as her peer. Elga follows this with a moral: "with respect to many controversial issues, the associates who one counts as peers tend to have views that are similar to one's own" (Ibid., p. 23).

Therefore, when looking specifically at what Elga calls "controversial issues"—but which I will refer to as *non-mechanistic issues*[3]—one potential hypothesis about peer classification is Elga's claim that we have an obligation to consider someone an epistemic peer on an issue only if they share our opinions on it. I will refer to this hypothesis as the *community hypothesis*, but I will claim that this hypothesis is problematic, especially among political communities.

The Problems of Polarization and Independence

Before explaining the theory's flaws, let us assume it is true. In a world where epistemic hierarchies are created by this thesis, disagreement among epistemic peers is vacuously avoided (except in mechanistic cases)—I will never disagree with someone I consider my epistemic peer. My opinion on an issue will be altered only by interactions with epistemic superiors who hold opinions different from my own,[4] in which case, I will adopt theirs. Recall that an epistemic superior is more likely to arrive at the correct conclusion than you are, this is a desirable result.

However, there is inconsistency in assuming the existence of epistemic superiors with competing opinions in a system that excludes them from being your peer. Under Elga's view, Ann was not Beth's epistemic peer, precisely because Beth could judge—on the basis of fringe issues—that Ann had frequently come to the "wrong" conclusion. Beth would thus consider Ann her epistemic inferior, rather than the other way around.

Under the community hypothesis, then, an agent either holds the same views as you do or is your epistemic inferior, in which case you have no reason to regard their opinion as even remotely valuable. With no reason to bend on any non-mechanistic issues, disagreement then amounts to nothing except angry Facebook posts. This seems to undermine an important function of disagreement, as implied in both Condorcet's Jury Theorem and the Median Voter Theorem: that of generating consensus. The community hypothesis allows individuals to reinforce their own convictions with self-selected epistemic peers while simultaneously removing themselves from their opponents, eliminating the need for compromise altogether. This is the problem of polarization.

Other critiques of this hypothesis seem compelling.[5] Consider one central problem in the abortion scenario: When Ann judges Beth based on her opinions on closely related issues, she is judging Beth inferior on claims that 1) are not fully separable from the issue at hand and 2) are not verifiable; therefore, they are not evidence of Beth's incompetence or irrationality. When considering non-mechanistic issues, such as politics, people tend to form opinions consistent across issues in order to minimize cognitive dissonance.[6] Therefore, judging someone an epistemic peer (which asserts that *independent of your reasoning*, a colleague is as likely to come to the right conclusion as you are) requires you to set aside these ideologies and examine only traits, such as the ability to think critically and rationally. By asserting that these "clusters of controversy" determine an epistemic hierarchy, Elga asserts that disagreement implies inferiority (Ibid, p. 23).

One potential rejection of the community hypothesis is that we have no special obligation to consider anyone an epistemic peer—call this the *stand-alone hypothesis*. Yet this too is problematic. After all, it does seem right to conclude that a person who shares your view on an issue and has no external qualifications for superiority or inferiority should be considered your epistemic peer. Under these conditions, considering her a superior would be to invalidate the independence of your opinion forming, while considering her an inferior would be to invalidate hers. The stand-alone hypothesis thus seems to

violate some core of individual opinion forming by not calling a duck a duck. Call this the problem of independence—without some obligation to view members of a like-minded community peers, these communities devolve into bandwagons.

Peer Classification Outside of Like-Minded Groups: The Conversation Hypothesis

As an alternative to both the community and stand-alone hypotheses, I propose a middle ground: While we ought to consider those sharing our views as epistemic peers,[7] there additionally exist circumstances under which we are obligated to consider those *outside* of our like-minded communities as epistemic peers. My argument for this claim— call it the *conversation hypothesis*—is summarized here:

1. Agent *A* considers agent *B* her epistemic inferior if and only if *A* is more likely to reach the correct solution to *I* than is *B*. A similar definition holds for epistemic superior.
2. With regard to an issue *I,* any agent must consider other agents either her epistemic inferior, superior, or peer.
3. An issue *I* is considered *non-mechanistic* or *controversial* if a third party possessing the sum total of mankind's knowledge could not assess the correct response to *I*.
4. *A* is more likely to reach the correct answer on an issue *I* if and only if *A* has better training, experience, evidence, or history[8] than *B*.
5. When *I* is non-mechanistic, training and experience are irrelevant in determining the likelihood that an agent will reach a correct conclusion. Similarly, a non-mechanistic *I* prohibits there being any history of correctness.
6. Therefore, if agents *A* and *B* share the same evidence *E* on a non-mechanistic issue *I*, it is neither the case that *B* is an epistemic inferior to *A*, nor that *A* is an epistemic inferior to *B*.
7. Thus, when two agents share the same body of evidence on a non-mechanistic issue *I,* they are epistemic peers.

Premises 1–3 are definitional, and I take them as assumptions. Notice that the first is simply a restatement of the definitions of superiority or inferiority as used by a wide base of the philosophy of disagreement literature (e.g., Elga 2006 and Shoenfield 2014, among others). The second statement provides a complete framework for the binary relation of "epistemological knowledge" and follows immediately from the first definition. If I take any other person and any issue *I*, that person is either more likely, less likely, or equally as likely as I am to reach the correct solution to *I*. Therefore, that person must be exactly my epistemic superior, inferior, or peer.

Premise 3 restricts our attention to the class of issues relevant to our discussion, making it the most important definition. It is important to note that we are not considering issues that have no solution, but those that have no evidence-based or mechanistic solutions. In general, these issues fall into two classes: moral issues and unanswered questions. Elga's abortion case is an excellent example of the first class, as the legality of abortion depends almost entirely on the respective moral weights placed on a person's

freedom to choose and another person's right to life.[9] It may well be the case that, objectively, the freedom to choose ought to be given more moral weight than a person's freedom to life, or it may be that no such objective ranking of moral weights exists. In either case, no amount of knowledge or experience can equip a person to come to a foolproof awareness of their existence or nonexistence, so the issue is non-mechanistic. Similarly, some questions have objective answers that reside outside our current scope of knowledge; it is objectively true either that God exists or that he does not (but not both), but completely impossible for any person to provide a proof of either result, even if that person had access to the complete ken of human knowledge or experience. The class of non-mechanistic issues, then, represents a wide swath of issues, including (among other issues) important political questions related to same-sex marriage or the treatment of refugees. In a later section of this paper dedicated to responding to objections, I will further refine this definition to classify exactly which moral questions belong in the world of non-mechanistic issues.

Premise 4 is a collection of all possible ways one agent could label another as either superior or inferior to them on an issue; I have chosen these categories from the relevant literature on disagreement. No other categories appear to be relevant in determining the likelihood of arriving at the correct response to an issue. While there are some objections (e.g., *B* may have access to an epistemic superior in the issue to whom *A* does not, etc.), these can be handled readily by absorbing them into one of the categories (for example, access to a knowledgeable source on *I* could be considered part of the evidence *E*).

Premise 5 is the heart of the argument. Specifically, it asserts that in the class of non-mechanistic issues to which I have restricted my attention, only evidence can create an epistemic hierarchy. As this is the most important premise in the argument, I will attempt to present a thorough justification for its use. First, consider the roles of training and experience, which are common heuristics in assigning another individual to the role of epistemic superior or inferior. However, I claim that in the case of non-mechanistic issues, this reduces to a fallacious appeal to a false authority. Because a specific issue, say the morality of abortion or the existence of God, is non-mechanistic, there is no system of reasoning or training that has yielded a resolution to the issue. In such a scenario, distinguishing between differences in training or experience is impossible. Who is to say that a nurse who has performed hundreds of abortions is more or less likely to understand the correct moral weights assigned to an abortion than a professor of ethics, or that a physicist is more or less likely to be correct on the existence of God than a priest? Neither the training nor the life experience of any of these four occupations can be considered comparable, but there does not exist a clear way to assert dominance of one over the other. Furthermore, given that no clear path to resolution exists, it is not clear that years of experience or prestige of training would matter in assigning value to a person's opinion. If physics is not a confirmed path to understanding the existence of God, why should Stephen Hawking have any more moral authority on

the issue than a peruser of *A Brief History of Time*? For that matter, why should he have any more authority than a layman with no knowledge of physics whatsoever? Once the issue arises, it is self-perpetuating, and backward induction asserts that it is invalid in any potential comparison based on training or experience.

Second, consider the role of history in examining a non-mechanistic issue. Suppose that there was a history of correctness with regard to such an issue—were this the case, it would be included in the sum total of man's knowledge, and, therefore, reduced to a mechanistic issue, as a correct solution is known and accessible. Therefore, the second half of the premise follows merely from the definition of a non-mechanistic issue.

If, then, there are no other criteria by which a person can be labeled an epistemic inferior on some class of issues except through evidence, it follows that two agents who share the same body of evidence each ought to consider the other as their epistemic peer. Notice that this argument solves both the problems of polarization and independence: If these claims are true, and an agent outside of our community has the same evidence we do (even though he has reached a different conclusion), we still ought to consider him our peer. Similarly, those in our like-minded communities (that share evidences as one of their principal functions) will be considered our peers. Additionally, this argument has several critical implications, which I will explore in the subsequent section.

Application of the Conversation Hypothesis to Disagreement

Up to this point, I have been vague about the definition of the "correct" answer to an issue, especially non-mechanistic issues, for which the notion of correctness seems particularly obscure. Notice, however, that one implication of this argument is that the Equal Weights View leads to absurdity. Consider the case of Ann and Beth once more. Assuming, as Elga does, that Ann and Beth have had a lengthy conversation in which each has shared her views and evidences, there is no other criterion to separate Ann and Beth from peer ship. Yet the two have staunchly opposed opinions on the same issue and (at least, after the discussion) evidence that there is no intuition suggesting either will be dissuaded from her view.[10] Were the Equal Weights View legitimate, each would have to judge the other as equally likely to arrive at the correct conclusion, a process that would force both parties into agnosticism about the issue of abortion. Thus, the Equal Weights View, at least for non-mechanistic issues, for which hierarchy is impossible to form except by evidence, seems to imply widespread skepticism about the murkier moral questions.

Many could argue—and indeed, have done so—that these are precisely the questions that need to be answered the most. Indeed, almost every question on political policy can be reduced to a complicated question of balancing moral values.[11] Given that political policy questions (among others) require implementable solutions, skepticism is an undesirable goal. Therefore, it is unavoidable in non-mechanistic issues under the Equal Weights View that situations exist when it is rational to stick to one's original opinion even in the face of an epistemic peer who disagrees with you. This is precisely the hypothesis of *Steadfastness*. Additionally, this seems to imply[12] a form of *Working Permissivism*. That is, even if non-mechanistic issues have unique, correct solutions (and the

questions of existence and uniqueness are a matter for another paper entirely), it is not irrational for two agents to share a body of evidence and reach different conclusions. I will discuss important applications of this result in a later section. This result is interesting for two principal reasons, depending on the type of non-mechanistic issue at hand: For moral issues, it asserts that policy ought not to be based on unanimity, while for unanswered questions, it maximizes the chance of reaching a solution by jointly expanding the base of evidence. A multiplicity of unchanging viewpoints may lead to policy inefficiencies—it may be the case, for instance, that our policies around abortion cycle depending on fluctuations in a nation's majority opinion. However, allowing for this variability creates a space for educated conversation amidst the noise of impossible conversion. For moral issues, whose solutions are usually not theoretically attainable given any amount of experimentation or knowledge, working permissivism yields an important implication in political conversation and policy: listening with compassion. For unanswered questions, this is just as important but for a different reason. Allowing for the priest and the physicist to share information while preserving their own views is critical for interdisciplinary studies and the pursuit of further knowledge. Given that no path of training is more viable than another, allowing each to follow their own candidate path while learning from the triumphs and pitfalls of others increases the chance of success.

Response to Objections: Moral Absolutes and Subjective Evidence

I can see a few potential objections to this argument. The first major objection is empirical and rests on the apparent preponderance of unacceptable responses to non-mechanistic issues. Consider, for example, the issue of female infanticide and its proponents—should not the conversation hypothesis allow agents who hold morally repulsive views as epistemic inferiors? This objection can take two forms. The first is a mere rephrasing of the community hypothesis with the question: Is there a threshold level of ideological difference after which I can label an agent my epistemic inferior? The second is a more puzzling question: Is there an objectively identifiable set of opinions that are unacceptable on moral grounds?[13] The answer to the first question is an easy "no"; the epistemic hierarchy cannot be determined based on the distance between opinions but must be based on the likelihood of correctness. To do anything else would be to assert superiority in experience, training, or history, none of which is valid in answering a non-mechanistic question. On the other hand, the second question is a more difficult one; to examine it further, I will formalize it. Two agents—A and B—disagree on the issue of whether it is morally permissible to murder female infants but share the same evidence E on the issue. If agent B believes that female infanticide is acceptable, is agent A justified in labeling him an epistemic inferior?

It may seem immediately apparent that all issues of morality must be non-mechanistic; it appears that any arbiter possessing the totality of human knowledge would be unable to answer correctly any moral question. However, this seems too hasty. For example, it would appear to many that such an arbiter would clearly be able to identify the objectionable nature of murder and could decisively state that murder is morally

impermissible through an appeal to the human experience. If this is the case, then the objection can be rectified by revising premise 3 to say:

>   3. Agent *A* considers agent *B* her epistemic inferior if and only if *A* is more likely to reach the correct solution to *I* than is *B*, or if *B* has an opinion that requires the violation of a mechanistic moral question (such as the legitimacy of murder).

This revision of premise 3 is also desirable in order to address a similar but perhaps more general objection: one of extreme moral relativism. If we consider any moral issue non-mechanistic, then any response to the current body of evidence is a viable one and policy reflects merely the whims of the majority, whatever they may be. While I do not wish to make any normative claims regarding theories of moral relativism, interpretations of the conversation hypothesis that reduce the illegality of murder to merely the fact that more than 50 percent of the population favor not murdering others seems dubious. It may seem that in some instances, moral questions *can* be answered by enough human experience or experimentation—in such cases, the conversation would not apply, because a moral hierarchy could be built around experience and training, meaning that we need not be permissive about certain ideas, such as murder.

In order to respond fully to this criticism, then, we must address the question: When can a moral issue safely be considered mechanistic, and when can it not? It is likely that this is a question for future responses to the conversation hypothesis and the questions of political disagreement in general. However, I will offer an initial pass at an answer. Recall that a mechanistic issue is answered correctly by examining the complete body of evidence of human experience. We can similarly define a *morally mechanistic issue* as a moral issue that is answered easily by the application of a specific view of morality, such as utilitarianism or the categorical imperative. This is essentially equivalent of reducing moral issues to simple and compound issues; many (if not all) theories of normative ethics would conclude that murder is morally impermissible in the black and white, making murder a morally mechanistic issue. In fact, on simple moral issues, generally acclaimed theories of normative ethics tend to agree, making it almost natural to conclude that the question can be (at least very closely) determined. When context compounds the issues is when these theories diverge. Hence, when a moral issue becomes more non-mechanistic, for example, the "simple" question of murder can be generalized to the "compound" cases of abortion or the question of Jim and the Indians (Smart and Williams 1973). Given that different theories will reach different answers, the arguments above detailing the pitfalls of appealing to experience or training apply, making the issue more non-mechanistic.

A third and final critique lies in my (perhaps) flippant discussion of the evidence surrounding an issue. What exactly constitutes this evidence? Consider, as an example, the case of a jury in a criminal trial; while each juror receives the same evidence provided in the courtroom, might not previous life experiences alter the way evidence is perceived? In this sense, is it ever possible to share evidence completely?

This is exactly the type of problem that the community hypothesis was built to

address. Previous philosophers have relied on arguments such as this to advocate for uniqueness in the face of the empirical evidence of frequently divided juries. However, if we define evidence as strictly objective and tangible (e.g., the testimonies of the witnesses and arguments of the attorneys), the community hypothesis requires no raised eyebrows at the existence of a non-unanimous verdict. While it is true that subjective perception of evidence leads to differing views, this subjectivity is not a problem for non-mechanistic issues.[14] Even among differing perceptions, peer ship is not forfeited.

Applications and Conclusion

Therefore, while we may be under more obligation to view others as our epistemic peers than perhaps previously assumed, we are under less obligation to suspend our own beliefs in order to accommodate disagreement. This result has direct applications to almost any area of interest in the domain of political science, as political activities and behavior are inevitably built on discussion of non-mechanistic issues, such as abortion. In fact, current research in political science empirically substantiates the use of the conversation hypothesis by people utilizing social media sources such as Facebook and Twitter. For example, Kim, Hsu, and de Zuniga (2013) use Facebook network data in the U.S. to conclude that political posts on Facebook contribute to an increased level of "heterogeneity in discussion networks," or an increased likelihood that a political participant will engage in conversation with other political participants from across the political spectrum.[15] In general, the authors find these increased conversations from "across the aisle" lead to an increase in civic engagement.

One important gap in this research, however, is an examination of the results of such cross-partisan discussions. Specifically, current political science research has been silent on the ways these conversations affect the initial opinions of their participants. Future researchers may focus on analyzing these political conversations in a setting that allows them to determine whether the conversation 1) brought the participants to a "middle ground" solution, as predicted by the Median Voter Theorem; 2) polarized the participants even more, as suggested by either the community hypothesis or the stand-alone hypothesis; or 3) encouraged a socially beneficial conversation without shifting each agent's own political opinions, as predicted by the conversation hypothesis. This research would have important political implications on fostering political discussion among heterogeneous groups and answer questions such as ideological selectivity in social media use (Iyengar and Hahn 2009). Once established, these results extend to almost any aspect of political science that relies on political dialogue about non-mechanistic issues, such as policy formation, political economy, and international relations.

On issues the collective human spirit has not adequately addressed, we have no sufficient criteria for distinguishing epistemic superiority and, therefore, cannot close our ears or minds to the arguments of those who have reached different conclusions than us. However, this does not need to lead to widespread skepticism or the sur-

rendering of belief. Rather, the exposure to different views on the same evidence can have a purely positive impact on discussion and embodiment overall. As Nietzsche wrote, "let us not be ungrateful toward such resolute reversals of the familiar perspectives . . . to see differently in this way for once, to *want* to see differently, is no small discipline and [serves] . . . so that one knows how to make precisely the *difference* in perspectives and affective interpretations useful for knowledge" (*On the Genealogy of Morality,* III.12).

NOTES

1. I recognize that for many of the issues discussed in this paper, the use of the word "correct" to describe a solution to the issue may seem vague. Over the course of the paper, I hope to clarify this sentiment.

2. These definitions appear in broad strokes—there are many variations on a theme when it comes to the semantics of the debate. However, in order to focus on the more critical issues, I will attempt to use definitions that appeal to many philosophers' views.

3. As a pass at a definition of these types of issues (which I think synthesizes the views of philosophers such as Elga who have written on the subject), I assert that an issue is non-mechanistic if it is impossible to amass a body of evidence E large enough to arrive at a unique answer to the question being considered. According to another way of thinking, an issue could be considered mechanistic if we could reasonably conceive of a computer or robot possessing the sum total of humanity's knowledge answering the question for us, thereby eliminating the need to handle disagreement stemming from human error. I will clarify this definition further in the next section of the paper.

4. We can assume, for the purposes of this paper, that the discovery that an epistemic superior holds an opinion contradictory to your own is strong enough evidence for you to change your opinion. Whether this is true is a matter for another paper.

5. For one thing, it seems to be begging the question by arguing that a label of epistemic peer is conditional on minimizing disagreement.

6. The validity of this claim is questioned by some, who maintain that we merely *think* that we do this when, in fact, our opinions are quite inconsistent. While this may be true, it is the very perception of opinion-forming that I am concerned with, and so this objection is an irrelevant one.

7. Unless some obvious condition exists, which labels one of us as superior.

8. i.e., a better history of being correct on issues similar to (or exactly identical to, in the case of mechanistic issues).

9. This is a broad-strokes portrayal of the issue in order to avoid a larger tangent.

10. Even Elga avoids claiming this, as the bulk of human interaction would suggest the opposite.

11. As in the case of abortion, in which different opinions arise from assigning different preferences to the values of freedom and the right to life.

12. Although the argument for this conclusion would take substantially more work.

13. An even more basic question may arise concerning the labelling of any view as morally repulsive within the domain of a non-mechanistic issue. Here, we must remember that non-mechanistic issues are not unsolvable ones—opinions held by agents assessing a body of evidence are considered candidate solutions to an issue, and one agent's candidate solution may be viewed as repugnant or extremely off-base by another agent. Specifically, a view may be repulsive because of its implications for other, more mechanistic issues, rather than its contrast to the true solution (which is inaccessible for these questions). For example, within the question of abortion's morality, it may be the case that a particular agent holds that children hold no rights at all until they are able to contribute to society, thereby condoning not only abortion but also other forms of child abuse (such as female infanticide). This candidate solution would be viewed as morally repulsive by many, even though there is no true solution with which to compare the idea.

14. Indeed, no standard of objectivity exists.
15. This article also includes important references to other studies attempting to understand the role of political conversation on social media sites in elevating offline political participation, and the role of different personal characteristics in mitigating these results.

REFERENCES

Christensen, David. 2007. "Epistemology of Disagreement: The Good News." *The Philosophical Review* 116, no. 2 (April): 187–217.

Christensen, David. 2009. "Disagreement as Evidence: The Epistemology of Controversy." *Philosophy Compass* 4, no. 5 (September): 756–67.

Christensen, David. 2016. "Conciliation, Uniqueness and Rational Toxicity." *Noûs* 50, no. 3 (September): 584-603.

Elga, Adam. 2007. "Reflection and Disagreement." *Noûs* 41, no. 3: 478–502.

Foley, Richard. 2001. *Intellectual Trust in Oneself and Others*. Cambridge: Cambridge University Press.

Goldman, Alvin I. 2010. "Epistemic Relativism and Reasonable Disagreement." *Disagreement*: 187–215.

Iyengar, Shanto and Kyu S. Hahn. 2009. "Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use." *Journal of Communication* 59, no. 1: 19–39.

Kelly, Thomas. 2005. "The Epistemic Significance of Disagreement." *Oxford Studies in Epistemology* 1: 167-196.

Kelly, Thomas. 2010. "Peer Disagreement and Higher Order Evidence." *Social Epistemology: Essential Readings*: 183–217.

Kim, Yonghwan, Shih-Hsien Hsu, and Homero Gil de Zuniga. 2013. "Influence of Social Media Use on Discussion Network Heterogeneity and Civic Engagement: The Moderating Role of Personality Traits." *Journal of Communication* 63, no. 3 (June): 498–516.

Kopec, Matthew and Michael G. Titelbaum. 2016. "The Uniqueness Thesis." *Philosophy Compass* 11, no. 4 (April): 189–200.

Nietzsche, Friedrich. 1887. *On the Genealogy of Morality: A Polemic*. Translated by Maudemarie Clark and Alan J. Swensen, 1998. Indianapolis: Hackett Publishing.

Rinard, Susanna. 2017. "No Exception for Belief." *Philosophy and Phenomenological Research* 94, no. 1 (January): 121–43.

Schoenfield, Miriam. 2014. "Permission to Believe: Why Permissivism is True and What It Tells Us about Irrelevant Influences on Belief." *Noûs* 48, no. 2 (June): 193–218.

Sher, George. 2001. "But I Could Be Wrong." *Social Philosophy and Policy* 18, no. 2: 64–78.

Smart, J.J.C. and Bernard Williams. 1973. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.

Vavova, Katia. 2014. "Moral Disagreement and Moral Skepticism." *Philosophical Perspectives* 28, no. 1 (December): 302–33.

White, Roger. 2005. "Epistemic Permissiveness." *Philosophical Perspectives* 19, no. 1 (December): 445–59.

White, Roger. 2010. "You Just Believe that Because. . ." *Philosophical Perspectives* 24, no. 1 (December): 573–615.

Women vs. Men—Who Makes Better Use of Financial Aid?

Colton Keddington and Nichole Keddington

Introduction

Human capital is one's ability to perform labor to produce value (Goldin 2014). As marketable skills increase, human capital increases. Post-secondary education is a valuable component in human capital, because it increases the skill set that enables a person to work and make a living. In the last few decades, through post-secondary education, U.S. women increasingly have obtained more human capital (Executive Office of the President of the United States 2014).

Women's college attendance has increased substantially over time, and women in the U.S. now graduate from college at higher rates than men do. Ishitani's study shows that women are 56 percent more likely to graduate in four years than men (2006), and Astin points out that women also tend to have better grades in college even when controlling for high school grades (1997). Although once barred from higher education completely, women now make up the majority of college students in the U.S. (about 11.5 million females in 2017, compared with 8.9 million males, according to the National Center for Education Statistics). Women also graduate at higher rates than men do (about 1.8 million females in 2015, compared with about .8 million males). Women surpassing men in both college attendance and graduation rates is evidence of underlying differences in the factors that contribute to a person's college education. We further investigate these differences by analyzing how financial resources help members of each gender graduate from college.

We analyze how being a federal financial aid recipient contributes to a person's likelihood of graduation. We theorize that women who receive financial aid will be more likely to graduate than men who receive financial aid. This hypothesis can be viewed as a test of whether or not the economic development literature, which

is primarily tested in Third World countries, may apply to First World settings. We also theorize that females who receive financial aid are more likely to graduate than both females and males who do not receive financial aid.

Literature Review and Theoretical Framework

The existing research on women with financial aid versus men with financial aid is limited while comprehensive data on students' academic performance and financial aid statuses is essentially nonexistent. We do not claim that no research has been done on this subject, as some studies on women and financial aid in higher education do exist. Hossler, Hu, and Schmidt studied women and financial aid in higher education and found that women apply and enroll at lower rates when financial aid is not available (1999). This gap disappeared when controlling for family financial support, suggesting that women are more sensitive to tuition costs than men are. Consequently, women probably apply, enroll, and graduate at higher rates than men when given financial aid.

Some investigation has occurred on the effect of financial aid on retention and graduation. In general, Wohlgemuth et al. find that providing financial aid does increase the rate of retention (2007). Furthermore, Hossler, Gross, and Ziskin find that women are retained more than men are when institutional aid is offered (2006). However, this was not the main focus of their study, and their coding when gender is unknown is questionable. Furthermore, Hossler, Gross, and Ziskin include all missing data as female. The potential issues of this practice are many but easily stated. We can assume that at least some portion of those who do not list their gender are male. One possible example is men who fear they will not receive fair treatment due to their perception of institutions attempting to help females more than males. Perhaps realizing this issue themselves, Hossler, Gross, and Ziskin state that an investigation into women on financial aid in higher education deserves its own study. We hope to add substantively to these studies.

Among the strongest of the analyses on women and financial aid in higher education is that of Fenske, Porter, and DuBrock (2000). In their study, they analyzed the persistence of women majoring in science, engineering, and mathematics (SEM).[1] Measuring persistence by graduation rates, they found evidence suggesting that women have a lower attrition rate than other groups of financial aid recipients. Their analysis is highly focused on SEM majors and does not address the need for a more generalized study of male and female use of financial aid in higher education.

We seek to do an analysis similar to Fenske, Porter, and DuBrock but with a broadened scope from SEM majors to all students, with a larger emphasis on gender differences. We show how women who receive financial aid differ from both men and women who do not receive financial aid. In essence, we argue that women take better advantage of the financial opportunities given them and that those that receive financial aid are significantly different from those that do not receive financial aid. Women see themselves as responsible to graduate once they have received financial aid. Thus,

receiving financial aid is a strong motivator for women to complete their postsecondary education. When the opportunities and resources are available to them, women will be more likely to graduate from college.

Methods

To test this hypothesis, we will apply the 1987 National Postsecondary Student Aid Survey data and use logit regressions to determine the likelihood of different groups graduating from college. This data set includes information on 12,628 students who applied for Guaranteed Student Loans (GSL). Within this data set are variables indicating whether the student received financial aid ("Any Aid") and, if so, how much ("Aid Amount"). We limit our analyses to only students in need of financial aid. We do so by only including students who have applied for GSL. We recognize that not all students in need of financial aid apply for GSL. Therefore, our study is scientifically applicable to GSL applicants and arguably not to all students in need of financial aid. However, applying for GSL shows a need for financial aid, and the use of the GSL applicant data allows us to generalize onto all students with financial aid needs. Using GSL applicants provides a way to compare aid recipients to non-aid recipients, while eliminating bias that may arise from including students who have no need for financial aid.

Using many control variables, we will test to see if women complete their postsecondary education more than men when given student financial aid. Some of the key controls included are age, income, marital status, and race. Especially important is the control "Total Cost," which indicates the reported total cost per one year of school, depending on where that particular student attended college. If gender is unknown, the observation will not be included in the model. Doing so specifically addresses the methodological issue of Hossler, Gross, and Ziskin's study. These tests will allow us to analyze any relationship between gender and graduation rates among financial aid recipients and between recipients and non-recipients.

Limitations

The main limitation to this study is the available data, in particular the year collected. It is reasonable to argue that data from 1987 is irrelevant to today's world. While we maintain that the principles and hypothesis still hold true today, we admit it is a reasonable critique. The National Center of Education Statistics has done a similar survey as recently as 2012 but will only give out the data to those with a license. Unfortunately, as undergraduates we do not qualify to obtain a license and are forced to use older data.

Perhaps due to data age, the data set itself proposed some issues. The data started in ASCII format and was saved on the University of Michigan's Inter-University Consortium for Political and Social Research (ICPSR) web site. The data was publicly available for download and SAS instructions provided. A professional full-time staff member at the university converted the data from ASCII format to a format readable

by more modern statistical analysis software like Stata. Unfortunately, it seems that a portion of the data failed to convert. The observations that failed to convert were dropped, and the remaining 12,000 observations were used for this study.

Another issue with this dataset is the way some of the races are coded. The appendix contains the exact coding and tabulations of the race variable. The main issue with the race coding is the combining of Asian and Pacific Islander. Perhaps combining the two groups into one was accepted practice when this dataset was collected, but political science methodology has grown and combining the two groups is rarely, if ever, done anymore. As seen in all of the Appendix tables, the Asian variable is seemingly statistically significant. We choose to ignore these results because of the potential issues caused by the combining of Asian and Pacific Islander into one race category and the impossibility of disentangling the two in this data.

Also, there is a discrepancy between the data set and the codebook. Although the stated universe is "all" students who applied for and received GSLs, data exists for whether or not each student "received any financial aid." We assume, then, that loans do not qualify as financial aid, and we can make broader assumptions about recipients and non-recipients, even though all students have received GSLs. This makes our universe a sample of financial aid recipients and non-financial aid recipients who have all exhibited a need for aid. It does not include, however, students who did not apply for a GSL. We do argue that applying for GSL shows a need for financial aid and thus allows us to generalize our results onto all students in need of financial aid. Despite the limitations explained above, we maintain our general results are accurate and applicable.

Findings

The reported coefficients from our analyses is found in the Appendix. There, we also include demographic tabulations of the 1987 National Postsecondary Student Aid Survey data available to us. We began by using the simplest model possible by regressing if a person had "completed degree/course of study" on gender. The results were statistically insignificant. Such a finding seemed odd as much of the current literature and data indicate that women are graduating at higher rates than men are (Executive Office of the President of the United States 2014). The simple answer to this is to look at the longitudinal data on graduation rates. This 1987 data fits into the time period where most data indicate that men and women graduated at statistically the same rate (Ibid.). Please see Figure 5 in the Appendix for more information. Model 1 in Table 8 (see Appendix) is in line with Current Population Survey data as calculated and presented by the White House's Council of Economic Advisers. The consistency between existing literature and data on graduation rates in 1987 works to support the claim that our data set, although limited, is valid.

The lack of statistical difference in the gender variable in the simple model adds intrigue as it becomes significant in the simple model that tests our theory (see Appendix Table 8, Model 2). In this logistic model, we added the interactive term of "Female and Any Aid." This interactive term is the main variable of interest in Table 8 (see

Appendix). Here, we see that women have statistically different logged odds of graduating or completing their study than men given they receive any financial aid.

Table 1

VARIABLES	(2)	(7)
Female	-0.156** (0.0710)	0.0258 (0.119)
Any Aid	0.00164 (0.0781)	-0.278** (0.134)
Female × Any Aid	0.309*** (0.106)	0.364** (0.175)
American Indian		-0.482 (0.382)
Asian		-1.246*** (0.166)
Black		-0.178 (0.158)
Hispanic		-0.452*** (0.164)
Age		-0.0395 (0.0270)
Age2		0.000291 (0.000341)
Income		2.75e-09 (1.09e-07)
Married		-0.183* (0.0962)
Separated		-0.123 (0.277)
Total Cost		0.0001*** (2.83e-05)
Constant	1.758*** (0.0539)	2.413*** (0.514)
Observations	11,552	4,026

Models 2 through 7 in Table 8 (see Appendix) support the claim that women with financial aid graduate at higher rates than men do with financial aid. The simplest of these models, Model 2, shows that females receiving financial aid have an 84 percent likelihood of graduation, while males have a 78 percent likelihood of graduation. This 6 percent is statistically significant. Model 7 shows that after including all relevant controls, women receiving aid still have statistically significantly higher logged odds of graduating. Figure 1 is a visual representation of Model 7 Table 8 (see Appendix).

Models 2 through 6 in Table 9 (see Appendix) are conceptually important to our theory. Here, "Aid Amount" is added to the model. Doing so filters the observations to only those that received some aid. In effect, all variables in models 2 through 6 in Table 9 (see Appendix) are interacted with "Any Aid." Because Models 2 through 6 only contain

those that received financial aid, and control for how much aid was received and the total cost of one year of schooling, the gender variable becomes more telling of our theory. Every model except Model 3 in Table 9 (see Appendix) shows that women have higher logged odds of graduating than men given that both are receiving the same amount of financial aid and their schooling costs the same.
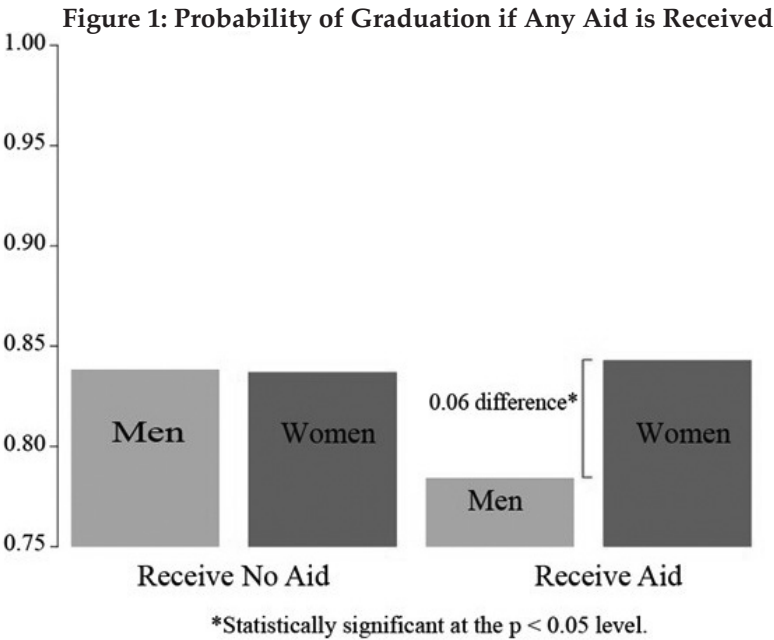
Figure 1: Probability of Graduation if Any Aid is Received

*Statistically significant at the p < 0.05 level.

Table 10 (see Appendix) attempts to uncover if there is one particular subgroup of women raising the odds of graduation or completion. In Table 10 (see Appendix), the observations are limited to women only and allow us to compare the women in this data set. Models 2 through 6 have the "Aid Amount" control variable and are filtered to only those observations that received some aid. The total cost of one year of schooling is very significant in all models in Table 10 (see Appendix). Interestingly, as the total cost of a year of school goes up, the woman is more likely to graduate. (See Figure 2 below.) This finding also appears to be true for men, as this control is statistically significant in all models of Table 9 (see Appendix) as well. One possible explanation for this trend is that only people that have the means and drive to finish a degree choose to attend a more expensive school. Those that question if they have the time, resources, or drive to finish would logically not risk as much money by attending a more expensive program. It is likely that those attending the more expensive school are systematically different from those that attend the less expensive schools.

Table 2

VARIABLES	(2)	(6)
Female	0.185** (0.0815)	0.363*** (0.135)
Total Cost	0.000142*** (2.29e-05)	0.000136*** (3.35e-05)
Aid Amount	-3.47e-06 (1.64e-05)	-3.18e-06 (2.38e-05)
American Indian		-0.275 (0.538)
Asian		-1.652*** (0.227)
Black		-0.256 (0.226)
Hispanic		-0.697*** (0.234)
Age		-0.0187 (0.0386)
Age2		-0.000112 (0.000492)
Income		4.09e-07** (1.93e-07)
Married		-0.257* (0.144)
Separated		-0.124 (0.367)
Constant	1.294*** (0.0960)	2.104*** (0.716)
Observations	5,240	1,904

Among women, the amount of aid is also significant. (See Models 2 through 6 in Table 10 in the Appendix.) In the table, the coefficients are very small, ostensibly indicating that the aid amount has no substantive effect. However, "Aid Amount" is calculated in single dollar increments, and aid is rarely if ever given in single dollar increments. For example, each additional $1,000 of aid given to a female increased her logged odds of graduation by 1 percent. The median amount of aid given in this data set is $2,370. A female student with the median amount of financial aid has, according to our results, 86.7 percent likelihood of graduating from college, as compared to females in the 25th percentile of aid (85 percent likelihood of graduation) or females in the 75th percentile of aid (88 percent likelihood of graduation). Realizing that the coefficient of "Aid Amount" will change by hundreds and thousands helps us understand the substantive effect is more than initially indicated by the small coefficient. (Figure 3 illustrates this point.)

Another variable that seems to be significant in both Table 9 and Table 10 (see Appendix) is race. White is used as the baseline for the race comparisons listed in

Figure 2: Probability of Graduation
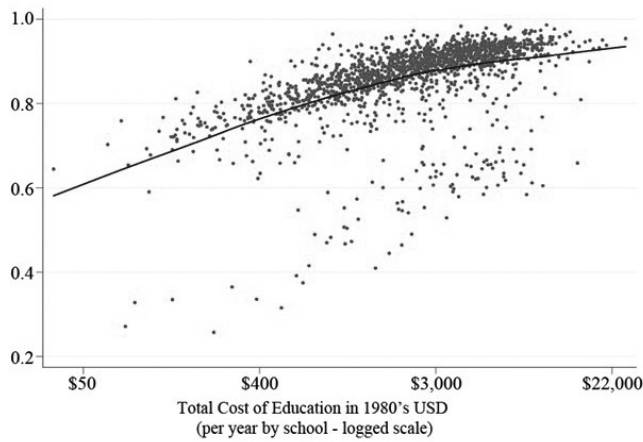
Women with Student Aid

Total Cost of Education in 1980's USD
(per year by school - logged scale)

Table 3

VARIABLES	(2)	(6)
Total Cost	0.000152*** (3.37e-05)	0.000178*** (5.54e-05)
Aid Amount	5.84e-05** (2.69e-05)	9.70e-05** (4.62e-05)
American Indian		-0.416 (0.825)
Asian		-1.591*** (0.396)
Black		0.0609 (0.300)
Hispanic		-0.761** (0.314)
Age		-0.0191 (0.0502)
Age2		5.95e-05 (0.000624)
Income		2.11e-07 (2.45e-07)
Married		0.0576 (0.194)
Separated		-0.124 (0.405)
Constant	1.300*** (0.117)	1.783* (0.952)
Observations	2,851	1,096

the Appendix tables. Hispanic is significant in Table 9 (see Appendix) in all models where race is included. In Table 10 (see Appendix), among women, Hispanic is not as significant until income is considered. The data seem to indicate that Hispanics have lower logged odds of graduating when compared to whites. (See Figure 4.)

The other piece of the race variable that comes out significant in every instance is Asian. We have not neglected to notice this occurrence, but we have reason to believe this finding is inaccurate. As noted in the limitations section above and in the Appendix, Asian and Pacific Islander are coded together in this 1987 dataset. While combining Asian and Pacific Islander may have been common practice in 1987, it is not today. Today, Asian and Pacific Islander are coded separately for many reasons. Because of this issue, we choose not to claim any valid results comparing the odds of Asians or Pacific Islanders to the baseline White category.

Figure 3: Probability of Graduation

Amount of Aid Given To Recipient
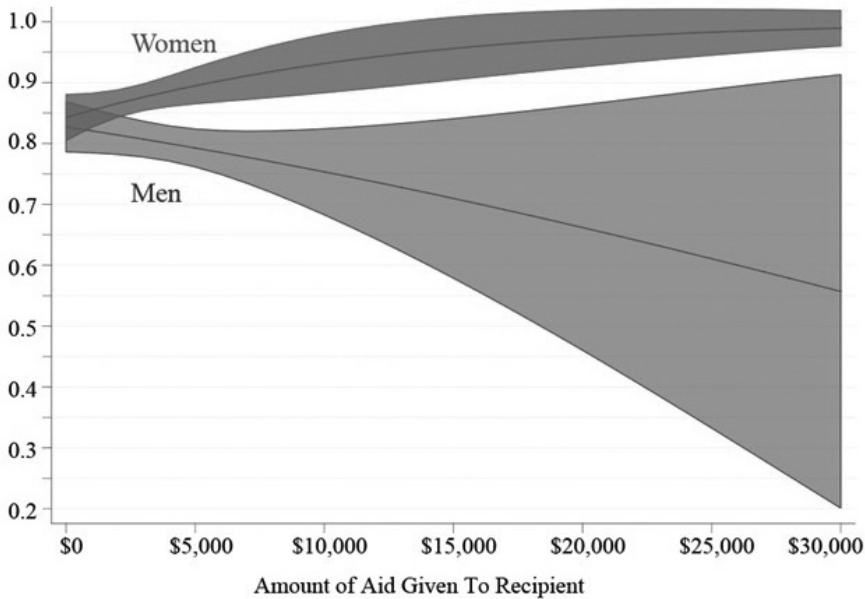
Conclusion

We find many noteworthy factors relating to likelihood of graduation. Among women, the amount of aid given significantly increased the likelihood of graduation. In both the simplest and most comprehensive models, we find that females with aid are more likely to graduate than females without aid and males, both with and without aid. This is our most substantive finding. It seems the data from the 1987 National Postsecondary Student Aid Survey support our hypotheses that women use student financial aid more effectively than men do.

Concluding that women use student financial aid better than men do supports the economic development literature referenced above. Wong and Psacharopoulos

appear to support the idea that women use resources more effectively than men do by showing how children improve when the woman has control of the income. We argue that this theory is not limited to development economics. We find evidence to suggest that the same theory is true in higher education aid within United States.

The implications of these results reach into both practical and academic spheres. For the institutions deciding who gets student financial aid, the results of these analyses may be particularly helpful. For academia, this study may aid in increasing the external validity of some gendered economic development theories.

Figure 4: Probability of Graduation
Women with Student Aid
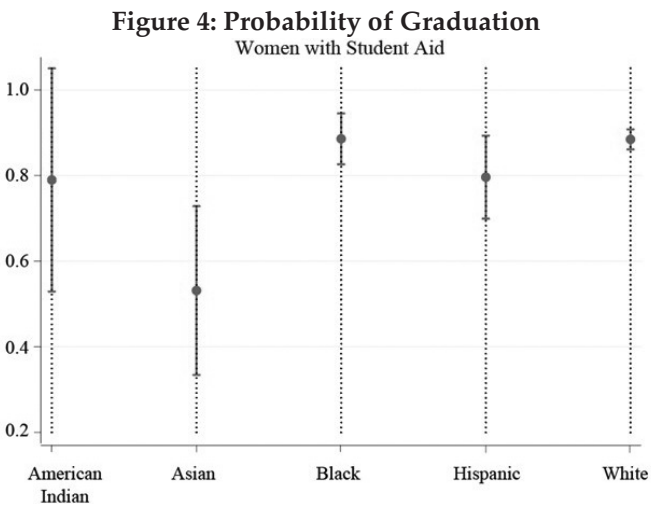
The limitations of this study are the data. However, the 1987 National Postsecondary Student Aid Survey provides a place to start. The more recent versions of this survey, done as recently as 2012, will likely have fixed the data and coding issues. Further research on more recent data would provide better insight into the current use of student financial aid. If our findings hold, there would also be a need to identify what specific mechanisms lead to women's increased graduation rates as a result of receiving financial aid. This information will help us know how to apply those factors to men so that members of both sexes can more effectively use the resources available to them.

NOTE
 1. Now referred to as "STEM."

APPENDIX

Description of data set

In our analyses, we included the variable age^2 to account for the distribution of ages in the population.

All observations received some amount of financial aid in this table.

Table 4

"Have completed degree/course of study"	Frequency	Percent
No	1,713	14.83
Yes	9,839	85.17

Dependent Variable: graduation status.

Table 5

Sex	Frequency	Percent
Male	5,761	45.62
Female	6,867	54.38

Key Independent Variable: female.

Table 6

Any Aid	Frequency	Percent
No	6,816	53.98
Yes	5,812	46.02

Independent Variable: any aid.

Table 7

Race	Frequency	Percent
American Indian or Alaska Native	102	.81
Asian or Pacific Islander	719	5.67
Black, not Hispanic	1,073	8.5
Hispanic	781	6.18
White, not Hispanic	9,953	78.82

Independent Variable: race.
Independent Variable: age.

Figure 5
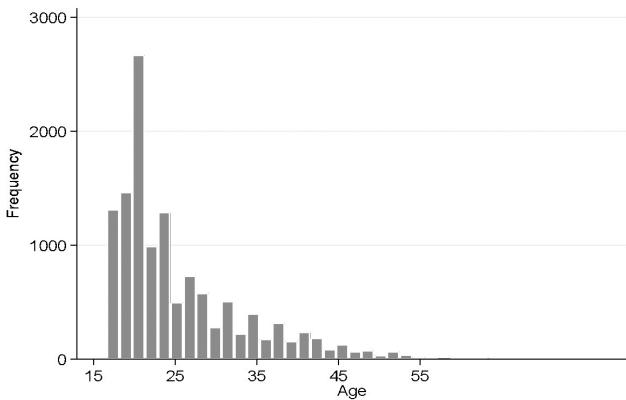
Table 8

VARI-ABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.0198 (0.0526)	-0.156** (0.0710)	-0.184** (0.0717)	-0.119 (0.0728)	-0.0225 (0.115)	-0.0223 (0.115)	0.0258 (0.119)
Any Aid		0.00164 (0.0781)	0.00711 (0.0786)	-0.0224 (0.0787)	-0.0431 (0.127)	-0.0676 (0.128)	-0.278** (0.134)
Female × Any Aid		0.309*** (0.106)	0.294*** (0.107)	0.242** (0.107)	0.362** (0.168)	0.352** (0.168)	0.364** (0.175)
American Indian			-0.196 (0.285)	-0.140 (0.284)	-0.380 (0.378)	-0.394 (0.379)	-0.482 (0.382)
Asian			-0.910*** (0.0906)	-0.942*** (0.0908)	-1.086*** (0.155)	-1.131*** (0.156)	-1.246*** (0.166)
Black			0.0935 (0.103)	0.125 (0.103)	-0.0350 (0.150)	-0.0964 (0.153)	-0.178 (0.158)
His-panic			-0.148 (0.107)	-0.138 (0.108)	-0.293* (0.158)	-0.325** (0.158)	-0.452*** (0.164)
Age				-0.0836*** (0.0163)	-0.0532** (0.0256)	-0.0404 (0.0262)	-0.0395 (0.0270)
Age2				0.0008*** (0.000244)	0.000414 (0.000328)	0.000268 (0.000334)	0.000291 (0.000341)
Income					-3.74e-08 (1.03e-07)	-2.60e-08 (1.04e-07)	2.75e-09 (1.09e-07)
Married						-0.230** (0.0932)	-0.183* (0.0962)
Sepa-rated						0.00207 (0.272)	-0.123 (0.277)
Total Cost							0.0001*** (2.83e-05)
Con-stant	1.759*** (0.0390)	1.758*** (0.0539)	1.849*** (0.0566)	3.398*** (0.254)	2.873*** (0.489)	2.782*** (0.490)	2.413*** (0.514)
Obser-vations	11,552	11,552	11,552	11,552	4,179	4,179	4,026

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9

VARI-ABLES	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0448 (0.0544)	0.185** (0.0815)	0.130 (0.0833)	0.147* (0.0841)	0.376*** (0.134)	0.363*** (0.135)
Total Cost	0.000179*** (1.64e-05)	0.000142*** (2.29e-05)	0.000147*** (2.34e-05)	0.000124*** (2.27e-05)	0.000136*** (3.38e-05)	0.000136*** (3.35e-05)
Aid Amount		-3.47e-06 (1.64e-05)	1.09e-05 (1.66e-05)	1.85e-05 (1.64e-05)	1.03e-06 (2.39e-05)	-3.18e-06 (2.38e-05)
American Indian			-0.411 (0.371)	-0.374 (0.370)	-0.264 (0.538)	-0.275 (0.538)
Asian			-1.356*** (0.138)	-1.386*** (0.137)	-1.600*** (0.224)	-1.652*** (0.227)
Black			0.123 (0.143)	0.118 (0.144)	-0.193 (0.221)	-0.256 (0.226)
Hispanic			-0.345** (0.156)	-0.344** (0.156)	-0.662*** (0.233)	-0.697*** (0.234)
Age				-0.0653*** (0.0245)	-0.0371 (0.0376)	-0.0187 (0.0386)
Age2				0.000599* (0.000363)	9.71e-05 (0.000482)	-0.000112 (0.000492)
Income					3.94e-07** (1.90e-07)	4.09e-07** (1.93e-07)
Married						-0.257* (0.144)
Separated						-0.124 (0.367)
Constant	1.249*** (0.0577)	1.294*** (0.0960)	1.401*** (0.101)	2.683*** (0.396)	2.293*** (0.713)	2.104*** (0.716)
Observations	11,126	5,240	5,240	5,240	1,904	1,904

Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 10

VARI-ABLES	(1)	(2)	(3)	(4)	(5)	(6)
Total Cost	0.000197*** (2.25e-05)	0.000152*** (3.37e-05)	0.000155*** (3.42e-05)	0.000141*** (3.37e-05)	0.000177*** (5.49e-05)	0.000178*** (5.54e-05)
Aid Amount		5.84e-05** (2.69e-05)	6.54e-05** (2.75e-05)	6.62e-05** (2.69e-05)	9.40e-05** (4.48e-05)	9.70e-05** (4.62e-05)
American Indian			-0.185 (0.553)	-0.154 (0.554)	-0.429 (0.818)	-0.416 (0.825)
Asian			-1.405*** (0.219)	-1.421*** (0.219)	-1.594*** (0.398)	-1.591*** (0.396)
Black			0.0760 (0.180)	0.0732 (0.180)	0.0351 (0.294)	0.0609 (0.300)
Hispanic			-0.386* (0.211)	-0.389* (0.210)	-0.778** (0.313)	-0.761** (0.314)
Age				-0.0429 (0.0311)	-0.0178 (0.0501)	-0.0191 (0.0502)
Age2				0.000437 (0.000454)	4.35e-05 (0.000621)	5.95e-05 (0.000624)
Income					2.15e-07 (2.45e-07)	2.11e-07 (2.45e-07)
Married						0.0576 (0.194)
Separated						-0.124 (0.405)
Constant	1.251*** (0.0626)	1.300*** (0.117)	1.388*** (0.125)	2.211*** (0.504)	1.793* (0.960)	1.783* (0.952)
Observa-tions	6,093	2,851	2,851	2,851	1,096	1,096

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
This table shows differences in factors influencing graduation rates among women only. Men are excluded from this table.

Figure 6

Postsecondary Attainment Has Risen Among Men and Women; Women Now Complete College and Graduate School at Higher Rates Than Men
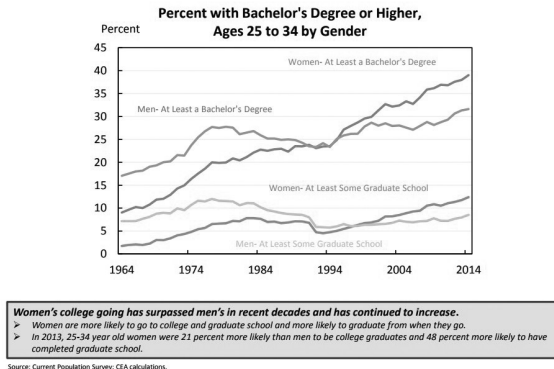
Figure is from the October 14, 2014, update of the Women's Participation in Education and the Workforce report by the Council of Economic Advisers. We include their graphic to show where our 1987 data fit in the overall trend.

REFERENCES

Astin, Alexander W. 1997. *What Matters in College?: Four Critical Years Revisited*. San Francisco: Jossey-Bass.

Executive Office of the President of the United States. 2014. "Eleven Facts about American Families and Work." https://obamawhitehouse.archives.gov/sites/default/files/docs/eleven_facts_about_family_and_work_final.pdf.

Goldin, Claudia. 2016. "Human Capital." In *Handbook of Cliometrics*, edited by Claude Diebolt and Michael Haupert, 55–86. Heidelberg, Germany: Springer Berlin Heidelberg.

Horn, Laura, and Katharin Peter. 2005. "Gender Differences in Participation and Completion of Undergraduate Education and How They've Changed over Time. Postsecondary Education Descriptive Analysis Reports." U.S. Department of Education.

Hossler, Don, Jacob P.K. Gross, and Mary Ziskin. 2006. "A Multi-Institution Analysis of the Effects of Campus-Based Financial Aid on Student Persistence at Public Four-Year Institutions." Indiana University.

Hossler, Don, Shouping Hu, and Jack Schmidt. 1999. "Predicting Student Sensitivity to Tuition and Financial Aid." *Journal of Student Financial Aid* 29, no. 1 (February): 17–33.

Ishitani, Terry T. 2006. "Studying Attrition and Degree Completion Behavior among First-Generation College Students in the United States." *Journal of Higher Education* 77, no. 5 (September-October): 861–85.

Jejeebhoy, Shireen J. 1995. *Women's Education, Autonomy, and Reproductive Behaviour: Experience from Developing Countries: Experiences from Developing Countries*. Oxford: Clarendon Press.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22, no. 9 (September): 1325–343.

Schultz, T. Paul. 1998. "Contents of the Handbook." Handbook of Development Economics 1.

Wohlgemuth, Darin, Don Whalen, Julia Sullivan, Carolyn Nading, Mack Shelley, and Yongyi (Rebecca) Wang. 2007. "Financial, Academic, and Environmental Influences on the Retention and Graduation of Students." *Journal of College Student Retention: Research, Theory & Practice* 8, no. 4: 457–75.

Wong, Yen Nee. "World Development Report 2012: Gender Equality and Development." *Forum for Development Studies* 39, no. 3 (November): 435–44.

Improving the Bureaucracy: What Leads Government Officials to Use Evidence-Based Reports?

Nicholas Moffitt

Introduction to General Problem Area

"Suppose you were an idiot. [Now] suppose you were a member of Congress, but I repeat myself" (Twain 1907). Although many people think negatively of politicians and governments in general, both the public and private sector are taking actions to improve the quality of programs produced by governments and, hopefully, improve people's lives and their perceptions of the government. To improve the quality of government programs, scholars have produced many assessments that precisely measure the impacts of public policies, but government officials typically do not use these evidence-based reports (EBRs) in designing policy (Head 2016). More rigorous policy evaluation started in the 1970s, but these policy evaluations only truly gained momentum in the 1990s (Head 2016, p. 472).

Much of the impetus of the EBR movement stemmed from economists pushing randomized control trials (RCTs) in public policy (Head 2016, p. 474). In government-sponsored RCTs, various subsections of a population, which ideally reflect each other and the entire population, each take part in unique public policies. After these RCTs run for a sufficient time, data scientists analyze the results of the experiment and compare the two groups. If the experimental groups share equal traits, the researchers can then infer that the interventions in the public policies caused the different results (Banerjee and Duflo 2009, p. 152). RCTs, therefore, more precisely measure policy outcomes than simple observational data on the program, because they control for other influencing factors through the experimental design. In the mid-1990s, when economists were beginning to lay much of the groundwork for experiments on public policies, they found that governments and developmentally oriented NGOs were spending thousands of dollars on programs that resulted in little real improvement.

For example, after running initial RCTs on educational programs in Africa, prominent economists found that policies geared at increasing years of education ranged in cost effectiveness from $3.25 to $6,000 per year of increased education (Banerjee and Duflo 2009, p. 153).

The results of these studies clearly showed that many governments were not effectively using their resources or measuring the results of their policies; they needed to apply serious analysis to their programs to evaluate which ones produced the best results for the least time and money invested. In turn, by using RCTs and then reporting those results in EBRs, government officials can better evaluate which policies produce the best results, and they can pinpoint the cost for each outcome produced. Unfortunately, despite the rising prevalence of EBRs in the 1990s, their use increased almost exclusively in the private sector and only slowly started influencing policies in the public sector. Clear examples of this varied implementation of EBRs comes from the health industry, where medical professionals started heavily using RCTs and EBRs in the 1990s, yet public policy regarding health hardly took notice of them (Black 2001). Other sources that birth public policies include ideology, partisan lines, and corruption. Because implementation of EBRs by government officials runs scarce in the developing world, this project will identify which characteristics of government officials and which types of hiring and monitoring processes of bureaucrats lead officials to have the most proficiency in, interest in, and intent in using EBRs. Because people elect members of parliament and other legislators rather than hire them, the term "officials" refers to bureaucratic employees hired through a merit-based application process.

The efficacy of any government relies heavily on the abilities of unelected officials, yet citizens often have no direct say in which unelected officials are hired. While well-educated, thoroughly trained, and closely evaluated officials perform the best work, many governments suffer from poorly prepared officials (Rajkumar and Swaroop 2008). Government officials often appeal to ideology, emotion, intuition, and unique personal experiences when making policy decisions (Banks 2010). However, governments should instead improve their effectiveness by precisely measuring the results of specific policies using evidence-based reports. Methods of using such EBRs have shown to increase program effectiveness in various fields including medicine, education, agriculture, technology, and others (Slavin 2008).

Unfortunately, a huge gap still exists between the academic world, with its evidence-based reports, and government policymaking. In fact, in a recent global survey of more than 3,000 scholars, 90 percent of them said, "There should be a larger number of links between academic and policy communities," and 76 percent of them said the gap between academia and policymaking has not improved in the last twenty to thirty years (Maliniak et al. 2012, pp. 67–69). Similarly, heads of governmental agencies in the past, such as Secretary of Defense Robert Gates, have said that governments should more frequently use the "untapped resources outside of government—resources

like those our universities can offer" (Gates 2008). In fact, a 2014 survey of policymakers showed that only 4.7 percent of government officials thought academics should not be involved in the policymaking process (Avey and Desch 2014).

Implications

If the usage of evidence-based reports increases in the public policymaking process, then the effectiveness of public policies will also increase (Banks 2009). Higher usage of EBRs in policymaking will improve policy results for several reasons. First, governments can easily identify which projects are and are not producing the outcomes they hoped for when they initiated the policies. Next, governments can empirically measure the exact impact of each policy and thus create a precise cost-benefit analysis of each program enacted. Finally, governments can perpetuate the cycle of successful policymaking by hiring officials who express more interest in EBRs and use them more frequently. If governments can pinpoint the officials most likely to use evidence-based reports, then they can focus their efforts on including those officials in policymaking. By hiring more officials who frequently use EBRs, governments can enact the most statistically successful and result-oriented public policy options (Avey and Desch 2014).

By measuring officials' knowledge and abilities, governments can identify which characteristics correspond with officials possessing sufficient knowledge and willingness to implement successful, result-driven policies. However, little research has been published on the topic of which characteristics cause individual officials to use EBRs more often. Therefore, by using survey data from three hundred government officials in developing countries on three continents, this investigation offers a unique insight into the effects that specific characteristics of officials have on an individual's interest in and reported use of EBRs, such as years of education, university major, gender, age, and income. With the knowledge about which officials are most likely to use EBRs, governments can further train those officials on such reports and put them in policymaking positions, thus increasing the number of policies based on reliable program evaluations.

Aside from personal characteristics, the survey also explores the effects of more structural procedures imposed by the government itself. The first structural procedure included in the analysis deals with the hiring/screening processes. To quantify the effects of governmental screening processes, the survey asked respondents if they completed a civil service exam to get their current job, a common method many governments use to evaluate the competency of candidates before hiring them. The second structural measure assessed in the survey involved the monitoring process, or the degree to which management promoted the use of EBRs. Two other important structural characteristics that could influence officials' use of or interest in EBRs include years worked in the government and workplace autonomy.

Overall, I found that female gender, higher income, post-graduate education, greater workplace autonomy, receiving a technical science degree, and completing a civil service exam lead to increased reported use of or interest in EBRs. Additionally, I

found that more years worked in the government and holding traditional family and religious values lead to decreased reported use of or interest in EBRs. Age had a quadratic relationship with reported use of EBRs: the youngest and oldest people reported using more EBRs the most, while middle-aged people reported using them the least. Management emphasis on EBRs appeared to increase employee interest, yet it did nothing to increase their use.

Theoretical Framework

Structural Characteristics and Procedures

CIVIL SERVICE EXAM

Both screening and monitoring methods exist to better control agents' or acting government officials' behavior and align their final policy outcomes with the outcomes desired by the principals (the people to whom the agents are accountable). Screening methods involve pre-hiring interviews, evaluations, and selection processes designed to ensure hiring the most qualified applicants. On the other hand, monitoring mechanisms focus on repeated evaluation, auditing, or supervision of agents after the hiring process to ensure they keep performing efficiently and are complying with what the principals desire. Kauppi and Raaj examine the effectiveness between the two mechanisms of controlling agent behavior and find that the pre-hiring controls of training and screening much more effectively reduced noncompliance in agents than the post-hiring mechanism of monitoring (2014). If civil service exams successfully screen employees and determine the most competent individuals for policymaking positions, and if all the most competent candidates understand how EBRs can improve policy outcomes, then I would expect to find that government employees who took civil service exams before being hired report higher interest in and use of EBRs.

MANAGEMENT EMPHASIS ON EVIDENCE-BASED REPORTS

Improper delegation to agents who lack proper knowledge in public policymaking leads to poor policy outcomes (Lupia 2000), but principals can and should attempt to prevent unknowledgeable agents from entering the policymaking arena through controls, such as screening before hiring them (Ibid. 2003). On the other hand, the post-hiring control of monitoring government workers' performance also provides a valuable tool for aligning the interests of the principals and the agents. In their study regarding public sector monitoring techniques, Sanders, Wright, and Horn show that teacher scores on performance assessments in Tennessee significantly predict student performance (1997). Just as teacher performance assessments can reliably predict student performance, governmental statistical evaluations in the form of EBRs can reliably predict policy outcomes.

Although the studies in Tennessee provide valuable insights into screening and monitoring processes, they also reveal some limitations in those procedures. First, although screening and monitoring of teachers can lead to hiring more effective employees, governments need to evaluate the effectiveness of their screening and monitoring

methods to find which assessments effectively hire teachers who will produce the best outcomes (Buddin and Zamarro 2009). Second, if school directors understand the basics of EBRs, they can use them to find additional effective methods of improving student performance by finding better teacher screening policies or classroom instruction programs (Sanders et al. 1997), while ignoring ineffective and costly methods of improving student performance, such as reducing classroom size (Hoxby 2000). To achieve the best policy results, programs must stem from previously implemented and tested policies proven to promote success. As shown from panel data in Latin America from 1974 to 2003, simply increasing government spending and creating extra government programs will not increase the welfare of the people (Lizardo and Mollick 2013). Increasing the quantity of evidence-backed policies will.

Officials still do not use EBRs often enough: "Although most practitioners claim to support the use of evidence relevant to their roles, their use of the best available evidence is patchy" (Head 2016, p. 471). Head goes as far as suggesting mandating the use of evidence-based research by governments, asserting that "the key task is to institutionalize rigorous processes for appraisal and evaluation [of policies in the public arena]" (2016, p. 476). To test the institutionalization of, or at least the emphasis of, management on using academic studies in public policymaking, our survey included a question asking the respondent: "What sources do you think your superiors want you to use in decision-making?" The respondents then chose from seven options, including the use of EBRs, and ordered them from most important to least important. Interestingly, 38.64 percent of respondents said their bosses would have ranked EBRs as most important on the list of seven decision-making factors, and 69.32 percent of them said their bosses would have put EBRs in the top three. In the end, if public servants listen to and follow their superiors, then I would expect to find higher reported use of and interest in EBRs by those who have workplaces where management promotes EBRs.

AUTONOMY IN THE WORKPLACE

To successfully match the interests of principals with the outcomes produced by agents, Fukuyama suggests "low-income countries [should] reduce bureaucratic autonomy while high-income ones [should] seek to increase it" (2013, p. 347). Low-income countries should rely on highly systematic bureaucracies instead of those with high autonomy, because fewer people obtain higher education in poorer countries, so those few with more education can mandate more effective policy processes such as requiring the use of EBRs to lower-level bureaucrats (Ibid.). According to the World Bank, Peru classifies as an upper-middle-income economy, having a GDP per capita of $6,046 US. India classifies as a lower-middle-income economy with a GDP per capita of $1,709 US, and Tanzania classifies as a low-income economy with a GDP per capita of $879 US (World Bank 2016). Therefore, based on Fukuyama's theory, if people from all three countries in the survey earn relatively low levels of income countrywide, which they do in India and Tanzania, then governments could create more effective policies through less autonomy in public servants. If part of hav-

ing less autonomy includes mandating EBR use, which would be the case if higher-ranked officials know the importance of EBRs in successful policymaking, I would expect to find that employees with lower levels of autonomy would use EBRs more often and would report higher levels of proficiency with them.

Personal Characteristics

EDUCATION (DEGREE AND LEVEL OF SCHOOLING COMPLETED)

Overall, policymakers believe that academic reports specifically from the disciplines of history, area studies, and economics help them the most in their policymaking processes. However, as the level of the agent's education starts to increase, they believe these studies will help policymaking less, especially in the fields of economics and political science. Additionally, policymakers overall gauge "sophisticated social science methods such as formal models, operations research, theoretical analysis, and quantitative analysis to be 'not very useful' or 'not useful at all'" (Avey and Desch 2014). However, policymakers did report that they highly value quantitative analysis of public opinion. Therefore, one policy implication to improve principal-agent compatibility could incorporate the implementation of frequent public opinion polls on public policy issues. The media frequently report on public opinion, but it usually does so in informal ways, so government-sponsored opinion polls on especially salient issues could increase agent accountability.

GENDER

In the scientific research field, some studies have shown that female researchers publish two less articles per five years than male researchers, controlling for all else (Prpic 2002). Structural factors in the scientific research field, such as position within their organization or network size, also appear to interact with women's success in publication, decreasing their success in relation to their male counterparts. However, other studies show that women's general level of lower productivity in the research field comes from a trend in women to specialize less than men do, rather than from institutional norms (Leahey 2006). Additionally, less women working exacerbate possible institutional problems that inhibit women from entering the research and innovation field and majoring in science, technology, engineering, and math (STEM). Although women constitute almost half of the work force and more than half of college graduates today, men still hold more than 75 percent of all STEM-related jobs (Beede et al. 2011). Therefore, if experience in the STEM field leads to more interest in EBRs and women on average go into STEM degrees and careers less often, I would expect women overall to express less interest in EBRs than men.

Because women major in STEM fields much less often than men do and STEM degrees use empirical evidence more often than other degrees, women might be less likely to use and show interest in EBRs. On the other hand, women already in the STEM field may show more interest in EBRs than men in STEM. Indeed, women tend to trust others more and look out for the common good more than

men do (Dollar and Gatti 2001). Additionally, governments with higher percentages of women in lawmaking bodies house less corruption (Dollar and Gatti 2001). Because corruption acts as an alternative source for policies, if a woman is making a policy decision, she is less likely to use corruption as an alternative to EBR findings. Therefore, if EBRs help to highlight corrupt public policies by showing their ineffectiveness and if women are more trustworthy and engage in corruption less than men, then women will be more likely to report use of and interest in EBRs than men.

AGE AND INCOME

Income works as a powerful motivating tool as well as a reliable reward system for diligence in the workplace. In the first half of careers, studies show that wage increases accurately reflect productivity gains by employees (Cardoso et al. 2011). Although worker productivity tends to decline or slow down as age increases, pay increases also slow down at an even faster rate. Therefore, I expect to see a positive relationship between income and reported use of and interest in evidence-based reports if more experienced officials know more about EBRs and if more experienced officials also receive higher wages.

As people grow older, they increase their work productivity and develop greater capacities to effectively fulfill their duties at work. However, at a certain point in their careers, their productivity begins to level off or even decline (Ibid. 2011). Their productivity may decline as their learning capacities decrease with old age or they cannot finish tasks as quickly. Interestingly, despite job performance decreasing later in life regarding new skills and intensive problem solving, evidence shows that older employees still maintain high productivity working where they have had extensive experience (Skirbekk 2004). Because governmental use of EBRs involving thorough statistical analysis has only recently received much attention, and this new skillset requires somewhat in-depth training, I hypothesize to see a quadratic relation with age and reported use of EBRs as described in Cardoso's analysis (2011). As people gain experience at the beginning of their careers, their likelihood of using such reports will increase. If the oldest people did not receive training on EBRs earlier in their careers and have smaller desires and capacities to learn innovative methods of research, such as those required to understand EBRs, and the youngest officials have yet to develop those skills, then I would expect to see the most use of and interest in EBRs among middle-aged government officials.

TRADITIONAL FAMILY AND RELIGIOUS VALUES

People perform their jobs with more productivity when they meet their physical, emotional, and even spiritual needs. In some cases, companies around the world are attempting to create workplace cultures that more openly accept spirituality and promote spiritual discussions. Empirical evidence from studies of these companies shows that increased spirituality in the workplace correlates with increased happiness and increased performance (Garcia-Zamor and Jean-Claude 2003, pp. 361–62). However, I do not specu-

late that holding more traditional values equates to both increased happiness and performance nor that increased job performance necessarily means increased use of EBRs.

Indeed, other studies have also shown that an increase in economic development on a macro-scale connects to an overall decrease in religiosity and an increase in more reason-based values instead (Inglehart and Baker 2000, p. 19). Interestingly, on a more individual level, national survey data shows this reflection of a stronger adherence to reason-based values and scientific knowledge among the less religious; people with weaker religious beliefs tend to support investment in groundbreaking medical technology more than people with stronger religious beliefs (Brossard et al. 2008). Therefore, if people with weaker religious beliefs tend to support reason and the quest for hard evidence over personal beliefs, and if the use of EBRs stems from the desire to believe in reason and hard evidence in place of personal beliefs, I hypothesize that people with less traditional or religious convictions will express more interest in and report more use of EBRs.

Methods

To evaluate the knowledge and skills of government officials in developing countries, a large team of three BYU faculty, nineteen BYU students, and nineteen local researchers administered a survey in Peru, India, and Tanzania. We posed dozens of questions about governance and demographics with multiple queries about officials' experience with and interest in evidence-based reports. The research team also trained more than six hundred government officials to use EBRs through a web site we developed, which contains more than four hundred such reports; each report graphs key findings from public policies on health, education, infrastructure, etc.

An initial summary of our survey data revealed that 20 percent of officials had never used reports in their policymaking decisions, and an additional 34 percent of all officials surveyed had not used EBRs in their policymaking processes during the last six months. Furthermore, despite only 8 percent of them saying they use EBRs in every policy decision they make, 92 percent of them said they believed that using EBRs would improve the quality of their work. This simple evidence from the survey confirms the literature that few government officials use EBRs despite the proven positive impact of using such studies and the acknowledgment that they would improve policymaking decisions (Slavin 2008). These initial findings highlight the importance of identifying the common characteristics of the few officials who do use EBRs so that governments can find more employees like them and focus their efforts on involving them in the policy making process.

After gathering and cleaning the survey data, robust statistical analysis revealed the final statistical relations stated in the introduction. However, some of the relations found from the survey differed from the original hypotheses and the literature as stated above in the theoretical framework section. For example, the literature suggested that females might use EBRs less than males, but in the survey, they reported to use more EBRs than males. The literature on autonomy suggested that lower autonomy for employees in poor

Figure 1: Importance of EBRs in Government
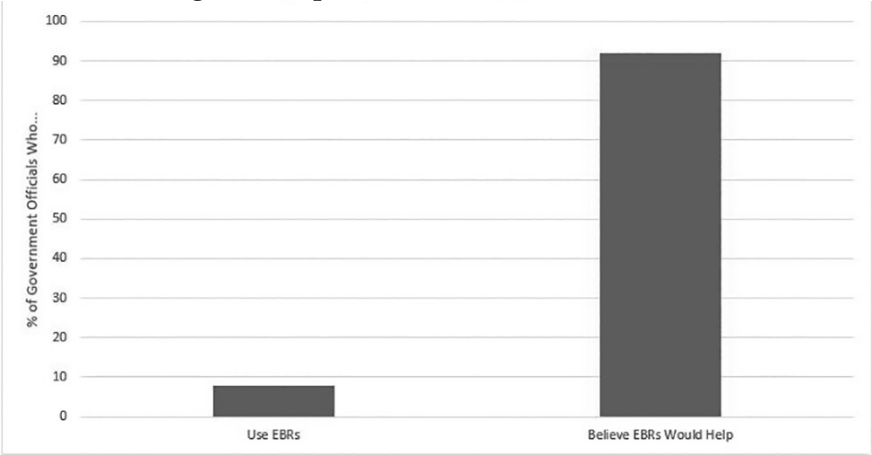
or developing countries might lead to better productivity and outcomes. Yet, in the survey, officials with higher autonomy reported more use of EBRs than employees with low autonomy did. Previous surveys of government officials showed that those with more education expressed less interest in academic studies of public policy, but our survey showed the exact opposite: officials with post-graduate degrees were more likely to report using EBRs than those without post-graduate degrees would. Additionally, age and years worked in the government seemed to have reversed relationships with interest in and reported use of EBRs from those suggested by literature about workplace experience and proficiency.

The independent variables discussed above in the theory section all come from survey responses. Equally, the dependent variables come directly from the survey: eleven of them are exact survey questions and four others are separate indexes of different combinations of those questions. The dependent variables about EBRs include questions such as the following: How often do you consult academic reports when making policy decisions? When was the last time you used an academic report in a policy decision? How interested are you in them? Do you think they help you make better policy decisions? (To reference all the questions used as dependent variables, please reference survey questions 17–27 in the Appendix. All questions in the Appendix remain precisely as posed in the survey given to the individual officials.)

By creating several indexes through the combination of up to ten survey questions in some cases, these indexes should more accurately evaluate the relations of specific independent variables with the use of and interest in EBRs. Factor analysis showed significant correlations between multiple survey questions that theoretically have a common theme and justified the creation of an index; the exact results from the factor analysis are in the Appendix at the end of the paper. To create each of the indexes, the responses to each survey question included in a specific index were transformed

into standard deviations so that each question would have equal weight in the index. After all the questions were transformed into standard deviations, they were combined into the indexes and collapsed to a smaller scale so that, once again, the indexes themselves would be in the form of standard deviations. Standardizing the indexes allows for analogous comparisons across indexes and simple magnitude tests for each independent variable because the independent variables' impacts on reported use of or interest in EBRs is given in terms of standard deviations.

In the results presented in this paper and elaborated on in the Appendix, 95 percent confidence determines statistical significance, meaning an individual characteristic or monitoring / screening process has a meaningful relationship with the official's likelihood of using EBRs. The tables attached below show all variables with their coefficients and standard deviations, but only the variables with 95 percent significance are considered to have a compelling relation with official interest in or reported use of EBRs. For each of the indexes, higher scores mean higher reported use of or interest in EBRs while lower scores mean lower reported use of or interest in EBRs. The compilations of the various indexes are as follows:

1) Master Index: The questions used to compile the master index pertain to interest in and reported use of EBRs. The theoretical reasoning for creating a master index comes from the idea that, although the statistical results from one individual dependent variable may not show strong enough evidence to result in statistical significance, many questions combined into an index gauging a person's interest and use may. Factor analysis of multiple variables with high correlations between each other also reveals an underlying common factor that an index can capture but a written survey question cannot. The index includes questions 17–25. Mean: 0.00; Range: -2.59 to 1.07; Standard deviation: 0.60.

2) Importance index: The questions used to compile the importance index reveal relevance of EBRs in the workplace. The theoretical reasoning behind an importance index stems from the idea that certain factors, such as a younger staff or a civil service exam requirement, may help to improve the overall perception of using new research methods, such as EBRs. In comparison to the other indexes that center more on the individual employee, the importance index takes a more general focus on the workplace attitude toward EBRs. The importance index includes questions 17 and 25–27, which ask about survey participants' personal attitudes toward EBRs and their perceptions of their co-workers' and bosses' views as well. Mean: 0.03; Range: -1.86 to .97; Standard deviation: 0.79.

3) Interest Index: The questions used to compile the interest index include only those regarding interest in and perceived importance of EBRs. The interest index provides a slightly different focus than the previously mentioned importance index, focusing on the perception of the individual rather than the entire workplace. Although the general workplace feeling toward EBRs as measured in the importance index gauges valuable overall views on EBRs, an individual's perception of

EBRs likely holds more influence in whether that individual will use them. Also, focusing on the individual, more exclusively in the interest and use indexes, more accurately measures the relations between characteristics specific to the individual and the official's interest in and use of EBRs. The interest index consists of questions 21, 23, and 24, which ask about overall interest in and perceived usefulness of EBRs. Mean: 0.00; Range: -3.83 to 0.66; Standard deviation: 0.86.

4) Use Index: The questions used to compile the use index include only those regarding reported use of EBRs. The use index provides a unique standpoint from the other indexes in that it focuses only on questions regarding frequency of physically using EBRs rather than simply believing in their importance or usefulness. In the end, academics make EBRs with the goal of affecting public policy, and those reports cannot influence policy unless officials use them. Therefore, the use index is a crucial indicator of which personal and systemic factors can lead to better public policy. The use index includes questions 18–20, which ask about the frequency the officials' report using EBRs. Therefore, although the measurement of EBR use relies on survey questions asking for reported use, many other high standard surveys use similar methods for measuring behaviors. This measurement of use does introduce some limitations because of systematically high reporting of socially desirable behaviors like using EBRs, but more accurate behavior measures would have significantly increased the survey's time and cost parameters. Mean: 0.00; Range: -1.59 to 1.62; Standard deviation: 0.83.

Aside from the indexes, each survey question acts as a dependent variable by itself. For survey questions with a mostly continuous set of responses, I simply used ordinary least squares regressions, checking for both interaction terms and quadratic relationships with other variables, including age and years worked in the government. Some academics suggest that questions with ordinal responses require more in-depth models, such as the ordered probit model. However, a simple use of OLS regressions can still create accurate predictions of significance for ordinal variables, especially when those variables have more than five possible ordered responses or are indexes created through the compilation of multiple survey questions. In fact, some scholars consider using ordinal variables as continuous variables a more powerful statistical approach, and in many cases, such an approach can reveal important relationships not shown in a model that keeps those variables as categorical (Pasta 2009).

Additionally, linear OLS models do not require perfectly even spacing between ordinal variables. Usually one-unit changes in continuous variables do not equal exact linear changes in another variable, because relationships are never perfectly linear, most "results are remarkably insensitive to the spacing of an ordinal variable, except in the most extreme cases" (Ibid.). Therefore, the final model for each dependent variable involves OLS regressions using country fixed effects, clustering the standard deviations by country as well.

The final model of the master index includes eighteen independent variables, and various other models also appear in the Appendix, some having fewer control variables and some having more interaction and quadratic effects included as well. Recall that to form the different indexes, survey responses were measured first in terms of standard deviations so that a compilation of questions with differing scales would not skew the indexes. Then the results from the different questions were added to the index with equal weight for each question, and the final number divided by the number of questions in each index. Therefore, each coefficient in the master index, as well as the other three indexes reported, is in terms of how many standard deviations the index moves with a one-unit change in the independent variable.

Results

Structural Characteristics and Procedures

Appraising which government employees express the most interest in and report the most use of EBRs becomes more important as the chain of delegation grows longer. As the consumers of public policies and programs, citizens often do not vote directly on what types of policies they want implemented or for the bureaucrats who enact those policies. Instead, citizens vote for political candidates who then choose the bureaucrats

Table 1

Independent Variables	Master Index	Use Index	Interest Index
Female	Pos	Pos	
Age	Parabolic	Parabolic	
Years in Government		Neg	
Post Graduate	Pos		
Civil Service Civil Service Exam	Pos	Pos	
Traditional Values (1-4)	Neg	Neg	
Autonomy (1-4)	Pos		Pos
Perceived Importance		Pos	
Witness of Corruption (1-4) Witness	Pos	Pos	Pos
Importance of Studies to Boss	Pos		Pos
Number Ministries Worked In	Pos	Pos	
Income (1-6)	Pos	Pos	Pos
Business Degree			
Technical Science Degree	Pos	Pos	
Stats/Math Degree		Neg	
Humanities Degree			
Other Degree (Mostly Communications)			
Peru (Countries Compared to India)	Pos	Neg	Pos
Tanzania			Pos
*Only recorded if significant at 95% level			

Showing positive and negative relations between the IV and DV. Only relations with at least 95 percent significance after incorporating all the control variables shown below are reported.

Table 2

VARIABLES	(1) Master Index	(2) Interest Index	(3) Use Index
Female	0.128** (0.0646)	0.172* (0.0998)	0.243** (0.0970)
Age	-0.0445** (0.0205)	-0.00230 (0.00595)	-0.0713** (0.0308)
Age Squared	0.000464** (0.000220)		0.000787** (0.000330)
Years Worked in Govt.	-0.00313 (0.00434)	0.00392 (0.00667)	-0.0160** (0.00651)
Post Graduate	0.177** (0.0834)	0.102 (0.128)	0.208* (0.125)
Civil Service Exam	0.152** (0.0607)	0.0499 (0.0939)	0.272*** (0.0911)
Importance to Boss	0.0686*** (0.0148)	0.0676*** (0.0229)	-0.000810 (0.0222)
Traditional Values	-0.0848** (0.0394)	-0.0490 (0.0604)	-0.144** (0.0591)
Autonomy	0.145*** (0.0377)	0.138** (0.0582)	0.172*** (0.0565)
Perceived Importance	0.00673 (0.0118)	0.0141 (0.0183)	-0.00148 (0.0177)
Witness of Corruption	0.0937*** (0.0303)	0.100** (0.0468)	0.126*** (0.0454)
# of Ministries Worked in	0.0372** (0.0159)	0.0225 (0.0245)	0.0679*** (0.0239)
Income	0.108*** (0.0263)	0.171*** (0.0404)	0.0959** (0.0394)
Business Degree	-0.0994	-0.0810	-0.105
(Compared to Poli Sci)	(0.0910)	(0.141)	(0.137)
Tech. Science Degree	0.221*** (0.0676)	0.138 (0.104)	0.277*** (0.102)
Stats/Math Degree	-0.234* (0.131)	-0.156 (0.202)	-0.512*** (0.197)
Humanities Degree	-0.212 (0.155)	-0.225 (0.239)	-0.172 (0.232)
Other Degree	-0.247* (0.143)	-0.386* (0.222)	-0.208 (0.215)
Constant	-0.556 (0.465)	-1.913*** (0.406)	0.489 (0.698)
Observations	280	280	280
R-squared	0.493	0.412	0.296

Pictured above are the final fixed effects models clustering by country for each of the three main indexes with all controls included. Alternate models for each individual dependent variable are available in a separate excel spreadsheet. Remember that all the coefficients are in terms of standard deviations.

that realize policy decisions. In this way, the general population of the country acts as a principal in telling the agents what they want out of a public policy. However, as agents delegate policymaking power to more people down the chain of command—officials more enveloped in the bureaucracy—they become less accountable to the people. Therefore, screening and monitoring processes allow governments to evaluate which employees are most likely to enact efficient policies closely aligned with the constituents' interests and thus help reduce waste and misuse of power in the government (Lupia 2003). Table 1 shows a color-coded grid highlighting which independent variables, such as personal characteristics or screening/monitoring processes, have a statistically significant relation with the master index, use index, and interest index. After using a fixed-effects-model regression analysis using all the control variables shown in Table 2 (on previous page), the final models run for each of the dependent variables, each one clustering by country-fixed effects.

CIVIL SERVICE EXAMS (QUESTION 7)

In the master index that combines ten survey questions regarding interest in and use of EBRs, individuals who took a civil service exam prior to hiring are statistically .31 standard deviations higher in the master index than government officials who did not take the civil service exam. This result, along with all other results reported in the paper, exhib-

Figure 2
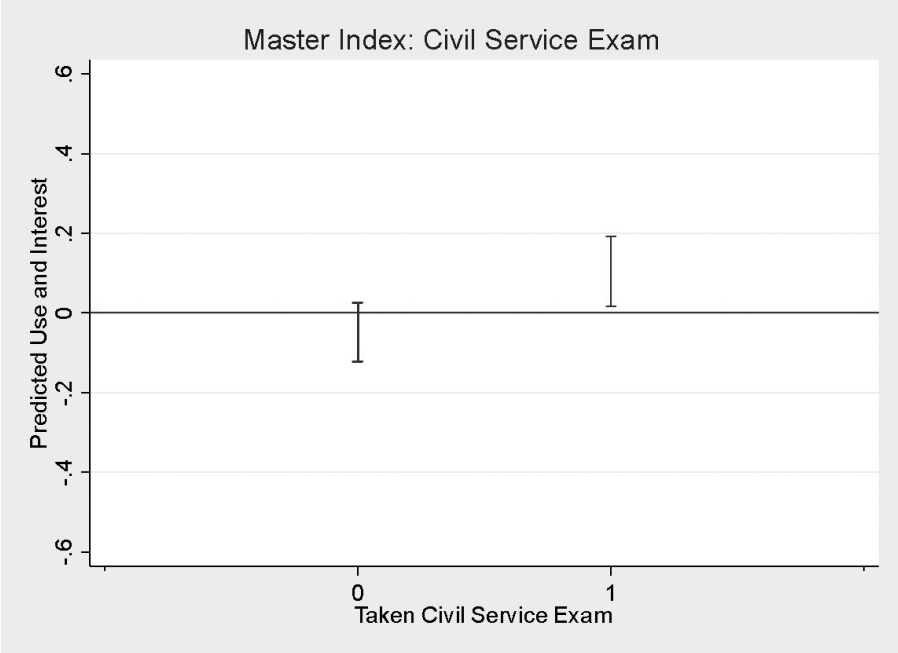
Civil service exam takers report more use of and express more interest in EBRs than non-civil service exam takers.

its significance at the 95 percent confidence level. The use index also shows a positive relation with a coefficient equal to 0.32 standard deviations, as do four other individual dependent variables in the survey, reinforcing its significance in leading to higher interest in and reported use of EBRs. These results show the importance of implementing screening practices in hiring; governments should use screening processes to more effectively eliminate less knowledgeable officials and should promote more effective policymaking by having people who will use evidence-based research with higher frequency.

MANAGEMENT EMPHASIS ON EVIDENCE-BASED REPORTS (QUESTION 27)

I focus on the use index and the interest index to measure the effectiveness of superiors placing emphasis on EBRs in the workplace, because the two indexes show different outcomes. The importance that superiors place on using evidence-based reports makes a notable difference in the interest of their employees in using EBRs. If superiors place more emphasis on such reports, then their employees report greater interest in them, think they are useful, and believe they will help them work better. In fact, an employee who says their superior rates evidence-based reports as number 1/7 of important factors in public policy decision making ranks 0.47 standard deviations higher on the interest index than someone who says their superior rates EBRs as 7/7 in importance. According to these results, repeated focus of bosses on using evidence-based reports positively influences the employees' views on using such reports. Aside from the interest index, management stress of EBRs also positively correlated with six other dependent variables. Regrettably, the survey results here do not necessarily mean causation because perhaps people who show greater interest in EBRs simply assume their superiors want them to value EBRs as well.

Unfortunately, that same positive impact of a superiors' influence on interest does not translate to increased use or knowledge in any of the four knowledge and use questions, and no significant relation appears in the use index either. These results suggest two important considerations for implementing future monitoring policy. First, superiors exert considerable power over what methods their employees identify as most effective in helping them to do their jobs. However, it also highlights the flaws in some monitoring techniques. Although government officials acknowledge greater importance in using EBRs as their superiors stress their value, officials still do not use them more often. Therefore, governments should research more effective monitoring techniques or oversight procedures besides simple management emphasis to ensure that governments use EBRs with greater frequency. Perhaps mandated use of EBRs for policy decisions, such as Head suggests, could lead to more frequent use (2016, p. 476). Additionally, this odd relationship between management emphasis on EBRs and increased interest without impacting use suggests the need for further research on how to ensure that government officials keep their actions more in line with their superiors' priorities.

NUMBER OF YEARS WORKED IN THE GOVERNMENT (QUESTION 12)

One concerning find was that, as bureaucrats work more years in the govern-

**Figure 3**

Interest Index: Management Emphasis on EBRs
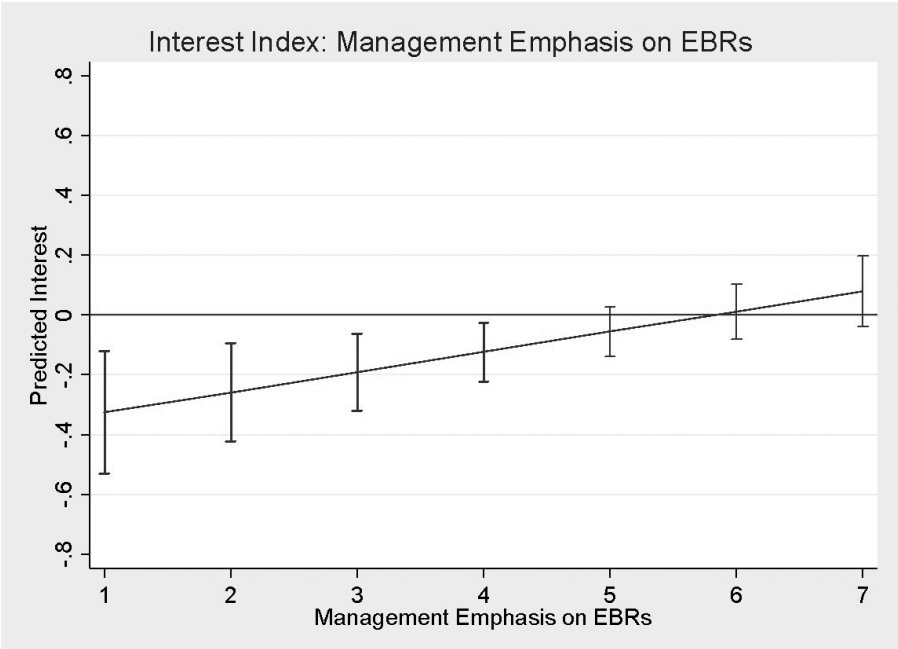
As bosses rank EBRs as more important, employees show more interest in using them. Note that the ranking goes from 1 being least important to 7 being most important.

ment, their interest in and reported use of EBRs decrease; three survey questions regarding interest and use, as well as the overall use index, confirm this relationship. Every ten years spent working in the government results in a 0.19 standard deviation decrease in the use index. That decrease in use may not seem significant initially. However, when considering the average number of years worked in a lifetime, usually somewhere around forty (Brandon 2014), the effect of years worked in the government on the likelihood of using EBRs becomes quickly apparent.

The negative relation revealed between years worked in the government and use of EBRs highlights the importance of government monitoring as discussed earlier. Not only did the post-hiring tool of management emphasis on EBRs show no effect in increasing EBR use, but even the years spent working in the government may decrease their use. These combined results accentuate the prevalent problem of continually monitoring and improving agent efficacy. Unfortunately, it appears that sufficient post-hiring methods do not exist in developing countries, at least using the post-hiring proxy measures of management emphasis on EBRs and years worked in the government, once again confirming what Kauppi and Raaj stated in their article about the lack of successful post-hiring mechanisms in governments (2014). If such mechanisms did exist and successfully increased productivity and innovation, then the survey should show that more years worked in the government leads to an increase in knowledge of, interest in, and use of EBRs. Policy

Figure 4

Use Index: Years Worked in the Government
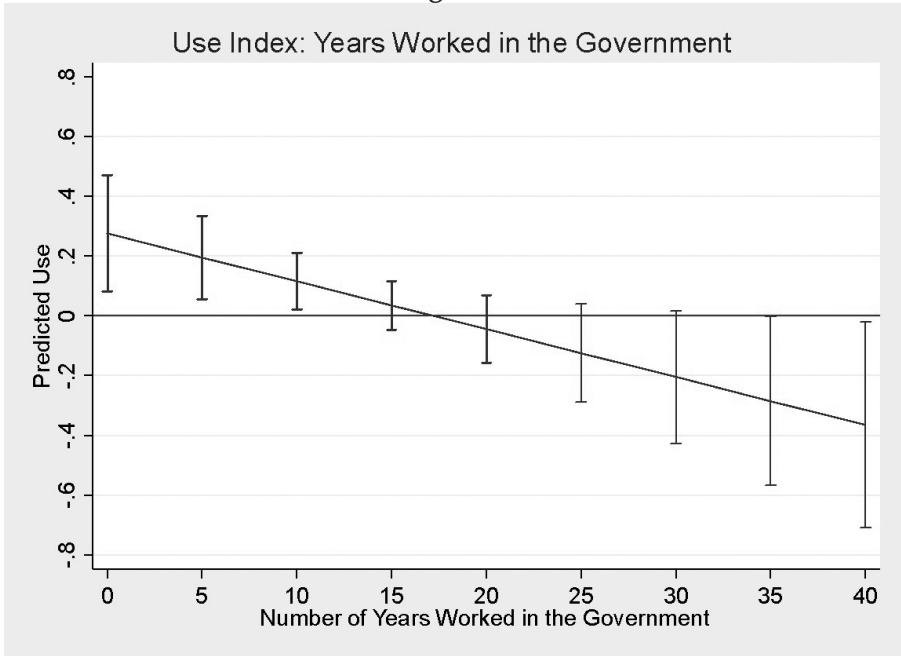
Officials who have worked more years in the government report to use EBRs less than newly employed officials.

implications to improve the use of EBRs as officials stay in the government longer include EBR training and possible bonuses or promotions for those using them with greater frequency in policymaking decisions.

AUTONOMY IN THE WORKPLACE (QUESTION 15)

In Fukuyama's earlier stated theory, he predicted that lower autonomy in government officials would lead to more effective bureaucracies and policies in less-developed countries. Using the evidence from earlier in the paper that higher EBR use leads to more effective policy outcomes, the survey data contradicts Fukuyama's theory, revealing that autonomy has a statistically significant positive relation at the 95 percent level with reported use of and interest in EBRs. Six individual dependent variables along with the interest index and the master index confirm this positive relation. A move from completely subservient to almost completely autonomous on the autonomy scale results in a positive shift of .45 standard deviations in the master index and a positive shift of .48 in the interest index.

These results could signify two patterns. First, perhaps there could be reverse causation. In fact, people most educated in and interested in using EBRs also rank higher up in the governmental hierarchy and, therefore, have more autonomy. Second, more workplace autonomy leads to trying new and different methods of research for public policies. As agents receive more freedom to devise their own public policy plans, they expand

their vision and use newer sources of information, such as EBRs. The idea of bureaucratic entrepreneurship and innovation when given more freedom receives direct support from evidence in the private sector of the same innovative principle, but just as in the private sector, bureaucratic entrepreneurship does not come immediately after receiving greater autonomy (Ibid., p. 34). For bureaucratic entrepreneurship to take place, agents must have high levels of autonomy for longer stretches of time. Additionally, agents must first develop a trusted capacity before receiving bureaucratic autonomy. Otherwise, with little knowledge and drive, the newly granted autonomy could lead to decreased productivity (Carpenter 2001, p. 14). Therefore, greater autonomy in agents may not lead to greater innovation and improved policy outcomes immediately, but improvements based on increased autonomy come from long-term investments in human capital leading to improved agent experience and abilities.

Figure 5
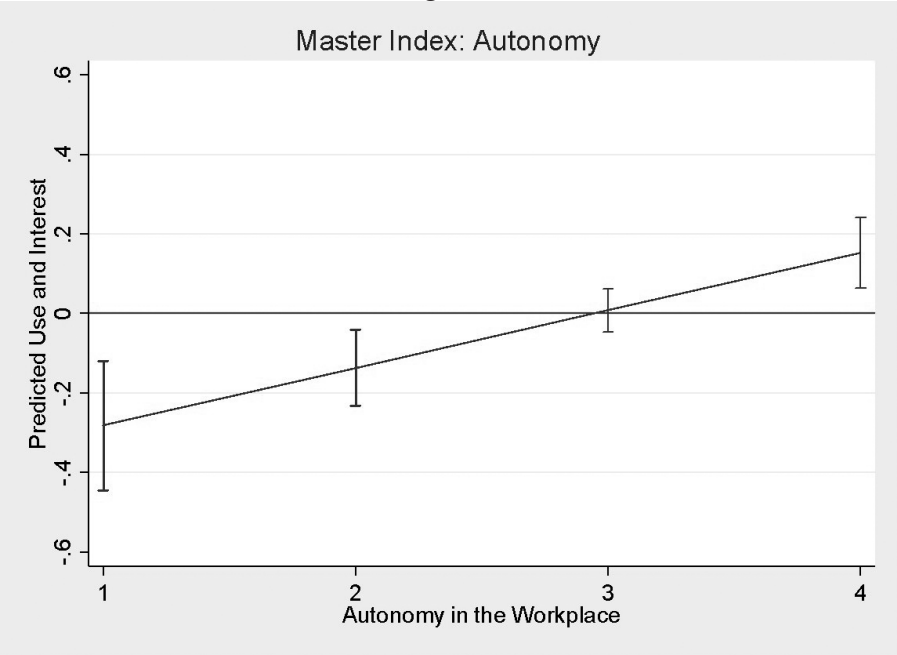
Officials with more autonomy report more use of and interest in EBRs. Note that on the autonomy scale, 1 means less autonomy and 4 means more autonomy.

Personal Characteristics

EDUCATION: DEGREE AND LEVEL OF SCHOOLING COMPLETED (QUESTIONS 3 AND 4)

In the survey data, the relations between collegiate discipline and interest/use of EBRs are compared to the political science discipline. Most degrees show no difference of interest in or use of evidence-based reports. However, officials with technical science

Figure 6

Master Index: Degree
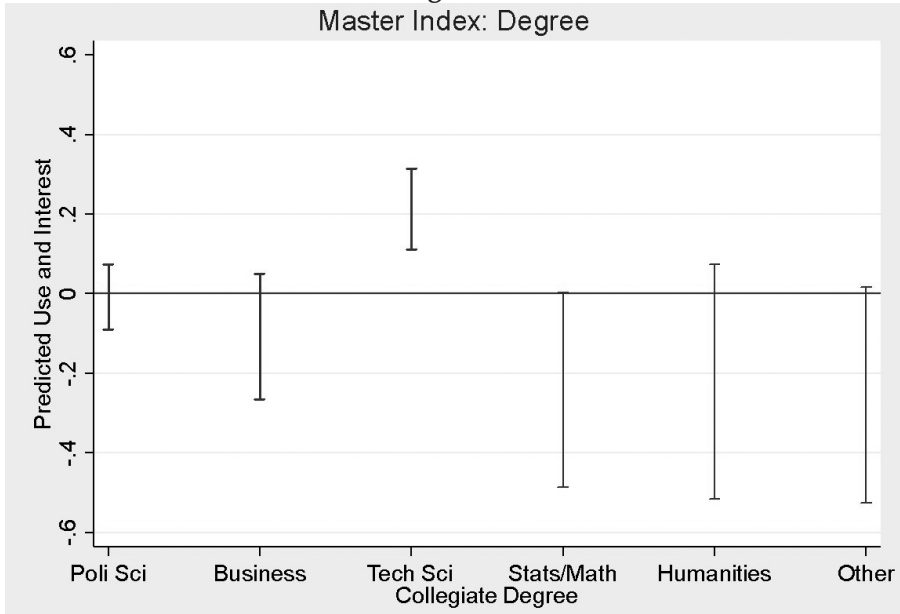
Officials who obtain technical science degrees report the most use of and interest in EBRs compared to all other degrees measured.

Figure 7

Master Index: Postgraduate Education
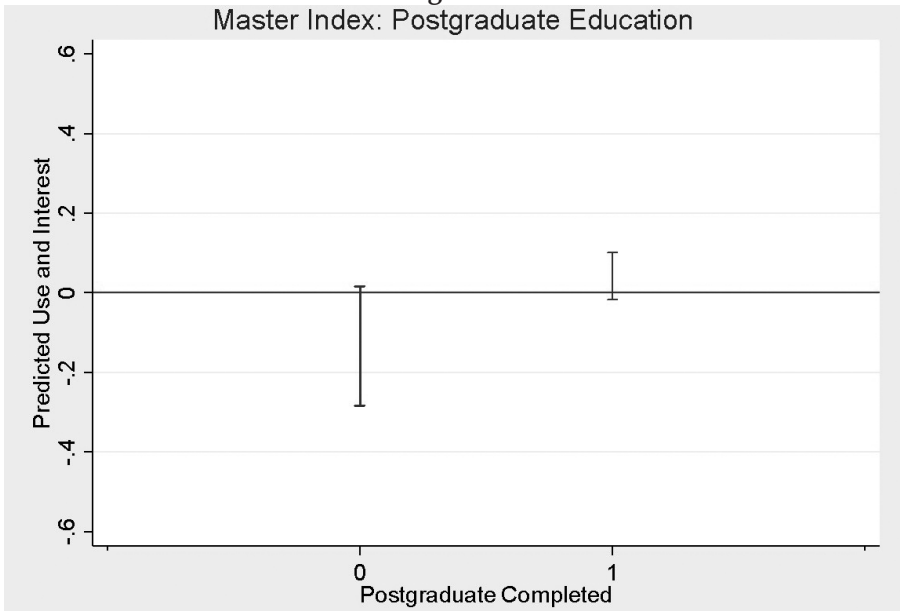
Postgraduates report more use of and interest in EBRs than non-postgraduates.

degrees do register higher on the use index, importance index, and master index with an increase in standard deviations of .33, .26, and .34, respectively. Because technical sciences use more hard evidence than other disciplines do, technical science graduates demonstrate more interest in reports based on empirical evidence. On the other hand, the master index shows that post-graduate education in general also shares a significant positive relation with interest in and reported use of EBRs, with an increase of .41 standard devia-

Figure 8
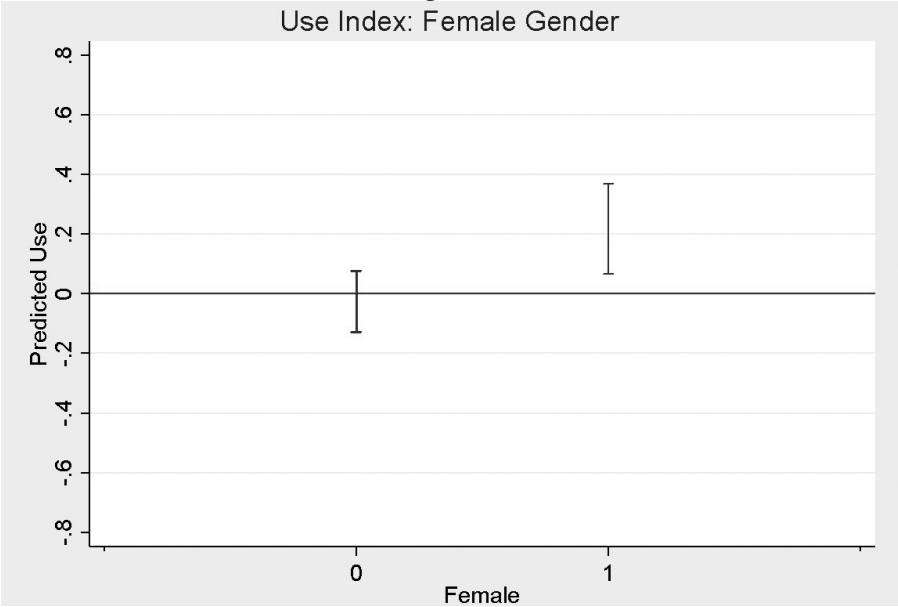
Females report higher usage of EBRs than males do.

tions compared to no post-graduate education. Officials with post-graduate degrees also report more interest in or use of EBRs as shown in four other dependent variables.

GENDER (QUESTION 1)

Contrary to what some background research suggested, survey data shows females are more likely to report seeing and using EBRs than their male counterparts. The use index shows an increase of .29 standard deviations in usage for females over males, and the master index shows an increase of .28 standard deviations for females over males.

INCOME (QUESTION 5)

Income by far appears to most strongly predict reported use of and interest in EBRs. The survey divides income levels into six categories. An official in the highest income bracket compared to an official in the lowest income bracket is .75 standard deviations higher in the master index and 1.0 standard deviation higher in the interest index. Income also has a significant positive relation in six other dependent variables.
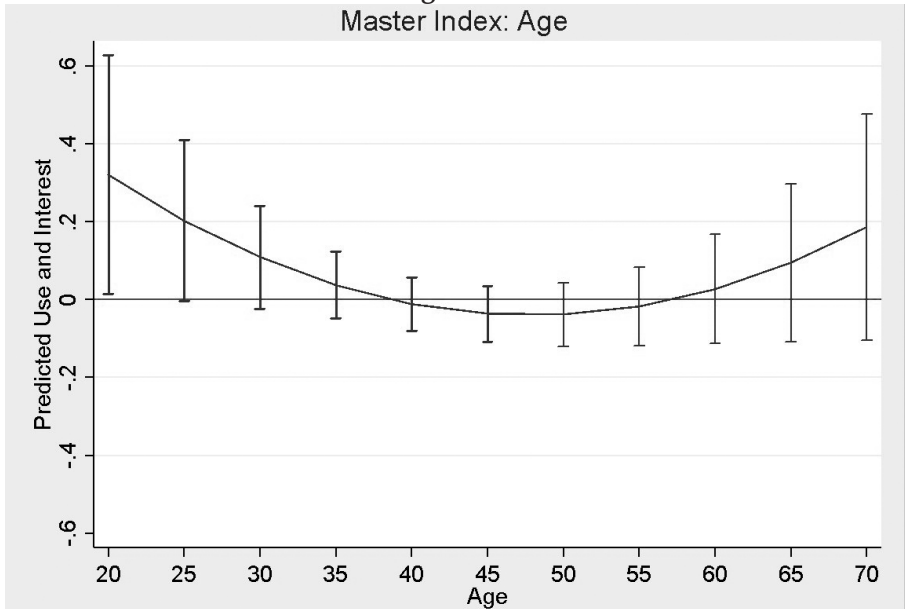
AGE (QUESTION 2)

Figure 9
Master Index: Income
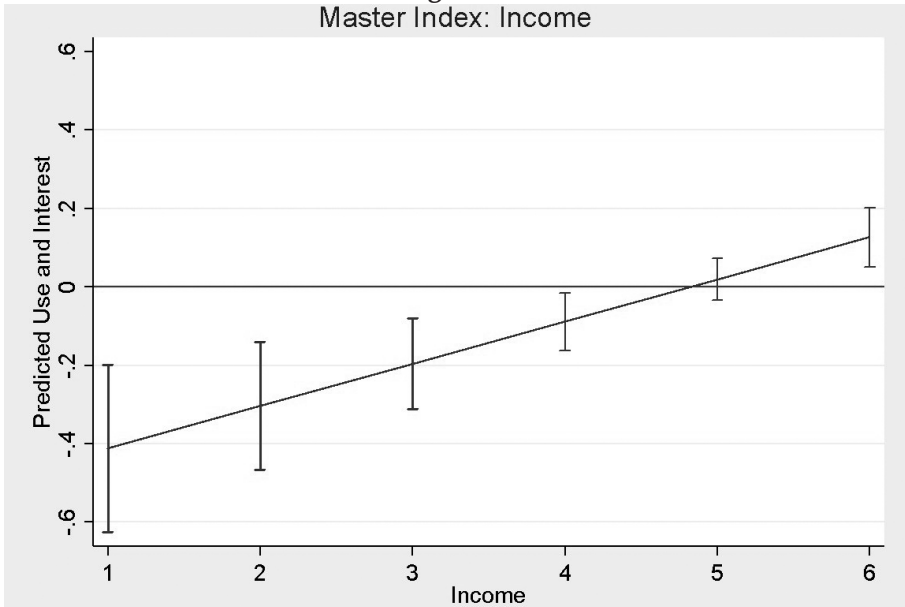
Officials with higher income levels report more interest in and use of EBRs.

Figure 10
## Master Index: Age

The youngest and oldest officials report higher interest in and use of EBRs while middle-aged officials report lower interest in and use of EBRs.

The survey results confirmed previous studies done by Cardoso that age maintains a quadratic relationship with productivity and learning (2011), yet the survey shows the opposite quadratic relation from that predicted by Cardoso. As officials grow older, they report using EBRs less and less. However, usually around the age of 45–50, they start using reports more often again. As shown in Figure 10 (on previous page), middle-aged officials report the least use of EBRs of any age group, and three other dependent variables confirm the same parabolic relation between age and reported use of EBRs.

TRADITIONAL FAMILY AND RELIGIOUS VALUES (QUESTION 6)

Just as the research suggests, higher levels of traditional religious values lead to lower levels of emphasis in logical, evidence-based research (Inglehart and Baker 2000, p. 19). In the survey, those who say they hold more traditional religious and family values report to use less and have less interest in EBRs. A person reporting to have the strongest traditional religious and family values compared to someone with the weakest values is

Figure 11
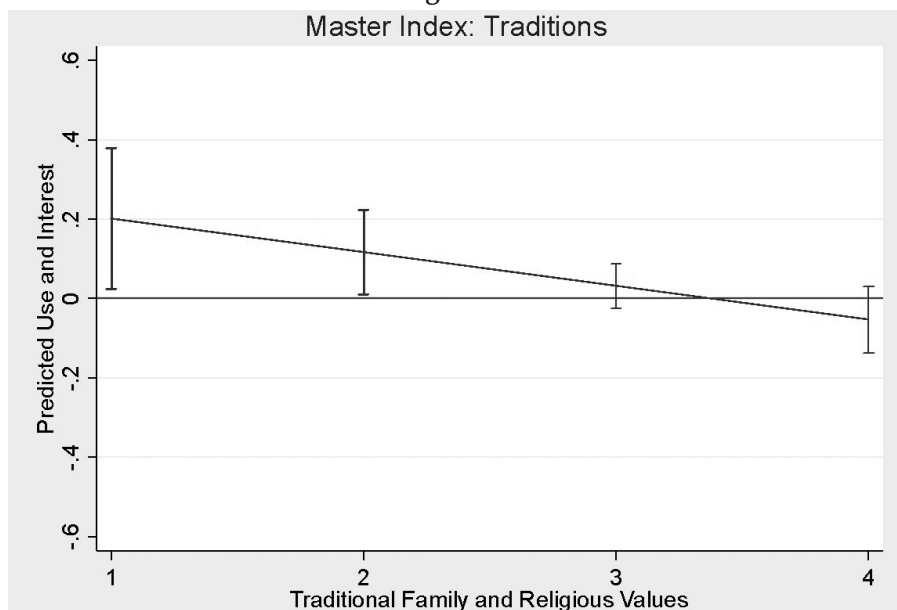
Officials who place more importance on traditional family and religious values report less interest in and use of EBRs. Note that on the traditions scale, 1 means less traditional and 4 means more traditional.

on average .55 standard deviations lower on the master index and .52 standard deviations lower on the use index. Overall, more traditional people also report lower use of/interest in EBRs in four other individual dependent variables.

Limitations, Case Selection, and Further Research

The survey used as the main evidence has several limitations. First, the survey gath-

ers quantitative data in a non-experimental manner. Therefore, all the statistically significant relations found do not necessarily mean causation. However, as stated in the theories section pulling from academic literature earlier, causal relations theoretically do exist even if not proven by the method of data collection. On the other hand, use of survey data adds a high level of external validity, or applicability across governments, especially in developing countries, because the survey responses come from government officials in three developing countries on three continents. Other contributing factors to high external validity include the completely randomized assignment of enumerators to administer the surveys and the wide distribution of demographics among those surveyed. Few studies of such magnitude have been conducted involving government officials from developing countries in the past, so even a sample size of three hundred provides significant new findings on what influences an individual official's interest in and use of EBRs.

More limitations to the study include any vagueness in the survey questions or dishonest responses that could skew the results and decrease the internal validity. Additionally, some of the questions in the survey asking about interest in EBRs use repetitive wording and may condition the participants to choose what they think the researchers want them to choose, once again skewing the results. Limitations from over-reporting a socially desirable behavior could skew the results of the survey if those who dishonestly answered as more likely to use EBRs all had shared characteristics that differed from those who reported being less likely to use EBRs (Zeglovits 2014, p. 225). Additionally, even if respondents did not lie to conceal a less socially desirable history of not using EBRs, they could simply misremember the proper frequency with which they used them and misreport due to memory recall problems (Krosnick 2002, p. 93). To possibly eliminate these reporting errors, we could have included a "don't know" option for all these questions on the survey, but because many studies have shown that such response options do not significantly increase reliability and will decrease sample size and, therefore, statistical strength (Ibid., p. 91), we decided to exclude a "don't know" option on the survey.

The responses to the survey may also have some selection bias. Because enumerators administered the survey only to government officials who had already accepted a visit to learn more about how they could use EBRs in policy making, the survey may include a disproportionate number of government officials who already use EBRs while ignoring those who have no interest. Despite the survey's limitations in internal validity, the high external validity of quantitative analysis justifies the use of statistical analysis. Because 280 officials completed the survey on three continents, the results of the survey should more accurately reflect the dispositions of government officials overall than if only several officials had received a more in-depth, qualitative interview. In the end, despite strong theories and robust statistical analysis, it is hard to infer causation, because observational data was extracted from surveys rather than from a randomized control trial.

Conclusions

Overall, I found that female gender, post-graduate education, receiving a technical science degree, higher income, greater workplace autonomy, and completing

a civil service exam increase reported use of or interest in EBRs. Additionally, I found that the number of years worked in the government and greater regard for traditional family and religious values decrease reported use of or interest in EBRs. Age had a quadratic relationship with use of EBRs: Middle-aged people reported using fewer EBRs to a certain extent, but the youngest and oldest people reported using more. Management emphasis on EBRs appeared to increase employee interest, yet it did nothing to increase employee use.

The survey evidence used in this report plainly outlines which personal and structural characteristics correlate with increased use of EBRs in developing countries today. Increased use of EBRs, according to both policymakers and academics, will objectively improve the outcomes of public policies. Governments are always trying to improve the effectiveness of their programs to better help their constituents. Because increased use of EBRs translates to improved policy outcomes and improved well-being of the impacted population, governments should use them with greater frequency and thereby improve their public policy decisions. Finally, by knowing which personal and structural characteristics correspond with higher use of EBRs, governments can design better hiring, monitoring, and training processes to ensure the inclusion of the officials most likely to use EBRs in their policy decisions.

APPENDIX

Regression tables

<div align="center">

Table 3

</div>

Use Index

Variable	Factor 1	Uniqueness
Question 19	0.7335	0.4619
Question 20	0.6596	0.565
Question 18	0.6844	0.5317

Interest Index

Variable	Factor 1	Uniqueness
Question 21	0.7019	0.5074
Question 23	0.7549	0.4302
Question 24	0.7725	0.4033

Importance Index

Variable	Factor 1	Uniqueness
Question 27	0.7233	0.4731
Question 26	0.6091	0.6198
Question 25	0.7537	0.4271
Question 17	0.7399	0.4568

Master Index

Variable	Factor 1	Uniqueness
Question 22	0.6227	0.5334
Question 19	0.5258	0.4432
Question 20	0.5754	0.5188
Question 18	0.4418	0.5212
Question 21	0.7268	0.4095
Question 23	0.7044	0.399
Question 24	0.6819	0.4228
Question 25	0.4349	0.5424
Question 17	0.4961	0.4045

For regression tables of alternate models, please request an extended excel file.

Factor Analysis

Survey Questions

Impact Evidence Survey

This research study is being conducted by Darren Hawkins, professor, Brigham Young University to determine how to best provide high-quality impact evidence to policy makers. You have been invited to participate because you are a government official in one of our focus countries. The study consists of 27 questions and will take approximately 10 minutes to complete. There are minimal risks for participation in this study. You may feel some discomfort in answering some of the sensitive questions on our survey. You may feel badly about not scoring better. Poor scores will never be revealed to others. The benefits of participating are learning more about what works in policy and what does not, which could be useful in your job and to recipients of government services. Involvement in this research project is voluntary. You may withdraw at any time without penalty or refuse to participate entirely. We will only report anonymous results from this study so you will not be identified in our research findings in any way. If you have questions regarding this study you may contact Darren Hawkins at (801) 422−5526 or dhawkins@byu.edu. If you have questions regarding your rights as a participant in research projects, you may contact: IRB Administrator, A-285 ASB, Brigham Young University, Provo, UT 84602, (801) 422-1461, irb@byu.edu. By advancing in this survey, you consent to participate.

Thank you for taking this survey! Your participation is completely anonymous and will help us learn more about how to provide helpful information to government officials.

Please enter the code provided by the research assistant

1. What is your gender?

  Male

  Female

2. What is your age?

3. What is the highest level of education you have completed?

    Primary school

    High school

    Vocational school

    College

    Graduate degree

4. What did you study for your highest degree?

    Social Sciences (Economics, Political Science, Sociology, Policy, Development, etc)

    Business

    Technical Science (Engineering, Agriculture, Chemistry, Biology, Physics, etc)

    Statistics or Math

    Humanities (Literature, Languages, Arts, Design, etc)

    Other: _____

5. Is your monthly income:

    20,000 Rupees or below

    Between 20,000 and 40,000 Rupees

    Between 40,000 and 60,000 Rupees

    Between 60,000 and 80,000 Rupees

    Between 80,000 and 100,000 Rupees

    Over 100,000 Rupees

6. To what extent do you agree with the following statement? "Tradition--customs handed down by one's religion or family--is important to me."

    Strongly Agree

    Somewhat Agree

    Somewhat Disagree

    Strongly Disagree

7. Did you take a civil service exam to get your current job?

    Yes

    No

8. Which factors help someone get the job that you have? Please count how many of these things matter in getting a job and tell us the number of relevant factors.

    a. Civil service exam score

    b. Political party

    c. Personal connections

    d. Corruption

    e. Prior experience and performance in other jobs

    f. Physical appearance

    g. Gender

    _____ How many of the above apply? (Provide just a count of the number that apply)

9. Have you ever seen evidence of corruption in the work you do?

    Yes, all the time

    Yes, sometimes

Yes, but rarely

No, never

10. Have you ever seen evidence of corruption in work others do in government jobs?

Yes, all the time

Yes, sometimes

Yes, but rarely

No, never

11. Please check all of the following that apply to your current position:

Policy or program formulation and design

Policy or program monitoring and evaluation

Policy or program implementation

Policy or program administration

Compliance

Accounting and budget

Human resources

Archives

Legal services

Equipment and facilities

Technical services

Other _____

12. How many years have you been employed by the government?

13. What sectors of the government have you worked in during the past 10 years?

Please choose all that apply

Agriculture

Commerce and Trade

Culture

Defense

Education

Energy

Environment and Natural Resources

Finance, Credit, Banking

Foreign Ministry

Gender and Family

Health

Housing

Industry

Justice

Labor

Macroeconomic Management

Public Safety

Social Welfare and Development

Sports

Transportation

Other _____

14. How many people do you believe your policies influence?

15. How much autonomy do you have in your job?

Almost full autonomy

Some autonomy

Little autonomy

No autonomy

16. What kind of decision-making process is used in your office?

One person makes the decision without the counsel of others

One person makes the decision, but takes counsel from others

A committee makes the decisions

Everyone has to agree with the decisions

Other _____

17. What source is most useful to you when making your policy decisions? Please order the following with 1 being the most useful and 7 being the least useful

_____ My intuition

_____ My experience

_____ My formal education

_____ Job-related conferences

_____ My coworkers

_____ Public opinion

_____ Reports, studies and other written material

18. International organizations and governments, often in partnership with academic researchers, are performing rigorous impact assessments of government programs and publishing reports with their findings. Have you seen such academic reports?

Yes, many

Yes, some

No, none

19. How often do you consult such academic reports when you make policy decisions?

Every time

Most times

Sometimes

Never

20. How recently have you used academic reports to make a policy decision?

In the past month

In the past six months

In the past year
In the past five years

More than five years
Never

21. How interested are you in learning from academic research?
    Very interested
    Somewhat interested
    Indifferent
    Somewhat disinterested
    Not interested at all

22. How well do you feel you understand academic research?
    Very well
    Somewhat well
    Not very well
    Not well at all

23. How helpful is academic research is in making policy?
    Very useful
    Somewhat useful
    Not very useful
    Not useful at all

24. Do you think using academic research makes your work better?
    Yes, definitely
    Yes, somewhat
    No, not really
    No, definitely not

25. What sources do you think are the most important for policy makers to use in decision-making? Please order the following with 1 being the most useful and 7 being the least useful
    _____ Their intuition
    _____ Their experience
    _____ Their formal education
    _____ Job-related conferences
    _____ Their coworkers
    _____ Public opinion
    _____ Reports, studies and other written material

26. What sources do you think your coworkers use in decision-making? Please order the following with 1 being the most useful and 7 being the least useful
    _____ Their intuition
    _____ Their experience
    _____ Their formal education
    _____ Job-related conferences

_____ Their coworkers

_____ Public opinion

_____ Reports, studies and other written material

27. What sources do you think your superiors want you to use in decision-making?

Please order the following with 1 being the most useful and 7 being the least useful

_____ My intuition

_____ My experience

_____ My formal education

_____ Job-related conferences

_____ My coworkers

_____ Public opinion

_____ Reports, studies and other written material

Survey Responses

The following list includes the different dependent variables in the survey used to measure government official proficiency in, interest in, and intent in using evidence-based reports. The definitions show the question as presented in the survey along with the possible answers officials could select. The percentages along the side of each answer show a tabulation of what percent of survey takers chose each response.

18. Seen Studies (1–3)

International organizations and governments, often in partnership with academic researchers, are performing rigorous impact assessments of government programs and publishing reports with their findings. Have you seen such academic reports?

Yes, many	20.34%
Yes, some	59.31%
No, none	20.34%

19. Use Studies (1–4)

How often do you consult such academic reports when you make policy decisions?

Every time	8.35%
Most times	24.08%
Sometimes	52.58%
Never	14.99%

20. Recent Studies (1–6)

How recently have you used academic reports to make a policy decision?

In the past month	23.89%
In the past six months	22.17%
In the past year	22.91%
In the past five years	7.14%
More than five years	3.94%
Never	19.95%

21. Interest Studies (1–5)

How interested are you in learning from academic research?

Very interested   61.58%

Somewhat interested   28.82%

Indifferent   4.93%

Somewhat disinterested   2.46%

Not interested at all   2.22%

22. Know Studies (1–4)

How well do you feel you understand academic research?

Very well   41.38%

Somewhat well   47.04%

Not very well   9.61%

Not well at all   1.97%

23. Useful Studies (1–5)

How helpful is academic research is in making policy?

Very useful   62.22%

Somewhat useful   29.14%

Not very useful   7.16%

Not useful at all   1.48%

24. Work Better (1–4)

Do you think using academic research makes your work better?

Yes, definitely   65.10%

Yes, somewhat   27.23%

No, not really   6.93%

No, definitely not   0.74%

The following rankings of 1–7 were later reversed in the indexes to make 1 least important and 7 most important:

17. What source is most useful to you when making your policy decisions (Rank of evidence-based reports, 1–7)?

1) 38.69%

2) 16.83%

3) 14.32%

4) 6.53%

5) 5.78%

6) 6.03%

7) 11.81%

25. What sources do you think are the most important for policy makers to use in decision-making (Rank of evidence-based reports, 1–7)?

1) 40.40%

2) 18.94%

3) 10.10%

4) 5.56%

5) 7.58%

6) 6.06%

7) 11.36%

26. What sources do you think your coworkers use in decision-making (Rank of evidence-based reports, 1–7)?

1) 25.59%

2) 14.12%

3) 12.65%

4) 10.88%

5) 9.41%

6) 9.12%

7) 18.24%

27. What sources do you think your superiors want you to use in decision-making (Rank of evidence-based reports, 1–7)?

1) 38.64%

2) 15.93%

3) 14.75%

4) 8.26%

5) 5.60%

6) 7.37%

7) 9.44%

REFERENCES

Avey, Paul C. and Michael C. Desch. 2014. "What Do Policymakers Want from Us? Results of a Survey of Current and Former Senior National Security Decision Makers." *International Studies Quarterly* 58, no. 2 (June): 227–46.

Banerjee, Abhijit V. and Esther Duflo. 2009. "The Experimental Approach to Developmental Economics." *Annual Review of Economics* 1 (September): 151–78.

Banks, Gary. 2009. "Evidence-Based Policy Making: What is It? How Do We Get It?" ANU Public Lecture Series, Australian Government Productivity Commission, Canberra, February 4.

Beede, David N., Tiffany A. Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E. Doms. 2011. "Women in STEM: A Gender Gap to Innovation." Economics and Statistics Administration, Issue Brief #04–11 (August). Accessed March 7, 2018. www.esa.doc.gov/sites/default/files/womeninstemagaptoinnovation8311.pdf.

Black, Nick. 2001. "Evidence Based Policy: Proceed with Care." *British Medical Journal* 323 (August): 275–79.

Brandon, Emily. 2014. "The Ideal Retirement Age—And Why You Won't Retire Then." *U.S. News & World Report*, May 12. Accessed November 15, 2017. https://money.usnews.com/money/retirement/articles/2014/05/12/the-ideal-retirement-age-and-why-you-wont-retire-then.

Brossard, Dominique, Dietram A. Scheufele, Eunkyung Kim, and Bruce V. Lewenstein. 2008. "Religiosity as a Perceptual Filter: Examining Processes of Opinion Formation about Nanotechnology." *Public Understanding of Science* 18, no. 5 (September): 546–58.

Buddin, Richard and Gema Zamarro. 2009. "Teacher Qualification and Student Achievement in Urban Elementary Schools." *Journal of Urban Economics* 66, no. 2 (September): 103–15.

Cardoso, Ana Rute, Paulo Guimarães, and José Varejão. 2011. "Are Older Workers Worthy of Their

Pay? An Empirical Investigation of Age-Productivity and Age-Wage Nexuses." *De Economist* 159, no. 95 (June).

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862–1928.* Princeton: Princeton University Press.

Dollar, David and Roberta Gatti. 2001. "Are Women Really the 'Fairer' Sex? Corruption and Women in Government." *Journal of Economic Behavior and Organization* 46, no. 4 (December): 423–29.

Fukuyama, Francis. 2013. "What is Governance?" *Governance: An International Journal of Policy, Administration and Institutions* 26, no. 3 (July): 347–68.

Garcia-Zamor, Jean-Claude. 2003. "Workplace Spirituality and Organizational Performance." *Public Administration Review* 63, no. 3 (May): 355–63.

Gates, Robert M. 2008. Address to the Association of American Universities. Washington, DC, April 18. Accessed November 15, 2017. http://archive.defense.gov/Speeches/Speech.aspx?SpeechID=1228.

Head, Brian W. 2016. "Toward More 'Evidence-Informed' Policy Making?" *Public Administration Review* 76, no. 3 (May/June): 472–84.

Hoxby, Caroline M. 2000. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics* 115, no. 4 (November): 1, 239–85.

Inglehart, Ronald and Wayne E. Baker. 2000. "Modernization, Cultural Change, and the Persistence of Traditional Values." *American Sociological Review* 65, no. 1 (February): 19–51.

Kauppi, Katri and Erik M. van Raaij. 2014. "Opportunism and Honest Incompetence - Seeking Explanations for Noncompliance in Public Procurement." *Journal of Public Administration Research and Theory* 25, no. 3 (July): 953–79.

Krosnick, J.A. 2002. "The Causes of No-Opinion Responses to Attitude Measures in Surveys: They Rarely are What They Appear to Be." In *Survey Nonresponse,* edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J.A. Little, 87–100. New York: Wiley.

Leahey, Erin. 2006. "Gender Differences in Productivity: Research Specialization as a Missing Link." *Gender & Society* 20, no. 6 (December): 754–780.

Lizardo, Radhames and Andre V. Mollick. 2013. "Can Latin America Prosper by Reducing the Size of Government?" *Cato Journal* 29, no 2 (May): 247–66.

Lupia, Arthur. 2003. "Delegation and its Perils." In *Delegation and Accountability in Parliamentary Democracies*, edited by Kaare Strom, Wolfgang C. Muller, and Torbjorn Bergman, 33–54. Oxford: Oxford University Press.

Lupia, Arthur and Mathew D McCubbins. 2000. "Representation of Abdication? How Citizens Use Institutions to Help Delegation Succeed." *European Journal of Political Research* 37, no. 3 (May): 291–307.

Maliniak, Daniel, Susan Peterson, and Michael J. Tierney. 2012. "TRIP around the World: Teaching, Research, and Policy Views of International Relations Faculty in 20 Countries." Williamsburg, VA: Teaching, Research, and International Policy Project.

Pasta, David J. 2009. "Learning When to Be Discrete: Continuous vs. Categorical Predictors." SAS Global Forum, Paper 248.

Prpić, Katrina. 2002. "Gender and Productivity Differentials in Science." *Scientometrics* 55, no. 1: 27–58.

Rajkumar, Andre Sunil and Vinaya Swaroop. 2008. "Public Spending and Outcomes: Does Governance Matter?" *Journal of Development and Economics* 86, no. 1 (April): 96–111.

Sanders, William L., S. Paul Wright, and Sandra P. Horn. 1997. "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation." *Journal of Personnel Evaluation in Education* 11, no. 1 (April): 57–67.

Skirbekk, Vegard. 2004. "Age and Individual Productivity: A Literature Survey." *Vienna Yearbook of Population Research* 2: 133–53.

Slavin, Robert E. 2008. "Cooperative Learning, Success for All, and Evidence-Based Reform in Education." *Education Didactique* 2, no. 2 (September): 149–57.

The World Bank. 2016. Data. GDP per capita (current US$). Accessed October 22, 2017. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

Twain, Mark. 1907. In *Mark Twain: A Biography* by Albert Bigelow Paine. The University of Adelaide Library. Accessed November 14, 2017. https://ebooks.adelaide.edu.au/t/twain/mark/paine/.

Zeglovits, Eva and Sylvia Kritzinger. 2014. "New Attempts to Reduce Overreporting of Voter Turnout and Their Effects." *International Journal of Public Opinion Research* 26, no. 2 (June): 224–34.

Religion, Nationality, and Empathetic Responses to Refugees

Reed Rasband

Introduction

Recent years have witnessed a dramatic surge in refugees seeking asylum throughout the developed world. While many of these asylum seekers originate from war-torn countries in the Middle East, large numbers have also left countries in Africa, Central America, and Southeast Asia. Contentious national responses to this surge have pitted supporters of refugee resettlement against those who express concerns over national security, vetting procedures, and threats to national cultural unity. In the UK, public discourse about refugee resettlement has been accompanied by rising rates of Islamophobic incidents (Adesina and Marocico 2015; Shepherd 2016; Batchelor 2017).

In these conditions, some opinion groups circulate anxiety-provoking information about refugees, while at the same time many media outlets choose to run stories portraying refugees as empathy-deserving victims of circumstance. Most research, however, focuses on the former and not the latter (Brader et al. 2008; Gadarian and Albertson 2014; Albertson and Gadarian 2015). This leaves us with much less certainty about whether empathetic media effectively shifts readers' attitudes toward refugee policies. We also do not know if these appeals are effective for the population as a whole (versus only a subset) or if appeals about certain types of refugees are more effective than others.

To answer these questions, I implement a survey experiment that exposes UK residents to an empathetic article about refugees, with descriptions that varied the refugees' religion and national origin. This experiment not only fills the current gap in knowledge about empathetic outgroup portrayals but also clarifies whether prejudice against certain immigrant groups reflects their religion or national origin by concurrently manipulating both. It also accounts for variations in response among individuals with differing degrees of pre-existing outgroup antipathy by measuring respondents' attitudes before treatment.

I find that empathetic media appeals about refugees are effective in generating empathy and shifting policy opinions for most individuals but hardly influence individuals with strong prior feelings of antipathy toward refugees. For UK residents overall, opinions are particularly intransigent when considering Muslim refugees. Articles about Christian refugees dampened active opposition (willingness to send a letter about the issue to a respondent's MP) among most people, and articles about non-Christian refugees motivated supportive action among individuals with little prior outgroup antipathy. Thus, empathetic appeals are ineffective in swaying the citizens who are perhaps their most important target: strongly prejudiced individuals.

Citizen Responses to Empathetic Appeals

Political psychologists frequently point to affect or emotion as determinant of political attitudes. In particular, a growing set of recent research has explored the connection between emotions and immigration attitudes, of which refugees are a subset. Research has shown that different immigrant ethnicities elicit differing levels of anxiety from Americans and Spaniards (Brader et al. 2008; Igartua and Cheng 2009). Anti-immigrant media can encourage negative feelings, such as anxiety or anger, which then encourages restrictive policy attitudes, while pro-immigrant media induces feelings such as sympathy and enthusiasm, which then discourages restrictive attitudes (Verkuyten 2004; Brader et al. 2008; Iguartua and Cheng 2009; Lecheler et al. 2015). However, experimental research into how positive emotions, such as empathy, influence support for immigration has been less extensive. Verkuyten (2004), Lecheler et al. (2015), and Adida et al. (2017) are among the few articles that take this approach directly. This lack of attention is surprising given the rising level of empathetic immigrant/refugee appeals currently promulgated through traditional and social media. These appeals are common, yet we have little knowledge of whether they are effective in influencing the American public.

Empathy is an emotional response defined as "the act of perceiving, understanding, experiencing, and responding to the emotional state and ideas of another person" (Barker 2008, quoted in Gerdes et al. 2010) or "the other-focused, congruent emotion produced by witnessing another person's suffering" (Batson et al. 1987). Generally, empathy is measured by self-report measures, where people report how strongly they are feeling a set of related emotions, such as compassion, sympathy, or tenderness (Batson et al. 2002, 1997; Gerdes et al. 2010). Typical experimental designs expose individuals to information portraying other people in emotional situations, with the expectation that subjects will absorb or reflect that same emotional state (Eisenberg and Miller 1987). Some authors also use exercises that ask respondents to imagine themselves in the place of another person, a treatment that perhaps adheres more closely to definitions of empathy as a concept but drifts away from the sorts of stimuli citizens would encounter in an everyday setting (Adida et al. 2017).

Before considering how empathetic appeals may change public opinion and political behavior, we must first establish that they actually provoke an empathetic emotional response. If they do, the following expectation would hold:

H1) Individuals exposed to media that presents refugees in an empathetic light will feel more empathy than people who are not.

Simply feeling an emotion, such as empathy, in response to media about an out-group does not guarantee that other variables, such as policy opinions and political behavior, will also change. Still, research has shown that these outcomes often do change in response to emotion-generating stimuli. For example, studies show that enthusiasm increases voting, anxiety increases political learning, and anger increases campaign involvement (Marcus and MacKuen 1993; Valentino et al. 2011). Anxiety increases people's propensities to seek further information about immigration threats, to trust certain political actors, and to favor restrictive immigration policies (Brader et al. 2008; Gadarian and Alberson 2014; Albertson and Gadarian 2015). These studies portray emotions as mediating variables between stimuli (news media or campaign ads about immigration) and public opinion or behavior around immigration issues.

It follows that empathetic appeals may also be effective in shifting these other outcomes. After all, research on empathy in non-political contexts has yielded results showing that empathetic treatments encourage prosocial behaviors, such as helping others, while also improving perceptions of stigmatized outgroups (Coke et al. 1978; Eisenberg and Miller 1987; Batson et al. 1987; Batson et al. 2002, 1997). "Taking the perspective of a person in need tends to increase one's empathic emotional response [and] empathic emotion in turn increases motivation to see that person's need reduced" (Coke et al. 1978). Within political science, Bansak, Hainmueller, and Hangartner's (2016) findings that citizens are more likely to approve hypothetical refugee applications if the refugee has experienced trauma, such as torture or the death of family members, also suggests that empathy can change attitudes and behavior. Adida et al. (2017) found that empathy treatments were effective in promoting inclusionary attitudes toward refugees and encouraging political action in favor of refugees, at least among certain subsections of the U.S. population.

If empathetic appeals indeed are effective in generating prosocial attitudes and behaviors toward refugees, the following expectations would hold in addition to the hypothesized relationship between these appeals and feeling empathy itself (H1):

H2) Individuals who encounter empathetic media treatments about refugees will (a) be more likely to hold opinions supporting open refugee policies and (b) will be more likely to act in support of those opinions.

Up to this point, the hypotheses have implied a uniform treatment effect across all individuals, yet there is reason to question that assumption. A recent experimental study that asked individuals to reject or deny hypothetical asylum applications found that 9 percent of European respondents rejected *all* applicants, regardless of their skillset or experienced vulnerabilities (Bansak et al. 2016). Adida et al. (2017) also noted a strong backlash against empathetic refugee appeals among sections of the U.S. populace. These findings indicate the presence of a subset of the population who dislike

refugees as a group. It is reasonable to suppose that emotional responses to sympathetic portrayals of refugees and other outgroups might differ for this subset.

There are several approaches to explaining how attitudes may form in this case. Theories of motivated reasoning indicate that existing affect continues to influence individuals' appraisals of new information (Redlawsk 2002; Gadarian and Albertson 2014). In this vein, various studies found that people with strong prior attitudes have greater difficulty processing concepts that they find affectively incongruent, are more skeptical of such material and tend to argue against messages that do not match their existing feelings (Lodge and Taber 2005; Lodge and Taber 2013; Taber and Lodge 2006). For these people, new empathetic information about refugees may be harder to incorporate into their existing beliefs and may generate counterarguments contradicting the intended message, proving less effective in changing opinions about refugee policy. This effect is similar to that predicted by Zaller's Receive-Accept-Sample (RAS) model, which posits that individuals are less likely to accept new information that is inconsistent with their prior beliefs (1992).

Another potential explanation for the hard-to-move attitudes among people with strong outgroup biases is that empathy may trigger distressing emotions as well (Eisenberg et al. 2014). Generally, political psychologists have incorporated this potential for distress into their theories through the concept of dissonance, a discordant emotional response that arises when people encounter information about an outgroup that contradicts their core beliefs or existing ideals about themselves, their own in-group, or an out-group (Gubler 2016; Gubler et al. 2016). These feelings might be generated, for example, if an individual encounters evidence portraying her own group as violent, or information portraying a stigmatized outgroup positively (Ibid.). Citing Harmon-Jones, Amodio, and Harmon-Jones (2009), Gubler suggests that dissonant feelings might be especially common "when the cognitive inconsistency suggests a new 'action tendency' that an individual is loath to adopt—e.g. an individual feels like he needs to act differently as a result of the new information but knows that behavior change will be costly or difficult" (2016). In this situation, it may prove more comfortable for the receiver simply to reject the new information or even generate self-justifying arguments that strengthen his or her own attitudes *against* the outgroup (Ibid.).

Whether we approach this issue from the angle of motivated reasoning or dissonance, the expected result is the same: Individuals with strong negative affect toward refugees (outgroup antipathy) will be less likely to change their opinions about refugee policies than other individuals. Inasmuch as empathetic media provokes either phenomenon in these individuals, they may become more likely to support closed refugee policies and to reveal those attitudes in their political behavior.

H3) Individuals will feel more empathy and exhibit stronger opinions and action supporting pro-refugee policies as their level of prior outgroup antipathy decreases.

In addition to variation in treatment effects across different types of individuals, it is probable that treatment effects will also vary for different types of refugees. Studies

of group attitudes such as prejudice or in-group centrism often find that these attitudes also relate to attitudes about specific outgroups (Brader et al. 2010; Valentino et al. 2013; Hutchings and Wong 2014; Gubler 2016; Gubler et al. 2016). Social identity theory suggests that, within specific cultures, some group attributes are more salient markers of in-group/outgroup categorization and comparison than others (Tajfel and Turner 1979). Because of this, we might expect empathetic appeals to be especially ineffective for high-antipathy individuals considering refugees whose religious, ethnic, or other characteristics particularly set them apart from the cultural norms of the receiving nation.

A large body of recent research in the U.S. and Europe has investigated how attitudes toward refugees and immigrants vary according to immigrant characteristics. These studies have established that receiving populations consistently prefer the immigration of high-status, educated, and skilled individuals who speak the local language instead of comparatively low-status individuals, who perhaps may have greater difficulty assimilating into the new culture (Harell et al. 2012; Iyengar et al. 2013; Hainmueller and Hopkins 2015; Bansak et al. 2016; Ostfeld 2017). Investigation of cultural factors, such as national origin and religion, generally uncovers bias against Middle Easterners and against Muslims (Igartua and Cheng 2009; Adida et al. 2010; Adida et al. 2017; Harell et al. 2012; Iyengar et al. 2013; Hainmueller and Hopkins 2015; Bansak et al. 2016; Adida et al. 2017).

Most of these experiments only vary one or two immigrant traits at a time, making it difficult to parse out which characteristics provoke differing responses (Hainmueller and Hopkins 2015; Bansak et al. 2016). This seems to be an especially relevant concern for generating conclusions about cultural cues like national origin and religion. Some studies look only at religion without looking at national origin (generally Muslim versus Christian) (Adida et al. 2010; Adida et al. 2017), while others only look at national origin without separating religion (Sniderman et al. 2004; Harell et al. 2012; Iyengar et al. 2013; Hainmueller and Hopkins 2015). Some authors have suggested religious stereotypes are the salient characteristics that provoke hostility toward immigrants from the Middle East but acknowledge that no clear conclusions on this issue are possible without experimentally manipulating both religion and national origin (Sniderman et al. 2004; Iyengar et al. 2013).

Bansak et al. (2016) is one of the only studies that manipulates both of these variables, in the context of a conjoint experiment asking Europeans to evaluate asylum seekers for admittance. They noted minor differences in asylum acceptance rates between countries of origin but significantly lower acceptance for Muslims vis-a-vis Christians.

In the present study, I will further elucidate the ambiguity between religious discrimination and national origin discrimination present in the existing literature. Following the general expectations of the authors above, I also expect that most ill will felt toward Middle Easterners actually reflects religious bias against Muslims, not concerns about a specific country of origin.

This hypothesis seems consistent with growing levels of Islamophobic incidents in the UK (Adesina and Marocico 2015; Shepherd 2016; Batchelor 2017) and would match the findings of Bansak et al. (2016). Furthermore, Tajfel reports that multigroup membership and overlapping social identities can help to break down discriminatory attitudes toward an outgroup (1982). Religion is a social identity category that can either associate refugees with an overlapping in-group trait (Christianity in the British context) or further differentiate the refugee outgroup from the ingroup. The degree to which citizens categorize refugees as an outgroup then determines their receptiveness to empathetic appeals about refugees. Refugees' nationality, in contrast, is a category that few British residents would share (or perhaps even recognize), making variation in national origin unhelpful for moderating identity categorizations when all other traits are held equal or when other, more familiar categories are referenced.

In the context of empathetic media accounts, Islamophobic bias would be manifest by less significant changes in opinion and behavior outcomes after respondents consider Muslim refugees than when they consider other religious groups. If this hypothesis holds, religious differences will predict support of refugee policies better than differences in national origin, especially for individuals who have strong pre-existing biases against refugees in general.

H4) Empathetic media about Christian refugees will encourage citizens to support open refugee policies more than media that references Muslim refugees, regardless of the refugees' national origin.

Experimental Procedure

In order to test these hypotheses, I conducted a survey experiment on a sample of UK residents. Preliminary results for this experiment come from an online survey performed on Prolific Academic, a UK-based online platform similar to Amazon Mechanical Turk that allows researchers to convenience sample a large and varied, though not necessarily representative, segment of the UK population (Peer et al. 2017). When used as response pools for experimental studies, online crowdsourcing markets such as these are known to yield results that are as valid as results obtained in laboratory or field experiments (Horton et al. 2011; Berinksy, Huber, and Lenz 2012). I fielded the experiment in September 2017 with a group of 1,421 adult UK residents. A quota mechanism along with block randomization by gender ensured there would be a balanced gender distribution across the entire sample and within treatment conditions. More details about the demographic characteristics of the sample are included in Appendix 1.

Successfully testing Hypothesis 4 requires a measure of outgroup antipathy, since emotional responses should depend in part on those pre-existing attitudes. To do this, respondents answered three questions asking them to rate, on a seven-point scale, whether they agreed or disagreed with a statement, such as "Refugees are more prone to violence than other groups." The statements were taken from a larger set asked by Gubler et al. (2016) and are included in Appendix 2.

These statements were presented in random order at the beginning of the survey. To minimize social desirability bias, the questions began with the statement, "There are no right or wrong answers to these questions. We're simply interested in your honest responses." In order to minimize priming effects from these statements that might interfere with the treatments, the respondents completed a number of other questions on unrelated issues before receiving the treatment or control articles and answering the remaining questions of this survey.

Respondents also answered a battery of standard questions measuring political psychological characteristics. These questions are primarily taken from the 2015 American National Election Survey (ANES) and the British Election Study and include measurements of political party affiliation, political ideology (left/right), interest in politics, egalitarianism, and authoritarianism (this last is presented through questions about values in child rearing).

Following these questions, respondents read either a news article that discussed a refugee family that resettled in London (the experimental treatment) or an article that discussed the nutrition value of fruit juice (the placebo control). Random assignment determined which article respondents read. Presenting the information about the refugees as a news article was done to mimic a real-world context in which UK residents might encounter similar humanizing appeals.

There were six versions of the treatment article, each of which varied the family's national origin and religion following a 2 × 3 design (Syrian Muslim, Syrian Christian, Syrian no religion, Sudanese Muslim, Sudanese Christian, Sudanese no religion). This allows analysis based on not only national origin but also religion. The no-religion condition only varied the national origin of the refugees, without any information about their religious beliefs. The refugees' names also varied to match the families' ethnic origins.

Sudan and Syria were chosen as the countries of origin due to their relevance in the British context and the similarity of the political situation in both countries. Britain has received a substantial number of refugees from both countries in recent years (UK Home Office 2017), and both countries have dealt with bloody civil wars that have received substantial recent coverage from UK and international news media. Besides this, the two countries have similar levels of average income and education and are equally distant from the UK, helping us to discard the possibility that respondents' assumptions about those characteristics will influence their responses. Importantly, for the sake of treatment credibility, Sudan and Syria both have populations of Christians and Muslims. Crucially, the ethnic and racial makeup of these countries is distinct, with most Sudanese asylum seekers being of African origin, while Syrian refugees are generally Arab. These characteristics allow us to test Hypotheses 1 and 2 without large concerns of confounding variables.

All other characteristics of the family remained constant in all six treatments. Respondents learned the refugee family had close family members that were killed and abducted during war in their home countries, had lost their material possessions, had

ing these questions, respondents responded to a battery of questions measuring their demographic attributes.

As a manipulation check, respondents answered two questions at the end of the survey (after completing items measuring their emotions, attitudes, and behavior) that asked them to identify the family's religion (from a list of options in random order including Muslim, Christian, Hindu, Buddhist, Jewish, or none) and nationality (from a list including Syria, Sudan, and some additional distractor options). This allows me to ensure that respondents were truly cognizant of the refugees' religion and nationality as they read the article. Results showed that 81.4 percent of treated respondents correctly remembered the religion and 85.5 percent correctly remembered the nationality of the refugees. This suggests that the manipulation was successful.

The end of the survey included a debriefing disclaimer informing individuals in the treatment groups that the religion and national origin of the family they read about may have been modified. All respondents also received a disclaimer that no message to their MP would actually be sent and that the question was for survey purposes only. Individuals received the debriefing material after all other data had been collected.

After data collection, successful randomization was confirmed using Wald tests to search for significant differences between treatment groups for all demographic and psychological characteristics (see Appendix 1 for details). Results show no systematic differences across treatment groups beyond what would be expected by random chance. In order to construct the empathy and dissonance dependent variables, factor analysis with varimax rotation of the fifteen emotion variables yielded two unambiguous factors (empathy and dissonance), which corresponded to my *a priori* expectations. Details on the factor loadings for the emotions outcomes are available in Appendix 1. Each of the antipathy, dissonance, and empathy scales were rescaled from zero to one for ease of interpretation. The opinion measure was also rescaled from zero to one.
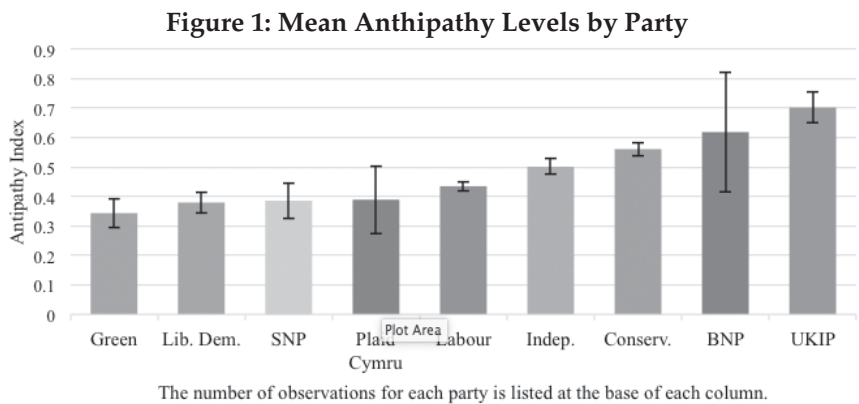
Results

Before sequentially discussing the treatment effects on empathy, opinions, and political behavior, it will be useful to discuss some details about the outgroup antipathy measure, since it is used as the basis for conditional analysis of all three outcomes.

Responses to the outgroup three antipathy questions yielded an approximately normally distributed index with a mean of .466, a standard deviation of .213, and Cronbach's alpha of .644. The somewhat low internal consistency of this scale, although not substantially different from typical standards, is understandable given that it is composed of only three questions taken from a larger index.

For ease of interpretation, and to avoid imposing a linear form on the relationship of the outcomes with antipathy, interaction effects with antipathy in the subsequent analysis are estimated with conditional effects models that separate the sample by antipathy level. In these models, high antipathy indicates an index score above .75 (N=130), mid antipathy includes index scores less than or equal to .75 but above .25 (N=1061), and low antipathy indicates an index score of .25 or less (N=230). Although

the smaller sample size of individuals with high outgroup antipathy is somewhat problematic for identifying treatment effects within this group, I show below that significant differences are still visible. Regression results that treat the antipathy variable as continuous show the same patterns as the conditional effects models. Expected values plots for the continuous models are available in Appendix 1.

Note that the statements used to generate the present results referred to "refugees" in general; respondents were not asked to elaborate on which assumptions, preconceptions, or suppositions guided their answers to the questions that made up the index. It will be up to future research (including an already-planned extension of this study) to delve more deeply into how British citizens define refugees as an outgroup.

Figure 1: Mean Anthipathy Levels by Party

The number of observations for each party is listed at the base of each column.

As a somewhat newer measure, it will be useful to compare outgroup antipathy with other political psychological indices. In this sample, outgroup antipathy correlates positively with authoritarianism ($\varrho = 0.375$) and ideology ($\varrho = 0.426$, where the variable increases from left to right) and negatively with egalitarianism ($\varrho = -0.363$). Adida et al. (2017) found that reactions to refugee policy differed by party in the U.S.; Figure 1 shows that within the UK context, antipathy levels also vary with party affiliation, with left-leaning parties showing less antipathy than right-leaning ones.

Treatment Effects on Emotions

Overall, reported empathy increased for individuals who had read the empathetic media article, as predicted by H1. H2, which deals with the opinion and behavior outcomes, is discussed in the next sections. As expected from H3, the strength of the treatment varied according to citizens' levels of prior outgroup antipathy. Figure 2 shows the expected values of empathy for each religious treatment condition for both the overall sample and conditional upon levels of prior antipathy. Figure 3 similarly reports the expected empathy values by nationality treatment condition. Regression tables for these and all following figures are provided in Appendix 1. Results show that low-antipathy individuals experienced the largest treatment effects, while mid-antip-

athy individuals saw milder but still significant increases in empathy. High-antipathy individuals' levels of antipathy were not distinguishable from the control.

Figure 2: Expected Values of Empathy by Religion Treatment
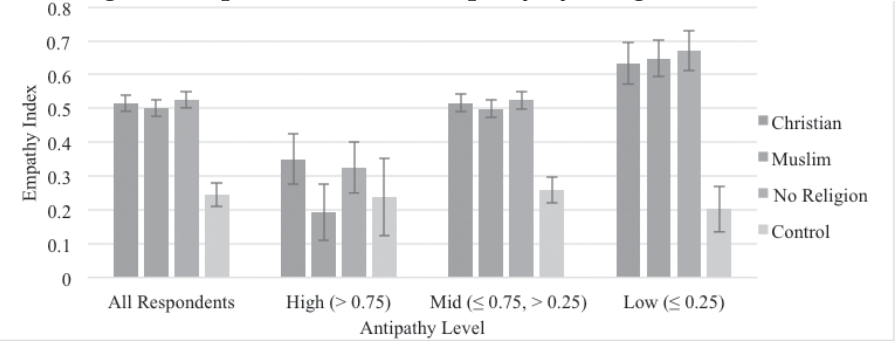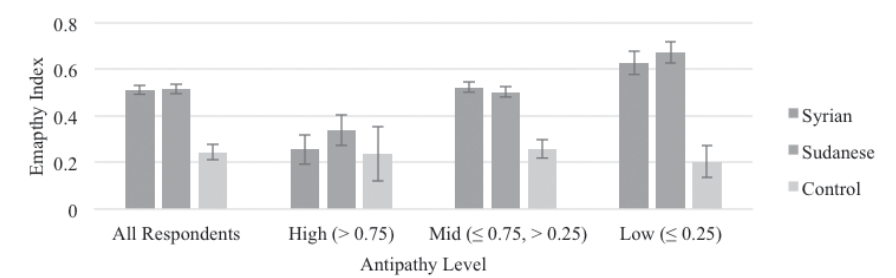

**Figure 3: Expected Values of Empathy by Nationality Treatment**

While Hypothesis 4 predicted differing treatment effects by religion for the opinion outcome, even in the empathy outcome we begin to see indications of religious bias. These differences are not apparent when observing the sample overall, nor are religious differences significant for low- and mid-antipathy individuals. Nevertheless, as shown in Figure 2, high-antipathy individuals' feelings of empathy are lower after reading about Muslim refugees. None of the three religious treatments is statistically different from the control for this subgroup. However, the Muslim treatments elicit significantly lower empathy than Christian treatments ($P<0.007$) and treatments where religion is not mentioned ($P<0.022$). Interestingly, this more detailed analysis also shows that treatment may not have been completely ineffective for the high-antipathy group: The coefficient for the Christian and no religion treatments was positive. The difference between the Christian treatment and the control group nearly attained established significance levels ($P<0.109$) and may have reached established benchmarks had a larger number of high-antipathy individuals been surveyed.

Further analysis tested whether the refugees' religion or nationality affected feelings of empathy. As shown in Figure 3, there are no significant differences between Syrian and

Sudanese treatments vis-à-vis the control, regardless of the respondents' antipathy level. Consistent with the expectations of H4, refugee nationality does not seem to affect significantly respondents' feelings of empathy in response to the treatment article.

Before continuing to the opinion outcome, I provide some basic analysis of reported dissonance in order to explore at least minimally the mechanisms behind the lack of treatment effects for high-antipathy individuals. Figure 4 shows expected values plots for each treatment condition and the control. As happened for empathy, treatment increased feelings of dissonance overall, although to a much milder degree than it did for empathy. Treatment increased empathy by about .27 points on the empathy scale but did not surpass an increase of .11 points on the dissonance scale. The significance of the findings decreased for low-antipathy individuals in Muslim treatment conditions and high-antipathy individuals in Christian treatment conditions. In other words, high-antipathy individuals felt no significant increase in distress when reading about Christian refugees but did feel this distress when reading about Muslim refugees. This pattern was reversed for low-antipathy individuals. The fact that all individuals, regardless of their level of prior antipathy, felt relatively similar levels of dissonance leaves open the possibility that respondents at either end of the antipathy scale are reporting the same emotions for different reasons. I discuss this possibility further in the discussion section.

## Figure 4: Expected Dissonance Levels
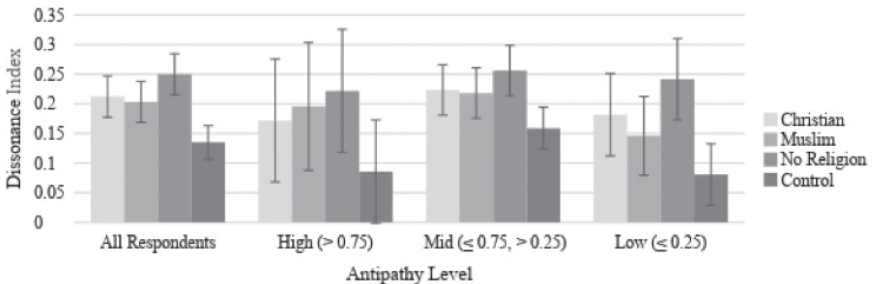
Treatment Effects on Opinions

Figure 5 shows expected values for opinions on refugee policy in the UK by religion and antipathy level. H2a predicted that empathetic media appeals would also affect respondents' opinions about refugee policy in addition to producing empathy itself. Supporting this expectation, results show a slight, yet significant, shift in public opinion away from reducing refugee intake.

H3 predicted conditional effects according to prior antipathy. Subdividing analysis by antipathy group, we see basically the same patterns for opinions as were visible for the empathy outcome, consistent with H3. Low-antipathy individuals shifted their already-welcoming opinions toward admitting even more refugees in response to all treatments. Mid-antipathy individuals also saw treatment effects on opinion but only for individuals who read about Christians or refugees without a specified religion; Muslim

treatments generated null effects compared to the control. High-antipathy individuals advocated a drastic reduction in refugee intake regardless of which article they read. Differences between the treatment groups are not significant for this subset of respondents, but average opinions in the Muslim treatment are lower than for the other treatment groups, just as they were for empathy.

Figure 5: Expected Opinion Levels
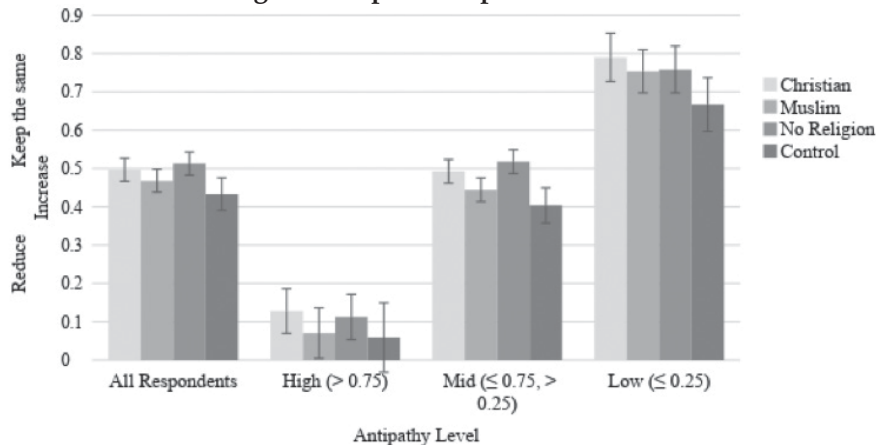
The lack of opinion change in the Muslim treatment condition for high- and mid-antipathy individuals provides evidence of the Islamophobic religious bias anticipated by H4. However, the fact that these differences were significant for the large group of mid-antipathy individuals contrasts with the observations for empathy and also suggests that anti-Muslim bias extends far beyond just a small segment of the population. Utilizing the same analytical strategy as for the empathy outcome, I found that refugee national origin was not a significant factor in changing citizens' opinions. As shown in Figure 6, the expected values of the opinion index for the Syrian treatments are virtually

Figure 6: Expected Opinion Levels by Refugee Nationality
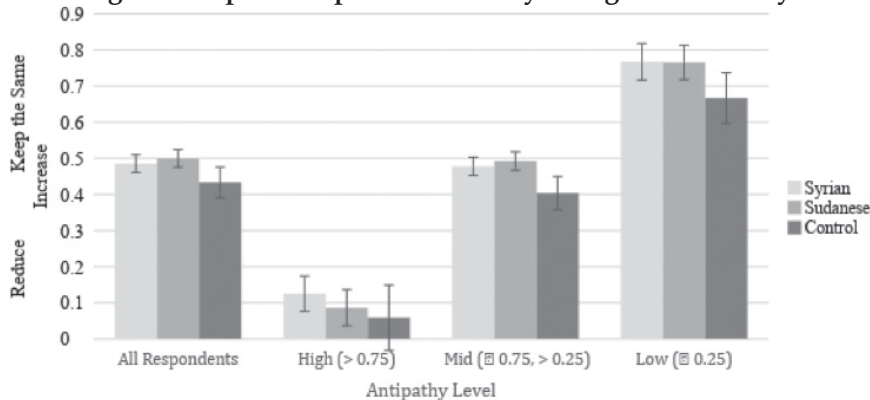
identical to those of the Sudanese treatments, regardless of respondents' levels of prior outgroup antipathy.

Treatment Effects on Behavior

With regard to the behavior outcome, by far the most common choice for responding citizens was not to take action on the proposed bill to allow admittance to refugees' families. Overall, 56.3 percent chose not to send a letter to their parliamentary representative, 29.9 percent chose to send a message in support of the bill, and 13.8 percent chose to voice their opposition to the bill. Because the behavior measure allowed for three potential outcomes, analysis was performed using multinomial logistic regression, with no action being the baseline outcome for comparison. Multinomial logistic regression was used instead of ordered logistic regression since the three behavior outcomes do not fit onto a single ordered schema; the support and oppose option both indicate greater willingness to act than the option to not send any message at all, but they do so in opposite directions in terms of openness toward refugees.

Figure 7 shows predicted propensities of active support for the pro-refugee bill for the whole sample and for each antipathy subgroup, as generated from the regression results. Figure 8 shows the propensities of active opposition to the pro-refugee bill. The results table for these figures is available in Appendix 1. The results show that the behavior measure is the only outcome variable for which no treatment effects are distinguishable over the entire sample, contradicting the expectations of H2b. Whether treatments are analyzed by religion or by nationality, no significant differences in action propensity are visible relative to the control. Wald tests between the regression coefficients of the nationality analysis, both overall and by antipathy subgroup, show that supportive action for the Sudanese treatment groups was slightly higher than the Syrian treatment groups. However, this effect only reached the 0.05 significance level when analysis was conducted on the sample overall (not in analyses divided by antipathy level). The fact that at least one difference in nationality treatment effects reached significance for the behavior outcome but not for opinion or empathy indicates that political behavior is a separate phenomenon from those outcomes that may be susceptible to nation-of-origin priming effects.

As shown, the propensity to actively support the bill increases as outgroup antipathy decreases, while the propensity to actively oppose the bill increases as antipathy increases. In most cases, however, this pattern only represents a prior attitudes effect, not a treatment effect interacting with prior outgroup antipathy as was observed in the empathy and opinions outcomes. The propensity to actively support or oppose the bill was statistically indistinguishable from the control condition in nearly all treatment conditions, regardless of the degree to which respondents reported prior outgroup antipathy.

However, when treatments are examined by antipathy subgroup, two conditional effects do emerge, as shown in the conditional effects plots from Figures 7 and 8. First, an action motivation effect is observed for low-antipathy individuals who read articles about Muslim refugees or about refugees without reference to religion (Fig. 7). These

individuals are more likely to request a letter be sent expressing their support for the bill than individuals who read about Christians or who read the placebo article. No treatment effects are observed for high-antipathy individuals. It is, however, notable that not a single high-antipathy individual chose to act in support of the bill after reading about Muslim refugees (this was the only treatment category that generated zero proponents among an antipathy subgroup).

Second, an action inhibition effect is observed for mid-antipathy individuals who received a treatment about Christian refugees (Fig. 8). These individuals are less likely to actively oppose the bill than are mid-level individuals who read about other refugees or received the placebo article.

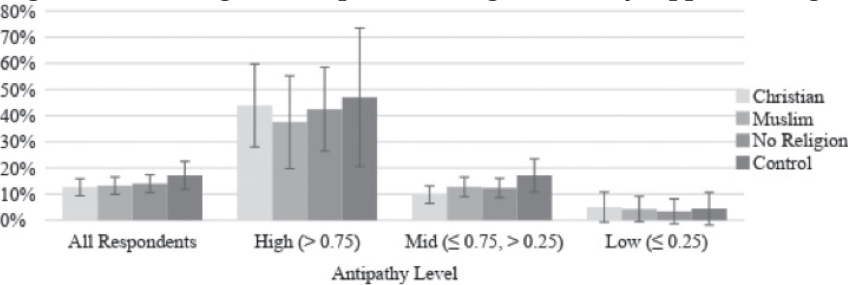## Figure 7: Percentage of Sample Choosing to Actively Support Refugees
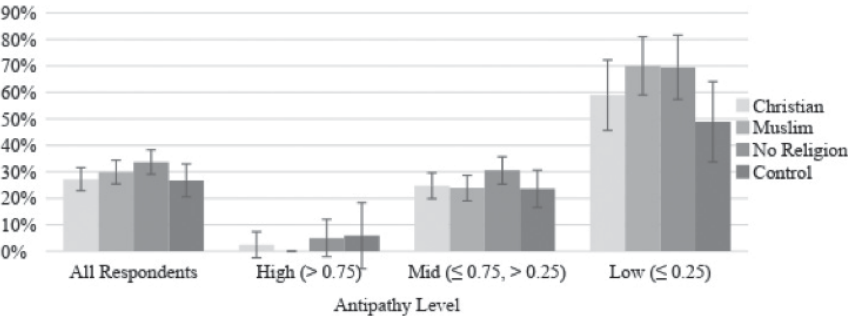
Figure 8: Percentage of Sample Choosing to Actively Oppose Refugees

Discussion

H1) Empathetic Media Promotes Empathy

While this hypothesis may seem simple, it was by no means a guarantee. Indeed, roughly 9 percent of individuals (those with high levels of outgroup antipathy) saw no statistically significant change in empathy vis-à-vis the control. Still, we can reject the null for this hypothesis with strong confidence, as reading an empathetic article about refugees, on average, produced a highly significant increase of 0.27 on the empathy scale.

H2) Empathetic Media Encourages More Open Policy Opinions and Pro-Refugee Action

The behavior results suggest that empathetic treatments are not very effective in modifying political behavior for the British public at large, but that they can have some results for subsets of the population. Interestingly, the same treatments that are the least effective in mobilizing pro-refugee action for low-antipathy individuals (that is, articles about Christian refugees) are the most effective in dampening anti-refugee action in the much larger group of individuals with moderate amounts of outgroup antipathy. More generally, these results suggest that opposition toward refugees in the general public is best palliated by media that present those refugees as sharing a dominant cultural trait with the receiving country (in this case, Christianity). Meanwhile, the subset of individuals that are already very open to refugees are most likely to act when primed with information about refugees that do *not* match wider societal norms, perhaps out of a feeling that these individuals in a religious or other minority are the most in need of active advocacy.

H3) Treatment Effects Are Conditional on Pre-Existing Levels of Outgroup Antipathy

Differences in treatment magnitude and significance according to respondents' levels of prior outgroup antipathy were the most consistent findings across all three outcome variables. In all cases, the high-antipathy group was the most impervious to change in aggregate emotions, opinions, or behavior after reading the treatment articles. While the small sample size of high-antipathy individuals means we should not immediately assume there is absolutely no treatment effect for this group (after all, mean opinions and empathy in the Christian and No Religion treatment groups were higher than the control average for this group), any effect is likely to be relatively small.

On the other hand, the largest effect sizes took place among low-antipathy individuals. Within the placebo control group, these individuals had a similar level of latent empathy levels as any other antipathy group, but they had the highest levels of empathy among any treated individuals. While this group was already the most likely to favor open refugee policies in the control condition, they evinced even stronger opinions and action in favor of refugees after reading the empathetic media article. Predictably, the results for mid-level antipathy individuals fall in between those of the high- and low-antipathy groups.

These findings support the expectations of Hypothesis 3. However, they do not indicate which (if any) of the mechanisms proposed above are responsible for the patterns we see here. Feelings of dissonance increased slightly for all individuals, although the patterns were slightly different for low- and high-antipathy groups. Furthermore, the measurement of dissonance itself does not explain *why* individuals reacted with those feelings. The negative feelings reported may have been due to a clash of new information with individuals' prior beliefs, but they may also be evidence of distress felt in response to the difficult situations reported in the treatment articles. Either or both of these pathways may have led to greater feelings of dissonance, and the pathway itself may have

varied according to prior levels of antipathy. Determining what theory best explains the phenomena capsulated by Hypothesis 3 will be a task for further research.

H4) Citizen Reactions Show Differences According to Refugee Religion

All three outcomes showed differing treatment effects according to refugee religion, although these differences were sometimes limited to specific antipathy subgroups. Overall, the results showed patterns of Islamophobia; high-antipathy individuals felt significantly less empathy after reading about Muslim refugees than they did after reading about Christians, and the Muslim treatments were the only treatments that did not elicit more open refugee policy opinions for the large group of mid-antipathy respondents (75 percent of the sample). For the behavior outcome, aggregate opposition only decreased for most individuals when confronted with Christian refugees, the group that shared Britain's dominant religion.

No significant differences existed between Sudanese and Syrian treatments for opinions and empathy, although some positive differences were visible in the behavior measure. These results within the fully crossed experimental treatment design confirm the suppositions of past authors that religion, not national origin, is indeed the more salient factor driving public opinion about refugee and immigrant admission (Sniderman et al. 2004; Iyengar et al. 2013).

Conclusion

The results of this study are preliminary, and questions linger over how responses may have differed if treatments varied additional refugee characteristics or if the experimental sample had been truly representative of the UK population. Furthermore, the comparison between Sudan and Syria alone does not allow us to determine what underlying opinions about those countries (apart from religion) are at work in influencing attitudes. Finally, we do not know exactly why respondents' pre-existing attitudes influence their appraisals of empathetic appeals, although a clear relationship exists. Still, we can maintain a high level of certainty in attributing causality to the relationship between empathetic media portrayals and respondent's feelings, opinions, and actions.

This experiment proposed to address several limitations in the current literature on attitudes toward refugees and immigrants. First, this study focuses specifically on refugees, who migrate for different reasons than typical economic immigrants and may elicit differing responses from citizens in receiving countries (Verkuyten 2004; Bansak et al. 2016). The results indicate that empathetic media accounts of refugees raise general feelings of empathy and shift opinions about refugee policy toward greater admittance. While methods of demonstrating variable mediation are contested and are not presented here (after all, empathy was not manipulated separately from the treatment itself), it is possible that positive emotions, such as empathy, drive changes in opinion and political behavior, as other scholars have shown happens with negative emotions (Brader, Valentino, and Suhay 2008; Albertson and Gadarian 2015). Regardless of the psychological pathway this effect takes, these results show that fear

and anxiety are not the only emotions that can sway the court of public opinion. Purveyors of media that empathizes with or humanizes outgroups may have confidence that their accounts are persuasive for the general public, although they may have less efficacy in promoting political action.

Second, by varying both the religion and the national origin of the refugee family in the treatment prime, this study allows us to distinguish between two cultural factors that may elicit differing responses in receiving citizens. This improves on past research that has only looked at variation in one or the other of these characteristics. Notably, accounts about Muslim refugees were the least effective in shifting opinions about refugee issues. This demonstrates the difficulty that opinion leaders face in combating religious prejudice in public opinion about immigration. It also suggests that, at least in the UK, the best way to encourage positive perceptions about refugees is to either emphasize a trait that the refugees share with the local populace (such as Christianity) or to avoid the topic of religion altogether.

Finally, this paper explores how empathetic media appeals about outgroups influence citizens' emotions, opinions, and actions. Beyond addressing the general lack of research on how positive affect influences opinions on refugees, this experiment allows us to distinguish between the empathy and dissonance felt by experimental subjects and to observe patterns between these feelings and individuals' prior outgroup antipathy. These results show that, for the general population, empathetic appeals successfully encourage positive affect and opinions about refugees. However, these appeals have little effect for people who feel the most negatively about refugees. This is especially concerning given that people with outgroup antipathy would be clear targets for empathetic appeals. Thus, while empathetic appeals may preach to the choir that already supports refugees and even win over a section of the public without strong set opinions, they are unlikely to change the feelings of the most prejudiced subset of individuals.

These individuals' attitudinal intransigence suggests competing policy responses: On one hand, opinion leaders seeking to promote greater empathy for refugees might need to recalibrate drastically their messages if they seek to persuade this group. Alternatively, they might develop messages that do not even attempt to sway the possibly unswayable but instead focus their energies motivating those who already favor refugee resettlement to advocate for them more vociferously (as was shown to be possible in these results). The results of this experiment were one-time and short-term; it is a possibility that a longer media campaign may be effective in wearing down hardened opinions against refugee resettlement. However, regardless of the tack one takes in responding to these individuals, the fact that a sizable percentage of the British public harbors unyielding prejudicial opinions about refugees is worrying. Its implications for refugee and immigration policy, as well as for democratic politics in increasingly diverse Western nations, should be the topic of further research.

APPENDIX 1

Additional Results, Tables and Figures

Randomization checks conducted using Wald tests to search for significant differences in observable traits between all treatment groups. Tests covered all demographic and psychological traits with more than one hundred observations. Of 357 pairwise comparisons, 2.8 percent were significant at the 95 percent level, well within levels of what would be expected by random chance.

Factor Loadings for the Empathy and Dissonance Scales

Factor analysis with varimax rotation of the fifteen emotion variables yielded two unambiguous factors, one which I call empathy ($\lambda = 5.58$) and one which I call dissonance ($\lambda = 2.21$). Both of these factors corresponded to my *a priori* expectations. A third factor, which I call positive affect, had an Eigen value of 0.80. Using the highest-loading emotions within each factor, I constructed two scales. The empathy scale ($\alpha = 0.9254$) includes the emotions "compassion," "warm," "softhearted," "tender," "sympathetic," and "moved." The dissonance scale ($\alpha = 0.8642$) includes "bothered," "tense," "angry," "anxious," "afraid," and "uncomfortable." Anger was the only emotion included on this scale that was not expected in the experimental design phase. Results for dissonance change very little with or without the inclusion of anger in the scale.

Regression Results of Treatment Effects by Nationality and Religion

As discussed in the main body of the paper, there are no significant differences between the Syrian and Sudanese coefficients for any outcome. For supportive action, the Sudanese treatment effect is statistically significant while the Syrian treatment effect is not. However, the confidence intervals around the coefficients overlap greatly. Tables 2 and 3 give the OLS regression results for the empathy and opinion outcomes, respectively. Table 4 gives multinomial logistic regression results for the behavior outcomes. Both overall sample estimates and conditional estimates are provided.

Tables 5–8 show regression results for models that test for religious treatment effects. Both overall sample estimates and conditional estimates are provided.

Table 1: Sample Demographic Characteristics

Gender	%	Marital Status	%
Female	50.46%	Married	38.63%
Male	49.54%	Single	31.81%
		Cohabitating	21.62%
Race		Divorced	3.38%
White	92.47%		
Black	1.27%	Annual Income	
Asian	4.01%	< £20,000	28.80%
Other/Mixed	2.32%	< £40,000; ≥ £20,000	38.24%
		< £70,000; ≥ £40,000	24.65%
Age	36.77 (mean)	≥ £70,000	7.30%

Table 2: Nationality Treatment effects on Empathy

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Syrian	0.269*** (0.0200)	0.0187 (0.0673)	0.265*** (0.0224)	0.424*** (0.0424)
Sudanese	0.271*** (0.0200)	0.101 (0.0678)	0.245*** (0.0225)	0.469*** (0.0415)
Constant	0.244*** (0.0173)	0.238*** (0.0592)	0.258*** (0.0197)	0.203*** (0.0344)
Observations	1,419	130	1,059	230
R-squared	0.128	0.031	0.123	0.378

OLS Regression with the placebo control used as the comparison condition. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Nationality Treatment effects on Opinion

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Syrian	0.0524** (0.0250)	0.0662 (0.0519)	0.0742*** (0.0267)	0.101** (0.0439)
Sudanese	0.0664*** (0.0250)	0.0275 (0.0522)	0.0888*** (0.0268)	0.0986** (0.0430)
Constant	0.433*** (0.0217)	0.0588 (0.0457)	0.404*** (0.0234)	0.667*** (0.0356)
Observations	1,419	130	1,059	230
R-squared	0.128	0.031	0.123	0.378

OLS Regression with the placebo control used as the comparison condition. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Multinomial Logistic Regression of Behavior on National Treatments

Action to Oppose the Bill

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Syrian	-0.284 (0.231)	-0.288 (0.568)	-0.471* (0.275)	0.405 (0.913)
Sudanese	-0.222 (0.232)	-0.298 (0.571)	-0.395 (0.276)	0.480 (0.914)
Constant	-1.192*** (0.196)	-0 (0.500)	-1.241*** (0.232)	-2.351*** (0.740)

Action to Support the Bill

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Syrian	-0.0722 (0.189)	-0.693 (1.287)	-0.0776 (0.234)	0.629 (0.383)
Sudanese	0.205 (0.186)	-1.355 (1.469)	0.207 (0.231)	0.915** (0.382)
Constant	-0.693*** (0.164)	-2.079** (1.061)	-0.922*** (0.206)	0.0465 (0.305)
Observations	1,421	130	1,061	230
Pseudo R-squared	0.002	0.005	0.004	0.016

Multinomial logistic regression with "No action" as the omitted outcome category and the placebo control used as the comparison condition. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Religion Treatment Effects on Empathy

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Christian	0.272*** (0.0213)	0.112 (0.0693)	0.259*** (0.0237)	0.430*** (0.0464)
Muslim	0.256*** (0.0212)	-0.0450 (0.0721)	0.241*** (0.0237)	0.445*** (0.0443)
No Religion	0.282*** (0.0212)	0.0873 (0.0695)	0.266*** (0.0237)	0.467*** (0.0459)
Constant	0.244*** (0.0173)	0.238*** (0.0583)	0.258*** (0.0197)	0.203*** (0.0345)
Observations	1,419	130	1,059	230
R-squared	0.129	0.070	0.123	0.376

OLS Regression with the placebo control used as the comparison condition. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Religion Treatment Effects on Dissonance

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Christian	0.0771*** (0.0174)	0.0860* (0.0518)	0.0645*** (0.0213)	0.101*** (0.0348)
Muslim	0.0685*** (0.0173)	0.110** (0.0539)	0.0591*** (0.0213)	0.0653* (0.0332)
No Religion	0.115*** (0.0173)	0.136*** (0.0520)	0.0971*** (0.0212)	0.161*** (0.0344)
Constant	0.135*** (0.0142)	0.0858* (0.0435)	0.159*** (0.0176)	0.0806*** (0.0259)
Observations	1,418	130	1,058	230
R-squared	0.031	0.054	0.020	0.094

OLS Regression with the placebo control used as the comparison condition. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Religion Treatment Effects on Opinions

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Christian	0.0637** (0.0265)	0.0692 (0.0544)	0.0891*** (0.0282)	0.124** (0.0478)
Muslim	0.0350 (0.0265)	0.0115 (0.0566)	0.0409 (0.0282)	0.0869* (0.0457)
No Religion	0.0798*** (0.0265)	0.0537 (0.0546)	0.114*** (0.0281)	0.0918* (0.0473)
Constant	0.433*** (0.0217)	0.0588 (0.0457)	0.404*** (0.0233)	0.667*** (0.0356)
Observations	1,419	130	1,059	230
R-squared	0.008	0.021	0.020	0.030

OLS Regression with the placebo control used as the comparison condition. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 8: Multinomial Regression of Behavior on Religious Treatments

Action to Oppose the Bill

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Christian	-0.369 (0.249)	-0.201 (0.592)	-0.661** (0.303)	0.454 (0.965)
Muslim	-0.266 (0.247)	-0.511 (0.619)	-0.364 (0.291)	0.560 (0.968)
No Religion	-0.126 (0.246)	-0.211 (0.597)	-0.292 (0.293)	0.272 (1.054)
Constant	-1.192*** (0.196)	-7.37e-06 (0.500)	-1.241*** (0.232)	-2.351*** (0.740)

Action to Support the Bill

VARIABLES	All	High Antipathy	Mid Antipathy	Low Antipathy
Christian	-0.0994 (0.200)	-1.012 (1.473)	-0.0502 (0.246)	0.454 (0.416)
Muslim	0.0504 (0.198)	-14.22 (772.1)	-0.0551 (0.247)	0.955** (0.411)
No Religion	0.252 (0.197)	-0.272 (1.293)	0.295 (0.242)	0.894** (0.424)
Constant	-0.693*** (0.164)	-2.079** (1.060)	-0.922*** (0.206)	0.0465 (0.305)
Observations	1,421	130	1,061	230
Pseudo R-squared	0.003	0.018	0.005	0.020

Multinomial logistic regression with "No action" as the omitted outcome category and the placebo control used as the comparison condition. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Linear Analysis of Treatment Interactions with Prior Outgroup Antipathy

Linear analysis of treatment effects by refugee nationality shows the same patterns as conditional analysis. No significant differences in response to Syrian and Sudanese refugees are visible. For supportive action, the Sudanese treatment effect is statistically significant while the Syrian treatment effect is not. However, the confidence intervals around the coefficients overlap greatly. Table 9 gives regression results for all three outcome variables. Figure A1 shows the expected values for empathy across all antipathy levels. Figure A2 shows the expected rates of support across all antipathy levels.

Table 9: Linear regression of national treatment effects on Empathy, Opinion, and Political Behavior

Behavior Outcomes

VARIABLES	Empathy	Opinion	Action to Oppose	Action to Support
Antipathy	0.111 (0.0701)	-0.833*** (0.0742)	5.032*** (1.190)	-3.527*** (0.875)
Syrian	0.561*** (0.0421)	0.123*** (0.0445)	0.00956 (0.876)	0.659 (0.443)
Sudanese	0.549*** (0.0417)	0.118*** (0.0441)	0.283 (0.878)	1.325*** (0.450)
Syrian × Antipathy	-0.617*** (0.0826)	-0.0965 (0.0874)	-0.719 (1.390)	-1.445 (1.057)
Sudanese × Antipathy	-0.608*** (0.0827)	-0.0953 (0.0875)	-1.003 (1.404)	-2.541** (1.076)
Constant	0.194*** (0.0354)	0.808*** (0.0374)	-3.991*** (0.743)	0.665* (0.362)
Observations	1,418	1,418		
R-squared	0.266	0.401		
Pseudo R-squared			0.153	0.153

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Empathy and Opinion results come from OLS regressions. Behavior results come from a single multinomial logistic regression with "no action" as the baseline.

Figure A1: Expected Values of Empathy by Refugee Nationality
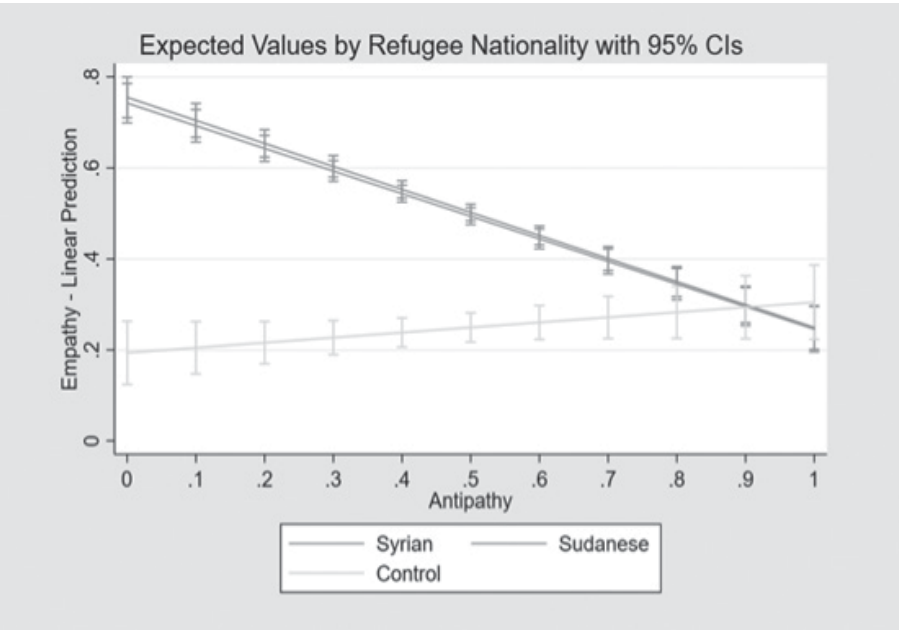
Figure A2: Expected Rates of Supportive Action by Refugee Nationality
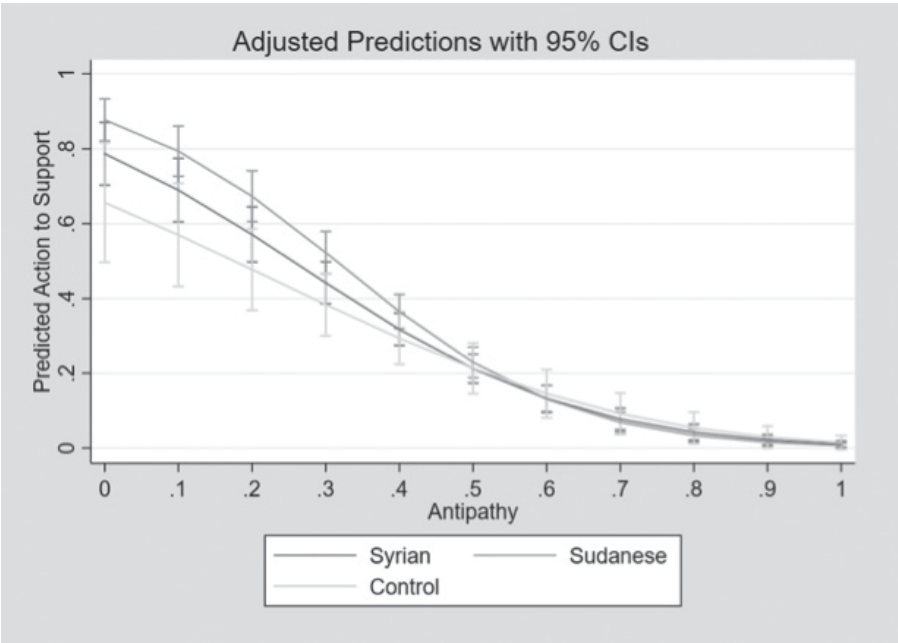
Results from analysis with antipathy as a continuous variable mirror the conditional analysis reported in the main body of the paper. Table 10 gives regression results for all outcomes. Treatment increases empathy, but to a decreasing degree as antipathy increases. Treatment also moves opinions toward supporting greater refugee admission, although anti-Muslim bias is less evident than in the conditional results. Null treatment effects are observed for opposing action to the refugee bill, but significant effects from the Muslim and no religion treatments are visible for supporting action. A negative interaction effect between the Muslim treatment and prior antipathy suggests that, when considering Muslim refugees, low-antipathy individuals are the most likely to actively support refugee admission, while high-antipathy individuals are especially likely to actively oppose it. Figures A3, A4, and A5 show the expected values for empathy, opinions, and supportive behavior, respectively.

**Table 10: Linear Regression of Religion Treatment Effects on Empathy,
Opinion, and Political Behavior**

VARIABLES	Empathy	Opinion	Action to Oppose	Action to Support
Antipathy	0.111 (0.0701)	-0.833*** (0.0739)	5.032*** (1.190)	-3.527*** (0.875)
Christian	0.530*** (0.0451)	0.146*** (0.0476)	-0.439 (0.954)	0.713 (0.484)
Muslim	0.573*** (0.0444)	0.116** (0.0468)	0.728 (0.925)	1.177** (0.479)
No Religion	0.558*** (0.0447)	0.101** (0.0472)	0.164 (0.932)	1.106** (0.485)
Christian × Antipathy	-0.551*** (0.0886)	-0.131 (0.0935)	-0.0967 (1.504)	-1.726 (1.160)
Muslim × Antipathy	-0.685*** (0.0876)	-0.147 (0.0924)	-1.847 (1.483)	-2.560** (1.149)
No Religion × Antipathy	-0.595*** (0.0881)	-0.0100 (0.0930)	-0.688 (1.481)	-1.753 (1.143)
Constant	0.194*** (0.0353)	0.808*** (0.0373)	-3.991*** (0.743)	0.665* (0.362)
Observations	1,418	1,418	1,420	1,420
R-squared	0.269	0.406		
Pseudo R-squared			0.155	0.155

Figure A3: Expected Values of Empathy by Refugee Religion
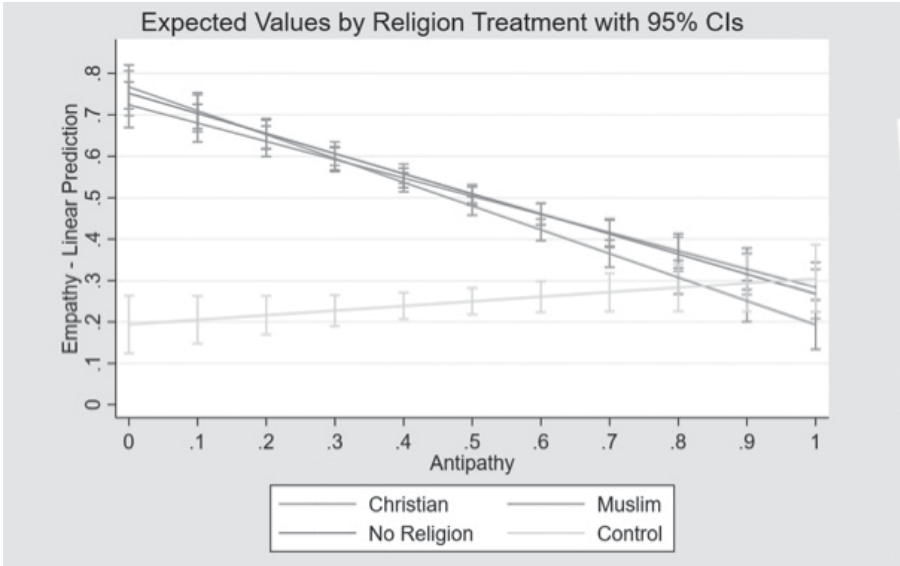
Figure A4: Expected Values of Opinions by Refugee Religion
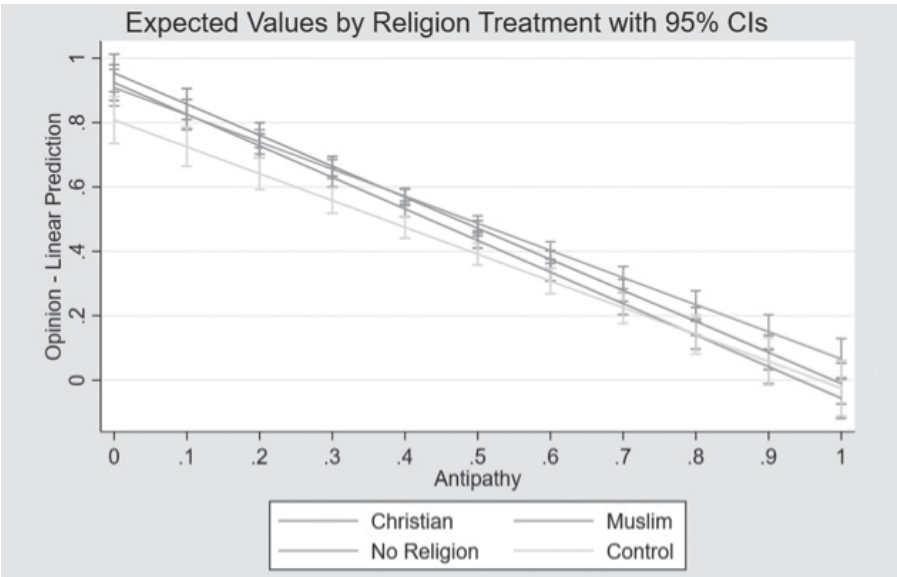
Figure A5: Expected Rates of Supportive Action by Refugee Nationality
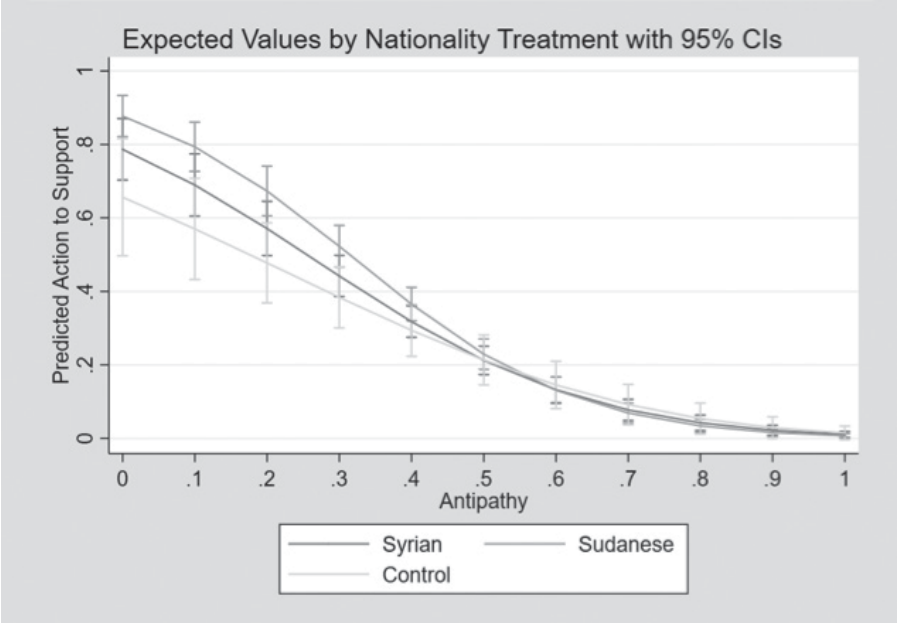
APPENDIX 2

Question Wording and Treatments

Outgroup Antipathy Measure

The following three questions were interspersed with distractor questions on other issues and then used to compose the outgroup antipathy scale. Respondents could answer using a Seven-point Likert scale ranging from "Strongly disagree" to "Strongly agree."

Below are a number of statements that you may or may not agree with. Please indicate to what degree you agree or disagree with each statement. There are no right or wrong answers to these questions. We're simply interested in your honest responses.

- Refugees are more prone to violence than other groups
- Providing increased opportunities (jobs, education) for refugees in the UK means decreasing opportunities for other residents.
- Refugees have moral values and customs from which UK residents could learn.

Treatment Articles

The following article was shown to individuals in the Syrian Christian treatment. In Sudanese treatments, all instances of "Syria," "Syrian," and "Homs" were replaced with "Sudan," "Sudanese," and "Karmuk," respectively. In Muslim treatments, all instances of "Christian," were replaced with "Muslim," or, in the eighth paragraph, "Islamic." All instances of "Jesus," "church," and "Bible" were replaced with "Allah," "mosque," and "Quran," respectively. In the no religion treatments, all religious vocabulary, along with the two sentences in the eighth paragraph that detail the family's worship practices, were omitted. Names varied in all treatments in order to realistically match names common within each religion and national community.

Please carefully read the following news article. You will be asked questions later about its contents:

"How one Christian refugee family from Syria is starting over in London"

Susanna Assali, a recently arrived Christian refugee, talks about what happened to her brother in Homs, Syria. He was pulled over at a military checkpoint, accused of being an anti-government activist, she says. He was tortured and nearly killed.

Her husband, Yakob Olikara also has his reasons — what happened to his brothers. One was killed by a landmine. Another brother went missing.

"We were looking in the hospitals and police stations," he says. "All of a sudden, a friend of ours called and he said, 'There's someone on the street who looks like your brother.'" It was. His brother had been tossed on the street dead, bound by his hands and feet. He worked at a sugar factory when Syrian soldiers — arbitrarily, Olikara says — rounded him and others up, suspecting them of being on the wrong side of the civil war. Olikara says his brother stayed clear of politics, but the violence is indiscriminate.

The next day, the couple fled with their young children.

Now, everything's different. In February, they arrived in London. Their journey was long, three years in all, after living in a refugee camp in Jordan. There were numerous interviews and security checks until, finally, they got word that they would be resettled in the United Kingdom.

"It's utter and pure luck," says Karen Thompson, executive director of the International Rescue Committee (IRC) in London.

Olikara says he and his family would have been happy anywhere they felt safe. "Sweden, the United States, Britain, wherever. We are so thankful to be here," he says. The couple says they believe Jesus has helped them to make this new life possible.

The family's only material possessions from Syria are a few photos. The couple's 9-year-old daughter, Sara, keeps them tidy in a small album. Pictures show the relatives left behind: Assali's mom and her two brothers, who are disabled and can't make the journey out of Syria.

It's a reality that weighs on the parents, but, like other Syrian Christian refugees, they're trying to focus on building a life here. They attend services regularly at their local church, and this has helped them make friends who share their Christian faith. They also enjoy studying the Bible once a week with other members of their community.

Olikara, who was a farmer and truck driver before, got a job fast as a cook at a local restaurant. The family also wants to meet people, and people here want to meet them. Their neighbor, Jared Park, who stopped by their flat on a recent afternoon, remembers seeing the family when they first moved in. "They're lonely," says Park "They don't know anybody around here. They don't have the language. Nothing, nothing."

But both parents say they want to return to Syria one day, when it's stable. Where they will live, though, is another question. They just got bad news about their home. "I received some pictures. Destroyed," says Olikara.

So they're focused on being here, and exploring the countryside in the family's first big purchase: a mini-van. A friend at their Christian church gave them a deal.

Tomorrow, Olikara takes the driver's test. It'll be a challenge for him, but nothing compared to what he and his family have survived.

Placebo Control Article

The following article was shown to all individuals in the control condition.

Please carefully read the following news article:

"Will 100% fruit juice make your child gain weight?"

Sugar can easily sneak into the diet, both for you and for your child, even through 100% fruit juices.

Many health experts have even expressed concerns that the content of naturally occurring sugars in such juices can have negative health effects on children, such as increasing the risk for obesity.

The relationship between 100% fruit juice consumption and weight gain has been analyzed in a study published in the journal Pediatrics on Thursday.

The study suggests that drinking 100% fruit juice is associated with a slight amount of weight gain in children 6 and younger who have one serving a day, but no association was found for children 7 and older who have one serving a day.

Yet the study has some limitations, and it recommends drinking 100% fruit juice only in moderation.

"I think caution is definitely in order and that when possible, parents should give whole fruit to kids, instead of fruit juice," said Dr. Brandon Auerbach, a primary care physician and instructor at the University of Washington's Division of General Internal Medicine in Seattle. "Water or low-fat unsweetened milk are other good alternatives to 100% fruit juice."

Auerbach, lead author of the new study, said he is the parent of a 7-month-old boy who soon may be offered fruit juices in day care and at school.

"I share the concern that 100% fruit juices have a lot of sugar, even though it's naturally occurring sugar," he said. "There are other health concerns about drinking 100% fruit juice, besides weight gain, especially related to risk of cavities and risk of future metabolic syndrome or diabetes."

The study was a systemic review and meta-analysis of eight previous observational studies on 100% fruit juice consumption and weight gain among children, based on their body mass indexes or BMI.

The analysis showed that consuming 100% fruit juice was slightly associated with weight gain in children 1 to 6 years old, but not enough to potentially harm health, Auerbach said.

The researchers found that in children ages 1 to 6, consumption of one daily serving was associated with a weight gain of 0.3 pounds or less over one year. In children 7 and older, 100% fruit juice was not independently associated with any weight gain.

"I was somewhat surprised by the results, given that some types of 100% fruit juice have comparable amounts of sugar as regular soda," Auerbach said.

He added, however, that the study certainly had some limitations.

"Although we combined evidence from the best available research, the studies were not randomized controlled trials," Auerbach said. "We did not examine other important health outcomes besides weight gain, such as diabetes risk, because too few studies exist on this topic in children."

The American Academy of Pediatrics recommends that children ages 1 to 6 drink no more than 4 to 6 ounces of 100% fruit juice a day. For children 7 to 18, juice intake should be limited to 8 to 12 ounces, or 2 servings, per day.

It's not recommended to give fruit juices to infants, Auerbach said. He concluded, "Our study findings support the current guidelines of the American Academy of Pediatrics on 100% fruit juice consumption."

As for adults, the US Dietary Guidelines for Americans recommend drinking no more than one cup, or 8 ounces, of 100% fruit juice a day.

Behavior Measures

The procedures section of this paper details the administration of the emotion and opinion outcomes. The following gives the exact wording of the behavior question:

There is currently a bill proposed before Parliament that would allow family members of refugees to enter or remain in the United Kingdom. The bill would also make available legal aid for these family reunion cases. Would you like us to send an email message to members of Parliament indicating either your support or opposition to this reform?

The text of the e-mail would read:

Dear MP,

It has come to my attention that a bill, the Refugees (Family Reunion) (No. 2) Bill, has been proposed that will allow refugees' family members to enter or remain in the UK. I (oppose/support) this bill and urge you to do the same.

Sincerely,

Your constituent

- Yes, please send a notice indicating my support for admitting the family members of refugees.
- Yes, please send a notice indicating my opposition to admitting the family members of refugees.
- No, please do not send a notice on my behalf.

REFERENCES

Adesina, Zack, and Oana Marocico. 2015. "Islamophobic Crime in London 'Up by 70%'." BBC, September 7. Accessed July 25, 2017. www.bbc.com/news/uk-england-london-34138127.

Adida, Claire L., David D. Laitin, and Marie-Anne Valfort. 2010. "Identifying Barriers to Muslim Integration in France." *Proceedings of the National Academy of Sciences* 107, no. 52: 22,384–90.

Adida, Claire L., Adeline Lo, and Melina Platas. 2017. "Engendering Empathy, Begetting Backlash: American Attitudes toward Syrian Refugees." *Immigration Policy Lab Working Paper Series,* No. 17-01 (May). https://ssrn.com/abstract=2978183.

Albertson, Bethany, and Shana Gadarian. 2015. *Anxious Politics: Democratic Citizenship in a Threatening World.* Cambridge: Cambridge University Press.

Bansak, Kirk, Jens Hainmueller, and Dominik Hangartner. 2016. "How Economic, Humanitarian, and Religious Concerns Shape European Attitudes toward Asylum Seekers." *Science* 354, no. 6309: 217–22.

Batchelor, Tom. 2017. "London Terror Attack: Huge Rise in Islamophobic Hate Crime Following Borough Market Stabbing, Police Figures Show." *The Independent*, June 7. Accessed July 25, 2017. www.independent.co.uk/News/uk/crime/london-bridge-attack-latest-rise-islamophobic-hate-crimes-borough-market-stabbing-terrorpolice-a7777451.html.

Batson, C. Daniel, Jim Fultz, and Patricia A. Schoenrade. 1987. "Distress and Empathy: Two Qualitatively Distinct Vicarious Emotions with Different Motivational Consequences." *Journal of Personality* 55, no. 1 (March): 19–39.

Batson, C. Daniel, Johee Chang, Ryan Orr, and Jennifer Rowland. 2002. "Empathy, Attitudes, and Action: Can Feeling for a Member of a Stigmatized Group Motivate One to Help the Group?" *Personality and Social Psychology Bulletin* 28, no. 12 (December): 1,656–66.

Batson, C. Daniel, Marina P. Polycarpou, Eddie Harmon-Jones, Heidi J. Imhoff, Erin C. Mitchener, Lori L. Bednar, Tricia R. Klein, and Lori Highberger. 1997. "Empathy and Attitudes: Can Feeling for a Member of a Stigmatized Group Improve Feelings toward the Group?" *Journal of Personality and Social Psychology* 72, no. 1 (January): 105–18.

Berinsky, A.J., Huber, G.A. and Lenz, G.S. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20, no. 3 (Summer): 351–68.

Brader, Ted, Nicholas A. Valentino, and Elizabeth Suhay. 2008. "What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration Threat." *American Journal of Political Science* 52, no. 4 (October): 959–78.

Campbell, Monica. 2015. "How One Refugee Family from Syria is Starting Over in California." Public Radio International, September 21. Accessed April 7, 2015. www.pri.org/stories/2015-09-21/how-one-refugee-family-syria-starting-over-california.

Coke, Jay S., C. Daniel Batson, and Katherine McDavis. 1978. "Empathic Mediation of Helping: A Two-Stage Model." *Journal of Personality and Social Psychology* 36, no. 7 (July): 752–66.

Eisenberg, Nancy and Paul A. Miller. 1987. "The Relation of Empathy to Prosocial and Related Behaviors." *Psychological Bulletin* 101, no. 1: 91–119.

Eisenberg, Nancy, Cindy L. Shea, Gustavo Carlo, and George P. Knight. 2014. "Empathy-Related Responding and Cognition: A "Chicken and the Egg" Dilemma." In *Handbook of Moral Behavior and Development*, edited by William M. Kurtines and Jacob L. Gewirtz, 63–88. Vol. 2, *Research*. New York: Psychology Press.

Elliot, Andrew J. and Patricia G. Devine. 1994. "On the Motivational Nature of Cognitive Dissonance: Dissonance as Psychological Discomfort." *Journal of Personality and Social Psychology* 67, no. 3 (September): 382–94.

Gadarian, Shana Kushner, and Bethany Albertson. 2014. "Anxiety, Immigration, and the Search for Information." *Political Psychology* 35, no. 2 (April): 133–64.

Gerdes, Karen E., Elizabeth A. Segal, and Cynthia A. Lietz. 2010. "Conceptualising and Measuring Empathy." *British Journal of Social Work* 40, no. 7 (October): 2,326–43.

Gubler, Joshua R. 2016. "When Humanizing the Enemy Fails: The Role of Dissonance and Justification in Intergroup Conflict." Working Paper.

Gubler, Joshua R., Christopher F. Karpowitz, J. Quin Monson, David A. Romney, and Mikle South. 2016. "Erasing Empathy: How Dissonance Disrupts Attempts to Humanize Outgroups." Working Paper.

Hainmueller, Jens and Daniel J. Hopkins. 2015. "The Hidden American Immigration Consensus: *A Conjoint Analysis of Attitudes toward Immigrants*." American Journal of Political Science 59, no. 3 (July): 529–48.

Harell, Allison, Stuart Soroka, Shanto Iyengar, and Nicholas Valentino. 2012. "The Impact of Economic and Cultural Cues on Support for Immigration in Canada and the United States." *Canadian Journal of Political Science 45*, no. 3 (September): 499–530.

Harmon-Jones, Eddie, David M. Amodio, and Cindy Harmon-Jones. 2009. "Action-based Model of Dissonance: A Review, Integration, and Expansion of Conceptions of Cognitive Conflict." In *Advances in Experimental Social Psychology*, edited by M.P. Zanna, 119–66. Vol. 41. San Diego: Elsevier Academic Press.

Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. "The Online Laboratory: Conducting Experiments in a Real Labor Market." *Experimental Economics 14*, no. 3 (September): 399–425.

Howard, Jaqueline. 2017. "Will 100% Fruit Juice Make Your Child Gain Weight?" CNN, March 23. Accessed April 7, 2017. www.cnn.com/2017/03/23/health/fruit-juice-weight-gain-kids-study/index.html.

Igartua, Juan-José, and Lifen Cheng. 2009. "Moderating Effect of Group Cue While Processing News on Immigration: Is the Framing Effect a Heuristic Process?" *Journal of Communication* 59, no. 4 (December): 726–49.

Iyengar, Shanto, Simon Jackman, Solomon Messing, Nicholas Valentino, Toril Aalberg, Raymond Duch, Kyu S. Hahn, Stuart Soroka, Allison Harell, and Tetsuro Kobayashi. 2013. "Do Attitudes about Immigration Predict Willingness to Admit Individual Immigrants? A Cross-National Test of the Person-Positivity Bias." *Public Opinion Quarterly* 77, no. 3 (Fall): 641–65.

Lecheler, Sophie, Linda Bos, and Rens Vliegenthart. 2015. "The Mediating Role of Emotions: News Framing Effects on Opinions about Immigration." *Journalism & Mass Communication Quarterly* 92, no. 4 (July): 812–38.

Lodge, Milton and Charles S. Taber. 2005. "The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis." *Political Psychology* 26, no. 3 (June): 455–82.

Lodge, Milton and Charles S. Taber. 2013. *The Rationalizing Voter*. Cambridge: Cambridge University Press.

Marcus, George E. and Michael B. MacKuen. 1993. "Anxiety, Enthusiasm and the Vote: The Emotional Underpinnings of Learning and Involvement during Presidential Campaigns." *American Political Science Review* 87, no. 3 (September): 672–85.

Ostfeld, Mara. 2017. "The Backyard Politics of Attitudes toward Immigration." *Political Psychology* 38, no. 1 (February): 21–37.

Peer, Eyal, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research." *Journal of Experimental Social Psychology* 70 (May): 153–63.

Redlawsk, David P. 2002. "Hot Cognition or Cool consideration? Testing the Effects of Motivated Reasoning on Political Decision Making." *Journal of Politics* 64, no. 4 (November): 1,021–44.

Sherwood, Harriet. 2016. "Incidents of Anti-Muslim Abuse Up by 326% in 2015, Says Tell MAMA." *The Guardian*, June 29. Accessed July 25, 2017. www.theguardian.com/society/2016/jun/29/incidents-of-anti-muslim-abuse-up-by-326-in-2015-says-tell-mama.

Sniderman, Paul M., Louk Hagendoorn, and Markus Prior. 2004. "Predisposing Factors and Situational Triggers: Exclusionary Reactions to Immigrant Minorities." *American Political Science Review* 98, no. 1 (February): 35–49.

Taber, Charles S. and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50, no. 3 (July): 755–69.

Tajfel, Henri. 1982. "Social Psychology of Intergroup Relations." *Annual Review of Psychology* 33: 1–39.

Tajfel, H. and Turner, J.C. 1979. "An Integrative Theory of Intergroup Conflict." In *The Social Psychology of Intergroup Relations*, edited by William G. Austin and Stephen Worchel. Monterey, CA: Brooks/Cole.

UK Home Office. 2017. *Immigration Statistics—October to December 2016: Asylum, Volume 1* by Chris Kershaw. Last modified February 23. Accessed November 29, 2017. www.gov.uk/government/statistics/immigration-statistics-october-to-december-2016-data-tables.

Valentino, Nicholas A., Ted Brader, Eric W. Groenendyk, Krysha Gregorowicz, and Vincent L. Hutchings. 2011. "Election Night's Alright for Fighting: The Role of Emotions in Political Participation." *Journal of Politics* 73, no. 1 (January): 156–70.

Verkuyten, Maykel. 2004. "Emotional Reactions to and Support for Immigrant Policies: Attributed Responsibilities to Categories of Asylum Seekers." *Social Justice Research* 17, no. 3 (September): 293–314.

Watson, David, Lee Anna Clark, and Auke Tellegen. 1988. "Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales." *Journal of Personality and Social Psychology* 54, no. 6 (June): 1,063–70.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

The Price of Scottish Independence: Why Remaining a Part of the UK Still Makes Sense in the Wake of Brexit

Lincoln Wilcox

Introduction

On 23 June 2016, the citizens of the UK voted to leave the EU in a historic referendum. The "Leave" camp obtained a mostly unexpected victory by a narrow margin of 52 percent to 48 percent. The results shocked the EU, since the UK was one of its most powerful member states. Analysts began providing various explanations as to how the "Brexit" would occur and what it would mean for the UK and the European community.

Although Great Britain as a whole voted to leave, a regional breakdown of the voting revealed a more complicated story. Majorities voted to leave the EU in both England and Wales—but Northern Ireland, and especially Scotland, voted strongly in support of remaining. Scotland in particular had a 62 percent majority vote to stay, suggesting the Scot's interests diverged from those of their southern neighbors.

These results are particularly significant given the recent independence referendum held in Scotland in September 2014 to decide whether Scotland should become an independent country, which resulted in a victory for the proponents of staying a part of the UK. The margin of victory, though clear, was not high, with 55 percent voting to stay and nearly 45 percent advocating Scotland's exit from Great Britain. The high voter turnout of 84 percent of the population signaled that people broadly and passionately cared about the consequences.

The results of the Brexit referendum may have changed the game for many Scottish voters. In the previous Scottish independence referendum, Scots decided whether the benefits of independence outweighed the costs of leaving the UK; however, they made this decision under the assumption of remaining a part of the EU in either case. Scottish voters have demonstrated their desire to remain a part of the EU, and Brexit forces them to weigh this desire against the costs of leaving the UK. Scottish politicians have, there-

fore, declared that Scotland ought to have a right to another referendum. Scotland's secession from the UK has the potential to cripple or perhaps dissolve the entity altogether. On the other hand, a second defeat in an independence referendum would likely put the debate surrounding independence to rest for a long period into the future.

A number of issues have dominated Scottish political discussion surrounding the possibility of leaving the UK. If Scotland wishes to become independent in hopes of rejoining the EU, it must consider the implications of separating from the UK. Issues to consider include economics, defense, energy, international relations, healthcare, nationalism, family and cultural ties, and the timing of a new independence referendum. In this paper, I analyze each of these issues in turn. Altogether, the optimal present solution is for Scotland to focus on maintaining unity with its UK neighbors until the Brexit negotiations take place and policy implications become clearer.

The Evolution of Scotland's Independence Movement

The original events that precipitated the current manifestation of Scottish nationalism and independence reform have roots that stretch back for centuries in British history (Broun 2007, p. 1). While these events will become more relevant in the discussion of nationalism later in the paper, this section's outline of relevant events will focus only on the more recent period leading up to the call by Nicola Sturgeon, the current first minister of Scotland and the leader of the pro-independence Scottish National Party (SNP), for a second Scottish independence referendum. This period began in 1997, when a successful devolution referendum won a significant victory for the Scottish independence movement. The referendum demonstrated wide support for the idea of increased autonomy for Scotland, and the British Parliament thereafter agreed to the 1998 Scotland Act, which created the Scottish Parliament (Flamini 2013, p. 61).

A push for independence grew in popularity in the following decade. The SNP, which promised an independence referendum as part of its electoral platform, became the largest party in the Scottish Parliament after the 2007 elections and won an overall majority in 2011. At that time, First Minister Alex Salmond negotiated with the British Parliament to hold an independence referendum, which London agreed to authorize. The referendum caused a sharp divide along a number of issues and ultimately amounted to the previously mentioned result of 55 percent to 45 percent in favor of remaining part of the UK.

The nature and dimensions of the Scottish push for independence have evolved over time. According to Mikesell and Murphy, the political aspirations of minority groups within a state can fall anywhere along a spectrum ranging from recognition, access, and participation to separation, autonomy, and independence (1991, p. 582). Within the context of this framework, one can observe how Scottish ambitions have developed. As full UK citizens, Scots have always enjoyed participation and access to British government and resources. Scotland's main aspirations during the 1990s fall under the categories of representation and autonomy, as it pushed for and successfully created a Scottish parliament that operated to direct some domestic affairs, albeit under the jurisdiction of the UK.

During the most recent years, however, Scottish sentiments have escalated to a large-scale push for separation and independence.

Mikesell and Murphy's model suggests that Scottish notions of self-determination have become extreme over time, but it remains to be seen if these sentiments will continue their upward trend or whether future Scots will look back on the recent calls for independence as a passing fad of the early twenty-first century. The trend of nationalism faces a defining moment in which it must decide whether its desire for autonomy will push it into a new era of independence or whether the price of such a move will ultimately make separation untenable.

The Issues Surrounding Scottish Independence

Without a doubt, the advent of Brexit has changed the political landscape for Scottish voters with regard to Scottish independence. A key issue during the debate leading up to the referendum surrounded whether Scotland would gain EU membership if it were to secede from the UK. The fact that the UK has voted to leave the EU means Scotland must leave, unless it can create a new option of gaining independence and then reentering the EU. As will be discussed in the upcoming sections, it remains entirely unclear whether or not such an option will ever become available. What is clear, however, is that Scotland no longer has the choice to retain membership in both the EU and the UK, as it did during the last referendum. This would create a new dilemma for Scottish voters in a hypothetical second referendum.

Despite this new development, most of the issues surrounding Scottish independence have not changed since the most recent referendum. Just as before, pivotal and divisive issues, such as defense, healthcare, citizenship, energy, culture, international status, and various economic concerns, will dominate discussion and ultimately determine Scottish popular opinion. This section will outline the details of these issues that sit at the center of the Scottish independence dilemma.

Defense

A clear issue involved in the creation of any new state is the management of its defense. Scots wish to remain a part of the EU but cannot do so as part of the UK due to Brexit. Scottish leaders must separate from the UK's defense system if it hopes to rejoin the EU. Scotland exists in a relatively safe neighborhood of states, but as a longstanding part of the UK—one of the region's most dominant military powers—it has to deal with carryover from its military involvement as part of Great Britain (Dorman 2014, p. 679). The primary defense issue that complicates a potential Scottish transition to independence is the location of the UK's Trident nuclear submarine bases in Scottish waters. Moving the bases would represent a large cost for the UK, but it seems unlikely that the UK would allow a separated Scotland to continue to operate its naval shipyards. Furthermore, the possibility of the creation of a new European state with nuclear capability, or the process of transferring or disarming nuclear weapons, invokes the participation of the international community in what would otherwise be a simpler independence transition.

The SNP's vision argues that an independent Scotland would have no need for nuclear weapons and advocates the removal of Trident. Scotland currently contributes tax revenue to the UK's nuclear program, and the SNP government estimates that Scotland would have to devote a further £100 billion in lifetime spending to updating its nuclear weapons were Scotland to remain a nuclear state (Scottish Government 2013, p. 232). The SNP argues that Scotland could make much better use of these funds were it to allocate them to other public needs, such as healthcare and education. On a practical and moral level, too, Scottish politicians have pointed to the example of a large number of European states—especially Nordic countries, which operate peacefully and effectively in the international community without any nuclear weapons (Bailes and Ingram 2013, p. 62).

One major complication with the SNP's vision remains its conflicting desires to join NATO while lowering its defense spending and military capability. The Government's White Paper report on Scotland's future outlines Scotland's membership in NATO as a clear priority of its future defense plan (The Scottish Government 2013, p. 234). The SNP highlights the savings an independent Scotland would accrue should it opt out of its joint support in funding the UK's defense spending. However, it overlooks the fact that NATO will continue to obligate Scotland to spend heavily on defense if it hopes to join (NATO 2016). Additionally, NATO would likely exert pressure to maintain the status quo of nuclear capability from the British Isles. Removal or disarmament of nuclear capability in Scotland would undermine NATO's priorities of maintaining a stable nuclear community, since the UK and France are Europe's only nuclear powers (Bailes and Ingram 2013, p. 63). Due to these pressures, Scotland cannot completely control its nuclear destiny.

Another consideration regarding the Trident nuclear program is the jobs it brings to Scotland's shipyards. While the government might save money in the long run by removing the nuclear facilities, it has to think about the short-term political and economic costs of removing jobs for thousands of Scottish workers (Good 2014). This reality may help explain why—despite perceived long-run benefits in national savings—support for Trident's continued operation has remained as strong as the support for Scottish withdrawal in opinion polls during the last year, with both camps hovering around 40 percent of respondents favoring the respective sides (ScotCen Social Research 2017). These polls show that SNP would have to deal with a significant amount of political backlash were it to move ahead with its plan to remove Trident. Overall, simply stating that an independent Scotland can and will quickly remove nuclear weapons overlooks both the feasibility of such a venture as well as the time frame involved, considering pressures from the UK, NATO, and Scottish workers.

Economy

Perhaps the most significant issues of Scottish independence, at least those that matter most in the minds of the Scottish voters, have been and will continue to be questions involving the future state of the Scottish economy. The SNP and other pro-independence

campaigners argue that UK headquarters in London has shortchanged Scotland on public spending, Scotland will benefit from increased economic autonomy, and it will maintain the British pound as its currency, ensuring stability and continuity. Upon closer analysis, however, economic uncertainty seems a more likely result than these supposed positive outcomes.

The border will likely prove one of the most obvious complications of Scottish separation from the UK. For businesses on both the north and south side of the dividing line, increased border security will raise the costs of transactions with businesses that serve customers in another country (Good 2014). The disruption the border will likely cause will simultaneously decrease Scottish support for independence in these regions while increasing the UK's political resolve not to allow a referendum. If Scotland wishes to avoid these issues it will need to negotiate an opt-out of the Schengen Area while applying for EU membership. This outcome seems possible, but since this agreement remains a fundamental value of the EU project, it still provides yet another complication on the pathway to membership (Banks 2014, p. 1).

One argument of the pro-independence camp has relied on the idea that the UK, centered in London and removed from Scottish interests, has inevitably led to unequal treatment in terms of public spending. In reality, Scotland actually benefits from spending arrangements that favor a smaller population spread over a larger geographic territory, as is the case of Scotland in comparison with its English neighbor. This has led to public spending in Scotland outdistancing England by £1,200 per capita annually (Martin 2014, p. 2). The supposed unequal spending by the UK has actually benefited Scots: If the North Sea oil and gas reserves turn out to be incapable of sufficiently supplementing government revenues, increasing public spending could become complicated.

A critical economic question for the future of an independent Scotland lies with the future of its currency. SNP leaders have often repeated the assurance that Scotland will continue using the British pound in the event of independence (Scottish Government 2013, p. 7). This continuity will bind interest and inflation rates to the rest of the UK and provide stability, while also ensuring continued ease of transaction between Scotland and the rest of Great Britain. However, this assumption of continued use of the British pound seems somewhat cavalier, given the reactions of leaders in Westminster and the potential restriction involved with joining the EU. Only a small handful of countries were allowed to join the EU without switching to the euro currency, suggesting that Scotland may at least run into difficulty negotiating this deal if it reapplies for membership (Flamini 2013, p. 62). British officials, such as Bank of England Governor Mark Carney, have used the example of Europe's troubled currency union to argue that a potential UK currency union with Scotland would be "incompatible with sovereignty" (Inman and Wintour 2014). Without extremely tight restrictions in place, forming a currency union with another country that makes its own financial decisions would put the UK at risk of Scotland undermining its economic stability.

Should the Scottish economy take a downturn upon gaining independence, the pound sterling, used throughout the rest of the UK, would suffer.

Other UK leaders have echoed Carney's concerns, making Scotland's future use of the pound very unlikely in the event of independence (Martin 2014). Without it, Scotland will undergo much greater risk of uncertainty and instability. Besides, polls show that only a small minority of Scottish voters would be willing to support the use of a currency besides the pound, while an even smaller number would be willing to use the euro, suggesting that the SNP will face a major uphill political battle in persuading Scot's to accept an alternative arrangement (ScotCen Social Research 2017). The SNP's oversight with regard to a proper plan for a stable currency was a major weak point leading up to the 2014 referendum and would assumedly receive a similar response in any forthcoming referendum.

The reaction of banks and financial institutions was another of the largest concerns that sank the possibilities of Scottish independence in 2014. During the last decade, Scotland has seen success in attracting large companies and banking institutions. In 2011, Scotland welcomed the entry of more foreign businesses than any other region in the UK, according to a report from Ernst and Young (Flamini 2013, p. 63). In the weeks leading up to the independence referendum, however, several large banks, including the Royal Bank of Scotland and Lloyds Banking Group, threatened to move their headquarters to London should the referendum succeed in producing an independent Scotland (Collinson, Treanor, and Jones 2014). The public statements of the banks' intentions to leave not only transmitted financial insiders' pessimistic opinions of Scottish independence to the public, but it also demonstrated the risk of a potential landslide of capital flight that could occur in a nascent Scottish state. Many wanted independence but fewer seemed willing to undergo the possibility of an economic recession to achieve it.

A case study comparison with Quebec during its 1995 independence referendum confirms that a violent and negative economic response is endemic to independence-related uncertainty. The fears of Scottish voters reflect the lessons learned by Quebec two decades earlier. Quebec, like Scotland, is a region within a larger country that enjoys semi-autonomous control over things like healthcare, education, and welfare. The referendum resulted in a very narrow (50.6 percent to 49.4 percent) defeat for proponents of independence. Despite its failure, it had reverberating economic consequences (Ibid., p. 1).

The build-up to the referendum in Quebec produced different kinds of uncertainty. First, people were unsure of the outcome. Second, people were unsure of what independence would mean should it obtain a victory in the referendum. This uncertainty produced a powerful economic response. Interest rates, exchange rates, and bond market ratings suffered for Canada as a whole, while Quebec-based businesses' performance decreased on the Toronto Stock Exchange. In fact, during the 1995 referendum campaign, the Toronto Stock Exchange suffered the sixth-largest drop in its history up until that point (Sayers 2014, p. 2).

The uncertainty created by Scotland's independence bid, as well as its potential success, could mirror the economic consequences of uncertainty that followed Quebec's independence referendum. These consequences will potentially cost Scotland in terms of financial instability and capital flight. Because of questions surrounding public spending, currency, and uncertainty, it seems clear that, at least economically, Scotland will remain in a better place if it maintains the status quo, rather than seeking for another independence referendum.

Energy

The critical components of the energy debate as it relates to Scottish independence fall into two categories: the production of renewable energy and the use of oil and natural gas. In both cases, pro-independence politicians have advocated separation from the rest of the UK as the best step forward in achieving maximum productivity and efficiency in these categories. Critics, however, have pointed out that independence may not provide the simple and immediate answers that the SNP and others have promised.

Scotland has emphasized its energy future as a key component of its independence bid, arguing that it has much more to gain than to lose in terms of energy production were it to become an independent country. The SNP government has stated that years of underinvestment in the field of renewable energy has harmed the UK's prospects going forward, and Scotland has an opportunity to become a regional leader in the industry. The long-term goal put forward by the SNP is to grow capacity in renewable energy production until Scotland can become an energy exporter, providing a boost for government revenue, jobs, and the economy as a whole, following the example of its North Sea neighbor Norway (The Scottish Government 2013, p. 296).

While the outcomes from renewable energy development described by the SNP government would likely produce excellent results for the Scottish economy, the SNP overlooks the possibility for increasing Scotland's energy productivity while remaining within the UK. The only real changes that need to take place in order for it to achieve its renewable energy goals are increased autonomy on energy policy and decreased regulations from London. Scotland could likely achieve both through negotiations with London without needing to obtain complete independence (Ibid., p. 295).

The second and more significant energy concern involved with Scottish independence is the future of the UK's oil and natural reserves in the North Sea. If Scotland separates from the rest of the UK, initial estimates suggest that upward of 19 percent of the oil and natural gas fields fall in offshore Scottish waters (Flamini 2013, p. 63). Although Scotland plans to rely heavily on its renewable energy in the future, these natural resources would provide an important source of energy security and, more importantly, revenue for the government, with which it could supplement tax revenue and make up the losses involved with leaving the much larger UK tax pool.

Scotland's rights to the North Sea's natural resources have provided one of the most convincing arguments for Scottish independence. Currently, the reserves belong to all of

the UK, meaning Scotland has to share resources that fall almost entirely within Scottish territory. Some estimate the offshore resources amount to as much as twenty-four billion barrels of oil, which could go a long way in subsidizing the government's expenditures in Scotland's transition to independence (Martin 2014, p. 3). It seems only fair for Scotland to receive a greater share of the resources that lie within its territory.

Some experts have cautioned that the estimates of pro-independence politicians may be over-optimistic. Sir Ian Wood of the Wood Group, which works in the North Sea oil industry, warned that he would place estimates of the reserves closer to fifteen or sixteen billion barrels, an estimate that has since been backed by British Petroleum and Royal Dutch Shell (Martin 2014, p. 3). While still a significant amount, the smaller estimate would force the SNP and others to work out far more austere measures for government spending without the additional income from oil and gas revenue. Others have gone on to point out that the North Sea oil fields have been harvested of their resources for many years, and the oil left, while still a large amount, will become increasingly difficult and more expensive to extract than it was previously (Good 2014).

Beyond the issues surrounding the SNP's overestimates of the North Sea oil and gas reserves, other critics have cited international law as another potential barrier to Scotland's unrestricted access to the resources. Estimations for Scotland's share of the oil and gas reserves are based on current administrative boundaries, but John Paterson, chair of law at the University of Aberdeen, has counseled the Scottish government to be wary of using estimates as the given reality of what will be the case after a Scottish separation. International law surrounding claims to the North Sea reserves, he says, will remain a matter of negotiation between countries, and the UK may well make a larger claim on the oil than it has previously. Because the conversation concerning the UK's natural resources will take place within the context of a much broader discussion concerning the division of debts and assets throughout the country, deciding who has the rights to the North Sea oil may not be as straightforward as some have imagined (Good 2014). Should Scotland attempt to retain complete possession of what have been the entire UK's oil reserves, the UK will argue that simply dividing resources based on geographic lines overlooks historical claims to the resources. British leaders can maintain that, just like two states would divide national debt equitably between them in the event of referendum, the UK and Scotland ought to divide its energy resources proportionally as well (Armstrong 2014, p. 299).

International Relations

A highly significant concern of a newly independent Scotland will be its relationship with both its neighbors and the international community as a whole. Scotland may have the highest stakes in play if it bids for independence, but its actions will interact with vital interests for the UK, the EU, and even the U.S., which would prefer not to see the strength of its close political and military ally undermined (Flamini 2013, p. 63). Scotland will have to decide the role it wishes to play, as well as plan strategically for the reactions other actors will have, should it separate from Great Britain.

Scotland will have to tread carefully to ensure that it does not permanently damage its relationship with England and the rest of the UK, since it will have to maintain important economic and cultural ties with its neighbors going forward. As previously mentioned, Prime Minister Theresa May has repeated her opinion that Scotland needs to focus, along with the rest of the UK, on unity as they negotiate with the EU, and any referendum on independence should occur after the negotiations with the EU take place. Sturgeon and other leaders will have to decide how far they can push London officials on the Scottish independence issue without alienating them, since a UK without Scotland will still maintain a stronger economic and political bargaining position compared to its new northern neighbor (Bailes and Ingram 2013, p. 42).

One of the biggest issues to decide with regard to the rest of the UK will be that of citizenship, as a large number of those born in Scotland now live elsewhere in the UK, and many from the UK now live in Scotland. Families extend across potential future borders. Scotland will face a tricky situation in deciding how to handle the citizenship of these people. It will likely end up having to let people decide for themselves but many will not be satisfied choosing to be only Scottish or only British (Flamini 2013, p. 62). The two countries could attempt to reach some kind of compromise that would allow citizens of each country to retain dual citizenship, but it seems unlikely that London will feel overly generous in doling out the benefits of citizenship to Scottish nationals very soon after they declare independence.

Scots will also have to decide their potential political relationship with the British royal family and whether they will follow the path of countries like Canada and Australia, who still maintain the British monarchy as the head of state (Martin 2014, p. 4). While the role of the monarchy would remain symbolic, the issue has divided Scots. Some wish to continue the traditional tie to Great Britain, while others think the British monarchy undermines Scottish nationalism (ScotCen Social Research 2017).

Arguably, the most pivotal outside actor in relation to Scotland's potential independence bid is the EU. The UK's vote to leave the EU remains the primary catalyst in the SNP's renewed call for independence so close on the heels of the 2014 referendum. The SNP has suggested that entry into the EU will prove a relatively simple process. Scotland will simply tweak a few treaties and reenter in the place that the UK will vacate. Some European officials, such as Guy Verhofstadt, the European Parliament's chief Brexit negotiator, have expressed sympathy for Scotland's cause, saying that it should not have to leave the EU when it voted strongly in favor of staying (Carrell 2017). Most, however, have maintained that a clear standard for application to the EU is set out under Article 49 of the European Treaties, and Scotland would have to join the queue of other applicant states, such as Bosnia, Serbia, and Montenegro. While Scotland would likely move through the process of application more quickly than some of these other states due to its ability to meet EU member criteria more easily, it would most likely still have to start in the same place and complete the same steps as any other applicant. Scotland would also have to apply as a new member to join the UN (Flamini 2013, p. 60).

While it seems unlikely that Scotland would achieve a simple and immediate entry into the EU as SNP leaders have suggested, the bigger problem it faces remains the question of whether the EU member states will be willing to let it enter at all. The EU requires unanimous approval of all member states in order for a new state to join. Both Spain, with Catalonia, and Belgium, with Flanders, combated secessionist movements within their territories and would rather not allow a precedent for successful secession to take place within the EU, which might encourage their respective separatist populations. These states could easily use their veto power to prevent Scotland from achieving EU membership (Flamini 2013, p. 60). After Sturgeon and May's meeting in late March 2017 to discuss a second Scottish referendum, Spanish Foreign Minister Alfonso Dastis made a public statement that "Spain supports the territorial integrity of the UK and doesn't encourage secessions or divisions in any of the member states" (Torres 2017). The Spanish government in Madrid moved quickly and decisively to repress the recent attempt of Catalonian leaders to declare independence by dissolving the regional government headquartered in Barcelona. Were Scotland to rejoin the EU, it would encourage Catalonia's hopes of doing the same. While Spain has yet to make any outright statements affirming it will veto a Scottish entry into the EU, its behavior toward Catalonia signals that it will not take lightly a move for Scottish independence.

In summary, Scotland faces, at best, a large degree of uncertainty regarding if or when it would join the EU. In the best-case scenario, it may well achieve membership, but it could only come after a long process of negotiations and will likely hinge on the status of secessionist movements elsewhere in Europe. If reentry into the EU remains one of the primary objectives of a second Scottish independence referendum, Scotland can only hope to find a way of placating the secessionist concerns of Spain and Belgium. This negotiation process will remain difficult as long as secessionist movements remain a concern for these countries, which seems unlikely to change in the near future. Scotland faces a long battle to become a member of the EU, if membership is indeed attainable at all.

Healthcare

A primary debate regarding Scottish independence in 2014 involved Scotland's National Healthcare System (NHS), and this debate will remain just as relevant in light of a new independence bid post-Brexit. Scotland gained independent control of its healthcare system and policy in 1999, during the same devolution agreements that saw the creation of a Scottish Parliament. The UK, however, still exerts some command over the system because it maintains control over the budget and spending on healthcare in all of Great Britain, including Scotland (Mathieson 2014). The healthcare debate, consequently, concentrates mainly on finances.

The argument of those supporting independence suggests that Scotland and the rest of the UK have begun to pursue increasingly divergent paths when it comes to healthcare. The British Parliament's policies have begun to focus more on privatization and competition within the industry, while the Scottish Parliament remains a strong propo-

nent of public healthcare. Leaders of the SNP and other pro-independence-minded politicians have warned that Scotland could face budget cuts for its healthcare in the future due to financial decisions and austerity measures made exclusively in London. They suggest that Scotland ought to have the ability to dictate its future with regard to both its healthcare policy and its budget (The Scottish Government 2013, p. 172). Without this power, Scotland's healthcare priorities will suffer from underfunding, and the country will have to adopt policy measures against its values.

This line of reasoning, however, remains subject to many critiques. To begin with, Scotland enjoys significantly higher spending per capita on healthcare by the UK government than England. According to a report during 2012–13, English residents were the recipients of £1,912 in annual government spending on health care on average, compared to £2,115 per person in Scotland, making it difficult to argue that Scotland is getting as bad of a deal as some politicians would like it to appear (Mathieson 2014, p. 3). More significantly, an independent Scotland would plan to obtain most of its tax revenue from North Sea oil (see section on Energy and resources). Between the depletion of these natural resources going forward, as well as a small population and an aging demographic, Scotland's ability to support its own healthcare system in the long run seems far less viable than its possibilities were it to stay connected to the tax revenue of all of Great Britain (Ibid., p. 4).

This argument not only undermines the SNP's logic for healthcare gains from independence but also provides a strong argument for the benefits of remaining a part of a larger, wealthier country. The climate among those involved in the healthcare industry prior to the 2014 referendum was one of pessimism and skepticism. A study surveyed 311 doctors, finding that 60 percent planned to vote against independence, with only 33 percent supporting independence (Rimmer 2014, p. 1). The doctors cited concerns over an economic downturn that would prevent an independent Scotland from recruiting enough medical professionals, while encouraging many to move south of the border to practice.

Others have raised concerns over waiting lists for operations or organ transplants. If Scotland must rely solely on its own healthcare system and a smaller population to sustain it, the logistics involved could increase wait times for life-saving surgeries and organ donations. Since EU health cards ensure access to healthcare in any EU member country, some pro-independence politicians have concluded worries about access to such emergency operations would not be a real concern were Scotland to rejoin the EU and regain access to the medical resources of other EU members. They argue that Scottish citizens have no need to rely on British healthcare (Good 2014). This argument provides little certainty for a country that may not rejoin the EU; the loss of access to the medical resources of its closest neighbor would certainly cause logistical problems for Scotland when it comes to issues like organ donations and emergency transplant operations. The country will take on more risk in this area if it chooses independence.

Nationalism

Scotland will evidently undergo heavy costs if it wishes to obtain independence. History, however, shows that this may not be entirely surprising. The bloody American Revolution, among others, demonstrates that nations seeking for independence sometimes value freedom even at high costs. Even though economics and political liberties matter to voters, it is possible that people might still view independence as more valuable than the sacrifices it will entail because of nationalism (Van Evera 1994, p. 5). If the feeling of nationalism remains strong enough within Scotland, its voters may continue to support independence, even at a high cost.

The roots of Scottish nationalism in relation to its southern neighbors stretch back centuries into British history. In A.D. 128, the Roman emperor Hadrian built a wall to keep out northern marauders, while the Scottish Wars of Independence in the thirteenth and fourteenth centuries gave rise to legends in Scottish national lore, like William Wallace (Broun 2007, p. 24). Unlike these wars in past centuries, the modern manifestations of Scottish nationalism have remained peaceful, but Scots clearly still feel nationalist sentiments. Still, Scotland would theoretically have to experience a widespread notion of nationalist Scottish identity in order to be willing, together as a nation, to sacrifice in order to obtain statehood (Nietschmann 1994, p. 241).

The trouble with Scotland, in terms of generating enough sentiment to become independent, remains that it faces division in the degree to which people feel strongly about Scottish identity throughout the country. According to census data, over a third of the country claimed to identify primarily as something other than "Scottish only," indicating the complexity of identity politics in Scotland (Scotland's Census 2011). As discussed previously, people who live close to the border between England and Scotland tend to see themselves as a kind of in-between people who identify with both nationalities, since the border has not been a true cultural barrier for a very long time. Furthermore, people in the border regions will be among those most hurt by Scottish independence and, therefore, will likely oppose nationalist sentiments to a greater degree. People who live far to the north of Edinburgh, in the Shetland Isles, also experience an identity crisis; when interviewed, some felt a greater connection to Norway than Britain, and most claimed to identify first as Shetlanders and second as British (Good 2014). Overall, it currently appears that Scottish nationalism remains too fractured to produce the cohesion necessary for achieving widespread support for independence.

Although Scotland remains characterized by a diversity of sentiments regarding independence at this time, proponents of Scottish independence ought to consider the role of age in identity politics. Data from 1999 to 2010 show that young people are more likely than older generations to support Scottish independence. Over this span, those between eighteen and twenty-four were 10 to 15 percent more likely to support independence as the best form of ruling Scotland (Schneider 2014, p. 56). While one cannot be certain about the causation of the support for independence among Scottish youth, this finding could be evidence of a generational shift that has begun to take place in Scotland

and provides a strong argument in favor of the SNP waiting a few years to carry out a second Scottish referendum, rather than doing so immediately.

Cultural and Familial Ties

Scotland has operated together with its southern neighbors as a single country since the Treaty of Union in 1707, which created the country of Great Britain. Their coexistence has led to the crossover one might expect to happen between neighboring regions of a single country during the past three hundred years, and the delineation between what it means to be Scottish and what it means to be English has become blurred over time.

At the time of the 2011 Scottish census, approximately 500,000 people, over 10 percent of Scotland's population, were born in England, Wales, or Northern Ireland, while approximately 700,000 Scots resided in other parts of the UK (Scotland's Census 2011). This overlap on a massive scale requires pro-independence politicians to account for the difficulty of deciding how to handle immigration and border rights with the rest of the UK were it to seek independence.

The issue of the border will also prove to be a difficult problem for politicians to resolve, primarily because it has not operated as a real border since the distant past. Most people who live in the border region have developed familial and business relationships on both sides of it, and many view themselves as neither English nor Scottish but as simply British (Good 2014). In fact, according to census data, only 62 percent of Scots viewed themselves as "Scottish only," with 18 percent of the population identifying with both "Scottish and British identities," and 8 percent choosing "British only" as their national identity (Scotland's Census 2011).

Furthermore, border-dwellers will have a tremendously difficult time adjusting to the realities of a real border, were Scotland to put it in place, as their livelihoods often require them or their clients to cross borders. Were Scotland to rejoin the EU and the Schengen Agreement, it seems impossible to imagine that Great Britain, which left the EU in large part due to populist opinion against free migration of immigrants into their country from elsewhere in Europe, would leave an open border with Scotland across which immigrants could travel (Good 2014).

The Timing of a Second Referendum

One of the most basic questions involved in proposing a second Scottish referendum lies in its timing. Nicola Sturgeon has announced her intentions to seek a second Scottish independence referendum, which she hopes will take place between the fall of 2018 and the spring of 2019 (Stone 2017). While the Scottish Parliament in Holyrood will back Sturgeon's appeal, British Prime Minister Theresa May and the UK Parliament in London have the final say in authorizing the referendum. May has repeatedly stated her objection to a second Scottish referendum. She insists that the process of Brexit negotiations call for national unity and that a Scottish referendum would undermine the negotiations' outcomes, creating a less favorable deal for Scotland and the rest of the UK (Taylor 2017).

Inherent to the timing dilemma is an element of risk for the pro-independence advocates. Should Sturgeon somehow succeed in bringing about a second referendum, the results must be final in one direction or another. If she fails in leading Scotland to independence, it becomes difficult to imagine she or any other politician will remain capable of mustering enough political will to push a third independence bid for years to come. This raises the stakes of holding a referendum and prioritizes the need to launch it at the right time.

The following line graph represents the data obtained by asking the question: "When do you think another Scottish independence referendum should be held?" Researchers compiled the data collected during the last nine months.

Table 1
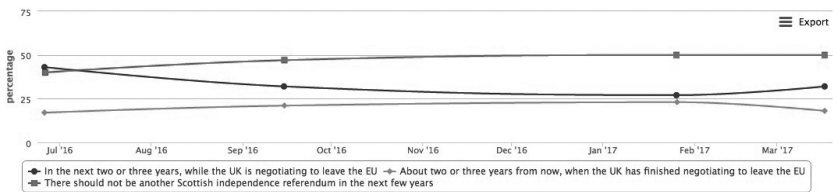
Although a brief moment existed immediately following the Brexit referendum in which Scots narrowly supported holding another referendum before the Brexit negotiations finish, the most dominant group in the polling clearly support not having a referendum in the next several years. Since February 2017, there appears to be some convergence, with those supporting a referendum after the negotiations declining, and those supporting a referendum during negotiations on the rise. Whether this trend will continue, and whether support for no referendum at all will decline, remains to be seen. Still, it seems clear that most Scots oppose Sturgeon's idea of an immediate referendum.

The data may confirm that Sturgeon's calculus is not entirely off. The strongest support for a new independence referendum has centered on the time of the Brexit referendum and the build-up to May's triggering of Article 50 to begin Brexit negotiations, suggesting that perhaps proponents of independence will have the best chance at earning popular support while Brexit remains fresh on voters' minds. However, even if this proves to be the case, the opportune moment of independence support may have already passed.

Conclusions and Policy Suggestions

The issues that have dominated discussion surrounding Scottish independence will ultimately determine when or if a second Scottish referendum will take place. Even though Scotland now faces the cost of leaving the EU in the wake of Brexit, the cost of leaving the UK is still higher. In the aggregate, the best solution for Scotland remains to wait patiently, and as Prime Minister Theresa May suggested, focus on unity with

the rest of the UK during the Brexit negotiations. While Scotland may well continue to feel that Brexit has changed the nature of its relationship with Great Britain, after the negotiations take place, Scotland will have a better idea of the situation in which it will exist going forward. This will enable voters to make decisions with more certainty than currently possible.

The takeaways from the discussion of most of the issues discussed in this paper—familial and cultural ties, healthcare, defense, energy, and economics—indicate that separation will prove difficult at any point, whether now or in the future. Those living near the border will be worse off in any situation that implements enforced separation between the states. Scotland can achieve its proposed healthcare and renewable energy reforms through negotiations to decrease London's central authority and increase Scottish autonomy, without needing to go as far as full independence. North Sea oil will remain a point of friction no matter what path Scotland chooses to take and will likely not provide a universal solution to government revenues should Scotland obtain independence. While it would be difficult to imagine London removing its nuclear submarine bases from Scottish waters without Scotland leaving the UK, the Trident nuclear problem will not gain a simple solution through independence either. Finally, all signs suggest that Scotland stands to lose in economic terms if it chooses independence.

Although perhaps less pessimistic in their outcomes than the other issues, the conclusions concerning nationalism, the timing of the referendum, and international relations suggest that the SNP would be wise to wait rather than immediately hurrying into a second independence referendum. Sturgeon does not currently hold the political support necessary, but age demographics and a spike in interest with Brexit events demonstrate that the most likely time for a referendum to find success would come in a few years. If the realities of Brexit's effects on Scotland are as unfortunate as predicted, then support for independence will only increase. Scotland's bid to join the EU will necessarily depend on the status of Belgium's and Spain's secessionist movements, while negotiations involving nuclear weapons and membership in NATO and the UN will also require a longer timeline. In general, Scotland's independence bid will likely benefit in these areas if it waits. If Sturgeon gets a referendum approved now, and then it fails, the SNP may take years or decades to build enough political capital to attempt another independence bid.

As a common trend throughout most of these issues, uncertainty remains the biggest issue in the current state of Scottish independence reform. Voters simply do not have the information necessary to make a comfortable decision regarding the future of their country. This information will only become available once the UK and the EU complete Brexit negotiations and the new status quo becomes known.

Scotland's best option, as things stand, will be to begin by negotiating with London for increased autonomy in its healthcare and energy objectives, especially in terms of budget and financing. Officials in London may be willing to give concessions to Scotland if it agrees to wait on an independence referendum until after the Brexit

negotiations. The time that passes will prove beneficial for Scotland, since more information will become available to voters as Brexit takes place, and Scots will grow more capable of making the decisions necessary to determine the future status of their country. Scottish patience will help the UK maintain stability as it negotiates Brexit, and this in turn will help Scotland maintain a better working relationship with the UK, allowing the SNP to move into a better position to negotiate and attempt a successful referendum.

REFERENCES

Adamson, Kevin and Peter Lynch, eds. 2014. *Scottish Political Parties and the 2014 Independence Referendum*. Cardiff, Wales: Welsh Academic Press.

Armstrong, Angus. 2014. "Assets and Liabilities and Scottish Independence." *Oxford Review of Economic Policy* 30, no. 2 (Summer): 297–309.

Bailes, Alyson and Paul Ingram. 2013. "Scotland the Brave New World." *The World Today* 69, no. 3 (June/July): 42–43.

Banks, Martin. 2014. "Scotland Risks Losing Opt-Outs on Euro and Schengen if It Splits From the UK, Warns EU Expert." *Guardian*, September 16. Accessed April 14, 2017. www.telegraph.co.uk/news/worldnews/europe/eu/ 11098835/Scotland-risks-losing-opt-outs-on-Euro-and-Schengen-if-it-splits-from-UK-warns-EU-expert.html.

Broun, Dauvit. 2007. *Scottish Independence and the Idea of Great Britain*. Edinburgh: Edinburgh University Press.

Carrell, Severin. 2017. "Independent Scotland 'Would Have to Apply for EU Membership.'" *Guardian*, February 10. Accessed March 27, 2017. www.theguardian.com/politics/2017/feb/10/independent-scotland-would-have-to-apply-for-eu-membership.

Collinson, Patrick, Jill Treanor, and Rupert Jones. 2014. "Banks Say: We'll Leave Scotland if the Independence Vote is Yes." *Guardian*, September 10. Accessed March 25, 2017. www.theguardian.com/uk-news/2014/sep/10/scottish-independence-create-mortgage-drought.

Curtice, John. 2006. "A Stronger or Weaker Union? Public Reactions to Asymmetric Devolution in the United Kingdom" *Publius* 36, no. 1 (Winter): 95–113.

Curtis, Chris. 2017. "Why Hasn't Scotland Changed Its Mind on Independence?" *Guardian*, January 27. Accessed February 11, 2017. www.theguardian .com/ commentisfree/2017/jan/27/shift-scottish-independence-yougov-nicola-sturgeon-balancing-act.

De Freytas-Tamura, Kimiko. 2016. "Scotland Says New Vote on Independence is Highly Likely." *New York Times*, June 24. Accessed February 11, 2017. www.nytimes.com/2016/06/25/world/europe/brexit-scotland-independence-referendum.html.

Dorman, Andrew. 2014. "More Than a Storm in a Teacup: The Defense and Security Implications of Scottish Independence." *International Affairs* 90, no. 3 (May): 679–96.

The Economist. 2016. "Scottish Independence: RIP UK?" *Economist*, September 11. Accessed February 11, 2017. www.economist.com/news/leaders/ 21616957-ditching-union-would-be-mistake-scotland-and-tragedy-country-it-leaves.

Fieldhouse, Edward A. and Christopher Prosser. 2016. "When Attitudes and Behavior Collide: How the Scottish Referendum Cost Labour." Paper presented at the Midwest Political Science Association Conference, Chicago, Illinois, April 7–10.

Flamini, Roland. 2013. "Scotland's Independence Bid: History, Prospects, and Challenges." *World Affairs* 176, no. 1 (May/June).

Geoghengan, Peter. 2016. "Scotland Eyes British Exit." *Politico*, October 12.

Good, Alistair. 2014. "Scottish Independence: Issues Concerning Voters." *Telegraph*, August 18. Accessed March 20, 2017. www.telegraph.co.uk/news/ uknews/scottish-independence/10616453/Scottish-Independence-the-issues-concerning-voters.html.

Gordon, Tom. 2017. "No Second Chance for Pupils if Scotland Focuses on Second Independence Vote." *The Herald*, February 12. Accessed February 11, 2017. www.heraldscotland.com/politics/15087535.No_second_chance_for_pupils_if_SNP_focus_on_second_independence_vote__warns_Davidson/.

Hazell, Robert. 2013. "Would Holyrood Pay for Independence in Euros?" *The World Today* 68, no. 7 (January): 40.

Howarth, Angus. 2017. "Only One Fifth of Scots Would Back New Currency After Independence." *Scotsman*, February 12. Accessed February 11, 2017. www.scotsman.com/news/only-one-fifth-of-scots-would-back-new-currency-after-independence-1-4364577.

Hunt, Alex. 2017. "All You Need to Know about the UK leaving the EU." *BBC*, March 30. Accessed February 11, 2017. www.bbc.com/news.

Inman, Philip and Patrick Wintour. 2014. "Scottish Independence: Currency Union 'Incompatible With Sovereignty.'" *Guardian*, September 9. Accessed March 28, 2017. www.theguardian.com/business/2014/sep/09/currency-union-independent-scotland-unworkable-bank-governor.

NATO. 2016. "Defense Expenditure of NATO Countries (2009–2016)." Information on Defense Expenditures, July 4. Accessed March 25, 2017. www.nato.int /cps/en/natohq/news_132934.htm.

Nietschmann, Bernard. 1994. "The Fourth World: Nations Versus States." In *Reordering the World: Geopolitical Perspectives on the 21st Century*, edited by George Demko and William Wood. Boulder, CO: Westview Press.

Mathieson, S.A. 2014. "How Scotland's NHS Would be Affected by a Yes Vote." *Guardian*, September 16. Accessed March 17, 2017. www.theguardian .com/healthcare-network/2014/sep/16/nhs-scotland-independence-yes-vote-policy.

Mikesell, Marvin W. and Alexander B. Murphy. 1991. "A Framework for Comparative Study of Minority Group Aspirations." *Annals of the Association of American Geographers* 81, no. 4 (Summer): 581–604.

Murkens, Jo Eric Khushal, Peter Jones, and Michael Keating. 2002. *Scottish Independence: A Practical Guide*. Edinburgh: Oxford University Press.

The Office for National Statistics. 2017. "Dashboard: Understanding the UK Economy." *National Statistical Institute*, March 11. Accessed February 11, 2017. www.ons.gov.uk.

Rimmer, Abi. 2014. "Doctors in Scotland Fear that Independence Would Affect Recruitment." *British Medical Journal Careers*, August 13. Accessed March 17, 2017. http://careers.bmj.com/careers/advice/view-article.html?id=20018822.

Ross, Tim. 2016. "Scotland Threatens to Leave UK if Forced Out of Single Market." *Bloomberg*, December 19. Accessed March 2, 2018. www.bloomberg.com/news/articles/2016-12-19/Scotland-threatens-to-leave-u-k-if-forced-out-of-single-market.

Sayers, Louise. 2014. "Scottish Independence: The Debate From a Czech and Canadian Perspective." *BBC*, September 17. Accessed April 17, 2017. www.bbc.com /news/uk-scotland-29126592.

Schneider, Anna. 2014. "Age Variations in the Attitude toward Scottish Independence." *Scottish Affairs* 23, no. 1 (Summer): 55–78.

ScotCen Social Research. 2017. "When Do You Think Another Scottish Independence Referendum Should be Held?" *What Scotland Thinks*, March 17. Accessed March 13, 2017. http://whatscotlandthinks.org/questions/when-do-you-think-anotherscottish-independencereferendum-should-be-held-asked-a#line.

Scotland's Census. 2011. "Ethnicity, Identity, Language, and Religion." Last modified April 9. Accessed March 15, 2017. www.scotlandscensus.gov.uk/ethnicity-identity-language-and-religion.

Scottish Government. 2013. *Scotland's Future*. Edinburgh: APS Group.

Stone, Jon. 2017. "Nicola Sturgeon Announces Second Scottish Referendum." *Independent*, March 13. Accessed March 13, 2017. www.independent.co.uk/news/uk/politics/second-

scottish-independence-live-referendum-nicola-sturgeon-brexit-speech-second-indy-ref-2-uk-eu-a7626746.html.

Taylor, Brian. 2017. "Theresa May and Nicola Sturgeon Meet Ahead of Article 50." *BBC*, March 27. Accessed March 27, 2017. www.bbc.com/news/uk-scotland-scotland-politics-39399390.

Torres, Diego and Tom McTague. 2017. "Spain to Scotland: You're Not Special." *Politico EU*, March 27. Accessed March 28, 2017. www.politico.eu/article/spain-to-scotland-out-is-out-catalonia-independence-referendum/.

Van Evera, Stephen. 1994. "Hypotheses on Nationalism and War." *International Security* 18, no. 4 (Spring): 5–39.

Woo, Stu. 2016. "Scottish Leaders Press for Second Vote of Independence From Britain." *Wall Street Journal*, June 26. Accessed February 11, 2017. www.wsj.com/articles/scottish-leaders-press-for-second-vote-on-independence-from-britain-1466967044.

KENNEDY CENTER INTERNATIONAL STUDY PROGRAMS   BYUPAS BRIGHAM YOUNG UNIVERSITY POLITICAL AFFAIRS SOCIETY   ΠΣΑ Pi Sigma Alpha Beta Mu Chapter   BYU POLITIKOS POLITICAL AND INTERNATIONAL STUDIES PUBLICATIONS