Substitution or Revolution?
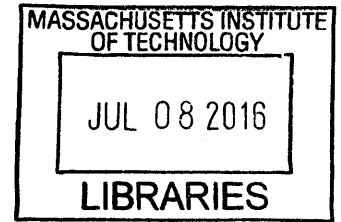An empirical analysis of the relationship between ICT and protests

by

Joseph P. Schuman

Submitted to the
Department of Mechanical Engineering
in Partial Fulfillment of the Requirements for the Degree of

Bachelor of Science in Mechanical Engineering

at the

Massachusetts Institute of Technology

February 2016

Signature of Author: _____ Signature redacted _____
Department of Mechanical Engineering
January 15, 2016

Certified by: _____ Signature redacted _____
Roger Petersen
Arthur and Ruth Sloan Professor of Political Science
Thesis Supervisor

Accepted by: _____ Signature redacted _____
Anette Hosoi
Professor of Mechanical Engineering
Undergraduate Officer

Substitution or Revolution?
An empirical analysis of the relationship between ICT and protests

by

Joseph P. Schuman

Submitted to the Department of Mechanical Engineering
on January 15, 2016 in Partial Fulfillment of the
Requirements for the Degree of

Bachelor of Science in Mechanical Engineering

ABSTRACT

Have cell phones and the Internet created a revolution in the way that protests are organized? Or are they merely a substitution for previous methods of communication? The literature on the topic is divided between the cyberphiles, who argue that information and communication technologies (ICTs) allow individuals to better organize and amplify social movements and the cyberskeptics, who reject the idea that ICT represents a different mechanism of communication and instead argue that protests stem from underlying structural issues. I analyze protests from the Arab Spring and the dissolution of the Soviet Union in order to answer the question of substitution versus revolution.

Through my empirical analysis, I find that cell phone subscriptions correlate positively with protest size, meaning countries with more cell phone subscriptions experienced larger protests, *ceteris paribus*. I did not, however, find the same result with Internet usage. My findings support the "amplification model" – that ICT amplifies existing social movements – but with some added nuance. I argue that the Internet amplifies information about protests to the wrong individuals – namely Internet users abroad who cannot participate in protests – while cell phones are used for domestic communication and thus increase the size of protests.

Thesis Supervisor: Dr. Roger Petersen
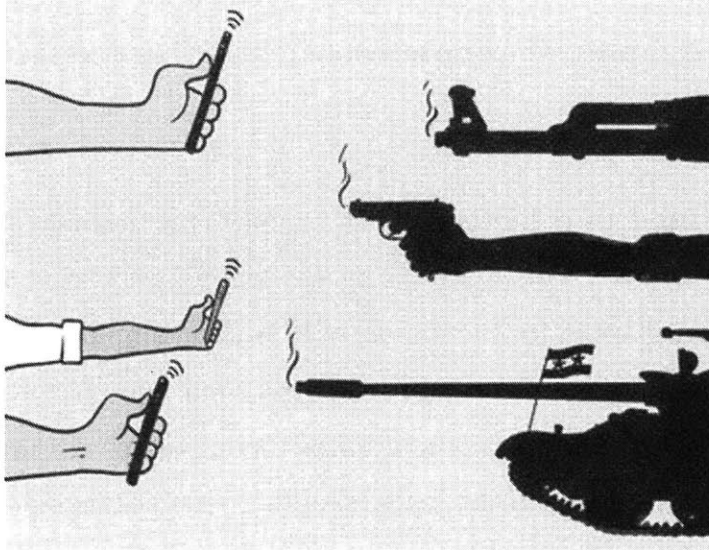Tile: Arthur and Ruth Sloan Professor of Political Science

Table of Contents

List of Tables

List of Figures

Introduction

The terms "Facebook Revolution" and "Twitter Revolution" have become synonymous with the Egyptian and Tunisian revolutions. It is common to attribute the magnitude and impact of the Arab Spring protests to information and communication technologies (ICTs) such as the Internet and cell phones. Yet, the ubiquity of these technologies, "does not automatically imply that [their] use will have any particular type or magnitude of aggregate consequences" (Agre 2001). Large-scale protests have existed throughout human history. Thus, the question remains, have ICTs created a revolution in the way that protests are organized or are they merely a substitution for previous methods of communication? The literature is divided on the issue. The "cyberphiles" argue that ICTs help to organize and amplify social movements while the "cyberskeptics" argue that underlying political and economic issues cause large-scale movements.

Aday (2012) notes that there are "major methodological challenges inherent in determining the actual causal impact of new media." Many of the existing sources only investigate cases within the Arab Spring and find correlations between ICT usage and protest size or state fragility. But, these types of analyses cannot differentiate between ICT existing as a new, improved tool of communication and ICT replacing older tools.

For this reason, my large-N sample includes twenty-one country-cases from the dissolution of the USSR in addition to the twenty Arab Spring cases. With these forty-one

cases, I am able to perform robust statistical analysis and draw distinctions between the 1989 cases and the 2011 cases, which no previous analysis has thus far been able to do.

Literature Review

There exists a significant body of literature that examines the mechanisms by which ICT facilitates large-scale protests. The literature agrees that ICTs act as a tool for organization and there exists some argument for the impact of ICTs as an amplifier. In contrast, there exists argument that maintains that certain structural conditions are more conducive to large-scale uprisings, even if they do not guarantee them. The two groups – sometimes referred to as the "cyberphiles" and the "cyberskeptics" – do not regularly acknowledge or investigate the work of the other and as such, the literature is largely indeterminate.

My literature review is divided into four sections. The first two highlight the principle arguments regarding the impact of ICT on protests: organization and amplification. The third section explains the arguments in the structural forces literature. Lastly, my fourth section attempts to reconcile the two bodies of literature using two additional sources.

ICT as an Organizer

Much of the literature argues that technology acts as an organizer of social mobilization. ICT's improvements on former means of social organization can be broken down into three categories: increased accessibility, decentralization, and acceleration. In terms of accessibility, Chebib (2011) notes that social media has incredibly low barriers to entry because accounts are free and granted to everyone. Social media is accessible through multiple means – at home and on a mobile device – and is easy to use as a result of easily understandable interfaces. Further, the accessibility of ICT gives eyewitnesses the ability to disperse highly credible information quickly and easily. Credible news is particularly relevant in states with repressive regimes, such as those analyzed in this paper, where a free media might not otherwise exist (Chebib 2011).

The second advantage of ICT as an organizer of social movements is its decentralization. El Tantawy (2011) notes that ICT facilitates communication that would otherwise have been restricted by financial, temporal or spatial constraints. Further,

decentralization enables "weak ties" that could not otherwise be formed between individuals (Chebib 2011). As a result, social movements waged over the Internet can be more protracted because social networks can sustain them in the absence of central organizations (Bennett 2003).

The third organizational advantage of ICT discussed in the literature is its ability to accelerate communication. Chebib (2011) argues that the "acceleration effect" of ICT has made large-scale social mobilization possible in a few weeks when it previously would have required a year. El Tantawy (2011) comments on the "swiftness" that ICT provides in organizational tasks such as planning, as well as receiving and disseminating information. Chebib (2011) argues that the accelerating effect of ICT enabled the overthrow of Mubarak in Egypt and Ben Ali in Tunisia in 18 and 28 days, respectively.

Empirically, there exists ample evidence of ICT as an organizer. Tufecki's (2012) survey of Tahrir Square participants shows that 46% of attendees used text messages to communicate about the protests, while 51% used Facebook, and 82% used mobile phones. Yet, there also exists a significant amount of nuance to these numbers. Tufecki also finds the continued importance of older forms of communication: 58% of attendees used print to communicate about protest and 92% heard about the protests through satellite TV. Even more staggering, 48% of demonstrators first heard about the Tahrir Square demonstrations through face-to-face communication. This plurality is significant when compared to the 28% of individuals hearing about the protests first through Facebook, the 13% hearing via telephone, and the 5% hearing though all other means including texting, e-mail, and Twitter (Tufecki 2012).

ICT as an Amplifier

There also exist proponents of the "amplification model," who argue that the Internet creates little that is qualitatively new in terms of organization, but rather amplifies existing forces (Agre 2001). Hussain (2013) supports the amplification model, finding that the Internet, "created regional and international news events that drew attention and sympathy from neighboring countries, and inspired others to join and celebrate the causes."

The primary question that must be answered by proponents of the amplification model is to whom does ICT amplify information. Using data on the number of clicks on Twitter links, Aday (2012) finds that social media outlets are more likely to spread information outside the region than inside it. Aday shows that over 75% of clicks on popular Twitter hashtags were from viewers outside of the region, while only about 10% of clicks were within the country where the uprising was taking place.

Some argue that the international dimension of the amplification model is a critical function of ICT (Stepanova 2011). Yet, as Aday (2012) instructively points out, the international amplification of communication only is impactful if it leads to a "boomerang effect," increasing pressure on autocratic regimes or helping reduce a regime's tendency to crack down violently on protests. Barring the boomerang effect, Aday does not believe that international amplification matters in terms of social mobilization.

Structural Issues

The literature focuses on three structural issues related to large-scale social movements: unemployment, education and youth demographics. As far as unemployment is concerned, the underlying theoretical connection with protest participation is opportunity costs. Campante (2012) argues that if human capital is more valuable in production, then individuals will be less likely to direct their human capital towards political participation because of the opportunity costs. Hoffman (2012) finds that the Middle East and North Africa have the highest regional rates of joblessness in the world. Similarly, Gelvin (2012) finds that the lack of employment opportunities in the Arab world has given rise to the phenomenon of "waithood," a period in which young adults remain indefinitely unemployed. Due to the low opportunity costs of participating in protests while jobless, unemployment may be a partial explanation for the large-scale uprisings seen during 1989 and 2011.

Next, education is commonly cited as a requisite for political participation. Campante (2012) notes that individuals with higher educational attainment, "consistently exhibit a greater propensity to participate in the full spectrum of political activities, from milder forms of engagement such as voting or discussing politics to more public forms of mobilization such as demonstrations." In fact, the relationship between education and

increased "civic skills" has been found to be causal in various randomized and quasi-experimental settings (Campante 2012). In the context of the Arab Spring, Özekin (2014) finds the expected years of schooling in the Arab states rose from 8 to 11.4 years from 1980 to 2012. This increase of education in the region might be another potential explanation of the mobilization seen during the Arab Spring.

Thirdly, the literature finds that increases in the youth population are especially conducive to social unrest. In what Gevlin (2012) calls the "youth bulge," approximately 60 percent of the population in the Arab world is under the age of thirty. In some countries in the Middle East, almost three-quarters of the population are under the age of thirty (Özekin 2014). Özekin (2014) argues that the youth bulge puts an enormous pressure on labor markets, social services and social stability. Muasher (2014) argues that younger cohorts are more likely to feel disenfranchised and more willing to chart a different discourse within their countries, when compared to older cohorts. Through these mechanisms, it is also possible demographic shifts including the youth bulge contributed to the large-scale mobilization seen during the Arab Spring.

These structural issues turn out to be even more conducive to social uprisings when they combine and interact. While dividing age groups into five cohorts, Hoffman (2012) finds that the youngest cohort in the Arab world, ages 18-24, is not only the most unemployed but is the second most educated cohort. Campante (2012) shows that this combination makes the youth cohort most likely to demonstrate. Campante also finds, in four out of four models, that the interaction variable of unemployment and youth population was statistically significant with regime change at a p-value less than 0.01.

Beyond the "big three" structural forces – employment, education, and youth – there exist other plausible variables that are more disputed in the literature. Economic factors, while intuitively pleasing, have come under significant scrutiny within the literature. Hussain (2013) includes two economic measures within his quantitative analysis – average income and wealth distribution – yet only finds that inequality correlates with state fragility in one of his four models. Özekin (2014) also illuminates some nuance by noting that some of the countries that experienced the largest movements, such as Egypt and Tunisia, were experiencing higher growth rates in 2010 than a decade earlier with 5

percent yearly growth, which exceeds European standards. Thus the effects of economic growth and income remain unclear.

Another variable discussed in the literature is governmental structure. Muasher (2014) notes that Arab governments had grown increasingly closed prior to the Arab Spring, banning and suppressing opposition. These governments had become associated with cronyism and the monopolization of wide sectors of the economy, both of which increased public perception of corruption (Muasher 2014). Gelvin (2012) argues that the grievances with government, especially the merging of the ruler, the ruling party, and ruling institutions, are conducive to large-scale social uprisings. The exception to this rule is the monarchies in the Arab world, which proved to be reasonably stable. Muasher (2014) argues that the monarchies are viewed as unifying institutions within their countries and thus protesters demanded change within the existing regime instead of a new regime altogether. Menaldo (2012) finds that the political culture in a monarchy solves the "credible commitment problem" and as a result, monarch experience less political instability.

Various sources within the literature identify other factors that correlate with large-scale uprisings within a country. For example, Hussain (2013) finds that nations with fuel-dependent economies are able to use the wealth gained from the practice to maintain social control. The oil-producing states can devote wealth from their exports to internal security services or to co-opt political opponents and thus are potentially less likely to have experienced large-scale social uprisings. The literature goes on indefinitely and any number of variables are potentially identified to correlate with uprisings. But, the major structural issues with sound theoretical background in the literature have been identified and will be tested in this paper.

Synthesis

Most sources in the literature investigate either ICT or structural variables. This division is understandable because the two explanations are at odds with each other – either ICT or the underlying structural issues caused the Arab Spring protests. However, as a result, the conflict between the two has yet to be resolved empirically.

Gelvin (2012) is one of two sources willing engage with these conflicting variables. Gelvin defines "cyberphiles" as those who argue that ICT played a central role in creating a community of protest in cyberspace. In the literature, cyberphiles present data on the widespread usage of Facebook, Twitter and other communication technology during the events. Gelvin also defines "cyberskeptics" as those push back on these claims by citing low Internet adoption rates and government censorship. Gelvin's cyberskeptic definition can also be expanded to include sources that argue that ICT is irrelevant when compared to the underlying political and economic issues within a country. Gelvin attempts to resolve the arguments of these two groups by noting that while social media played a role in the Arab Spring, it did not cause the uprisings. He states: "Like the printing press and telegraph before them, social media performed two functions in the uprisings: they facilitated communication among the participants and would-be participants who elected to take part in the protests, and they broadened the range of tactical options … open to participants." This view is consistent with the organization and amplification theories but acknowledges the existence of other structural factors that play an equal role in social mobilization.

The other paper that attempts to parse the impact of ICT and underlying conditions is Hussain (2013). Hussain sets up an empirical analysis including ICT variables such as mobile and Internet connectivity, and structural variables such as average incomes, unemployment, and demographic changes. Hussain finds that Internet use contributes to movement success in certain countries, particularly where unemployment is not high or where wealth distribution is unequal. Further, Hussain also finds that mobile phone adoption rates were less important for social movement success than the Internet. As far as structural variables are concerned, he finds that only unemployment and fuel exports significantly correlate with social movement success. From this analysis, Hussain argues that while digital media did not cause the Arab spring, one must acknowledge the strategic uses of digital media. As such, Hussain argues that digital media played a significant role in the Arab Spring by providing the infrastructure for deep communication ties and organizational capacity, connecting millions and bringing the narrative of the social movement across multiple authoritarian regimes.

Argument

My analysis will include three dependent variables: protest size, protest location and protest date. Based on the literature, there are two possible arguments regarding the effects of ICT on these protest variables. The first is put forward by the cyberphiles. Cyberphiles argue that ICTs act as an amplifier and as such one should expect to see larger protests in the Arab Spring cases where the presence of ICT is high. Cyberphiles also argue that the organizational ability of ICTs allows for protests to exist in multiple locations, particularly outside of the capital where protests have traditionally take place. Lastly, cyberphiles argue that the organizational capacity of ICT allows for the improved planning of protests, including the date of protests. If this were the case, one would expect to see more variance in the date of protest with higher ICT penetration, as this means that protests are concentrated on certain days of the week instead of randomly and evenly distributed.

The second possible argument, which I adopt, is the argument of the cyberskeptics. Cyberskeptics maintain that ICT offers no improvement and instead has merely replaced former means of communication. Thus, I do not expect to see differences in the size, location or date of protests in large-scale social uprisings. Instead, I expect structural variables such as unemployment, education, youth demographics and the economy to correlate with larger protests.

Data

The data used in my analysis is a compilation of five datasets. For my independent variables, I used the World Bank's "World Development Indicators." For my dependent variables, I used Beissinger's (2002) "Mass Demonstrations and Mass Violent Events in the Former USSR, 1987-1992" and Francisco's (2011) "European Protest and Coercion" datasets for the USSR cases. For the Arab Spring cases, I used Jenkins and Herrick's (2013) "Protest in the Arab Awakening" and Salehyan and Hendrix's (2014) "Social Confilct in Analysis" datasets. A summary table of my data is shown in *Table 1.* My analysis includes 41 cases with 21 cases from the USSR and 20 from the Arab Spring. Summary statistics tables for just the USSR cases or the Arab Spring cases are located in the appendix.

	Mean	Stdev	Min	Max
GDP Per Capita (PPP)	16,191.06	25,537.03	145.06	136,687.70
Tertiary Education	27.91	14.45	2.78	60.88
Secondary Education	86.31	22.73	7.35	122.90
Youth Unemployment	23.00	11.11	1.53	49.03
Unemployment	11.25	7.53	0.53	40.40
Inequality	34.80	16.38	11	87.35
Youth Under 14	28.96	8.89	13.67	47.33
Youth Under 24	31.45	8.62	17.50	47.67
Fuel Exports	34.06	36.58	0	100
Monarchy	0.20	0.40	0	1
Arab Spring	0.49	0.51	0	1
Internet Usage	19.41	27.27	0	85.00
Fixed Phone Subscriptions	12.64	7.68	0.61	40.52
Cell Phone Subscriptions	54.83	64.16	0	188.67
Protest Size	3.29	1.17	1	5
Protest Location	0.56	0.21	0.19	1
Protest Date Variance	0.01	0.01	0.0003	0.05

Table 1: Summary statistics of forty-one country-cases used in analysis

As far as the units of the structural variables in *Table 1*, GDP per capita adjusted for purchasing power parity is measured in dollars per person. This variable was used to account for the opportunity costs that prevent individuals from joining protests, as mentioned in the literature. I include a variable to account for inequality, measured by the Gini coefficient of each country, in order to control for wealthy countries where wealth is not distributed to the classes who would be protesting on the street. Tertiary and secondary education, or university and high school education, are included in my analysis to account for effects of education on political participation mentioned in the literature. Both types of education are analyzed in order to better understand which age group is included in the "politically aware" class. The units of these variables are the percent of the population that attend university or high school, respectively.

Next, I include unemployment, defined here as the percent of the population who is actively searching but unable to find employment, as well as youth unemployment, which is the unemployment of individuals ages 15 to 24. This variable, along with the two variables that account for the percent of the population under the ages of 14 and 24, are intended to capture the "youth bulge" mentioned in the literature. Lastly, I include a variable to account for fuel exports, which measures fuel exports as a percentage of the country's total exports.

All of these independent variables were coded by computing the average values during a three-year period. For the USSR cases, I took the average of the values from 1989 to 1991 and for the Arab Spring cases I took the average from 2011 to 2013. This was done in order to eliminate the data availability issues and to average out any potential outliers.

I include two binary variables in my analysis. The first is "monarchy," which takes the value of one if the country in question is a monarchy and a zero if it is not. The second is a binary variable for the Arab Spring, which is intended to control for differences between the 1989 and the 2011 cases that are not captured by my other variables.

Next, I have my three ICT variables: Internet users, fixed phone users, and cell phone subscriptions, all calculated as the number of users out of 100 people for each country. Graphs of the distribution of these variables are included in the appendix.

Lastly, my three dependent variables are protest size, location, and date. Protest size was defined as the size of the single largest protest in each country within the given three-year window of analysis. This value was then grouped into five categories based on the number of participants: (1) < 1,000 (2) 1,000 - 10,000 (3) 10,000 - 100,000 (4) 100,000 - 1,000,000 (5) > 1,000,000. This was done because only one dataset had exact estimates of protest size. Further, exact protest size estimations are unreliable, especially within totalitarian states and thus, grouping protests by order of magnitude helps to reduce some of this unreliability. A chart of protest size is shown in *Figure 1* and averages of the Arab Spring and USSR cases are shown in *Figure 2*. After running a t-test, I find a t-value of 3.05, which with 37 degrees of freedom results in a p-value less than 0.01, meaning there exists a statistically significant difference in protest sizes between the USSR and Arab Spring cases. This can be observed in *Figure 2*, where almost all of the USSR cases involve 10,000 people or more and almost all of the Arab Spring cases involve 10,000 people or fewer. A chart showing the maximum protest size by country is included in the appendix.

Maximum Protest Size
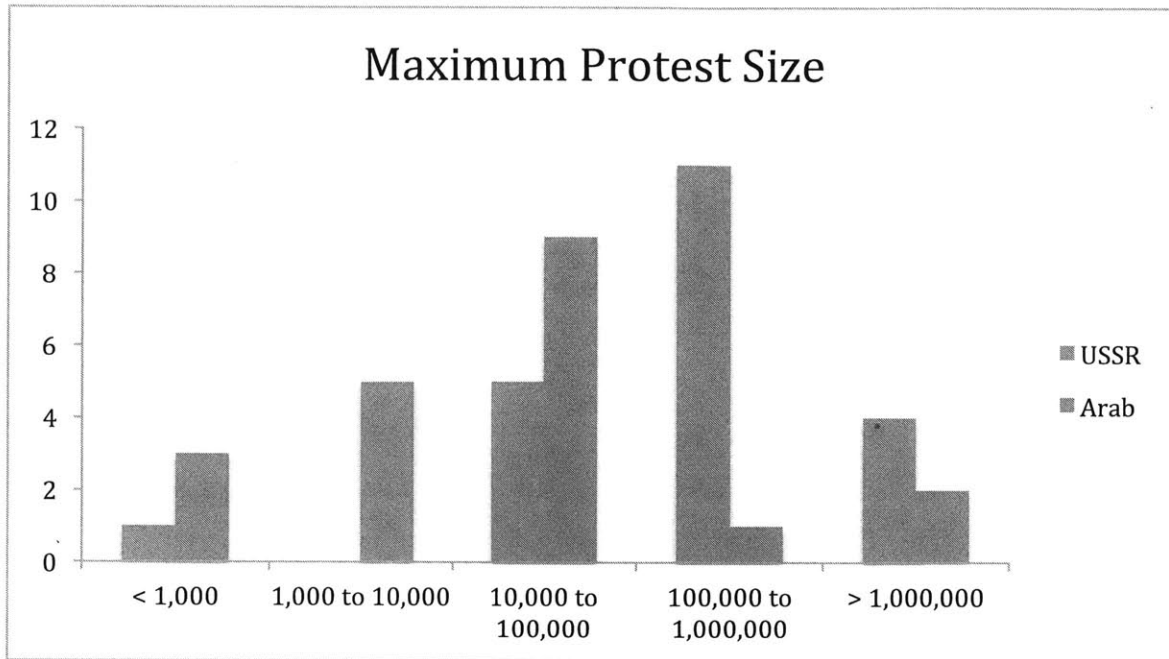
Figure 1: Histogram of maximum protest size per country-case

Maximum Protest Size
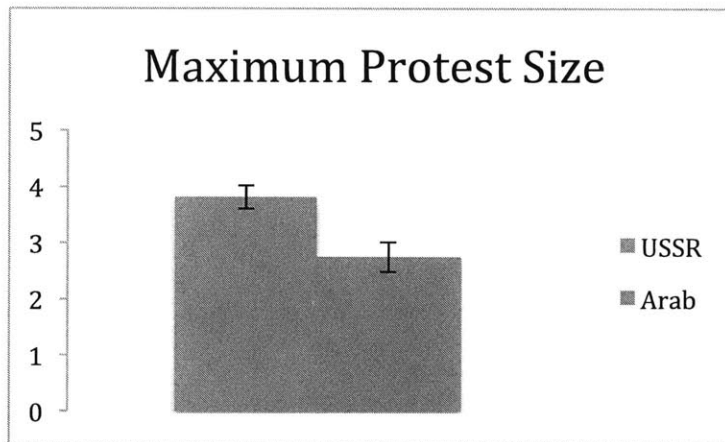
Figure 2: Average maximum protest size. The y-axis represents the five bins of protest size: (1) < 1,000 (2) 1,000 - 10,000 (3) 10,000 - 100,000 (4) 100,000 - 1,000,000 (5) > 1,000,000

Next, I investigate the relationship between protest location and ICT. Based on data availability, I had to use a proxy variable – percent of protests in the capital city – to measure changes in the location of protests. If ICT were improving organization and amplification, then I would expect to see more protests located outside of the capital city. *Figure 3* shows the averages of the two sets of cases. After running a t-test, I found a t-value of 0.45, with 28 degrees of freedom, which results in a p-value greater than 0.1 and no

statistical difference between the two sets of cases. As can be seen in *Figure 4*, there is no observable difference between the locations of protests in the two sets of cases.
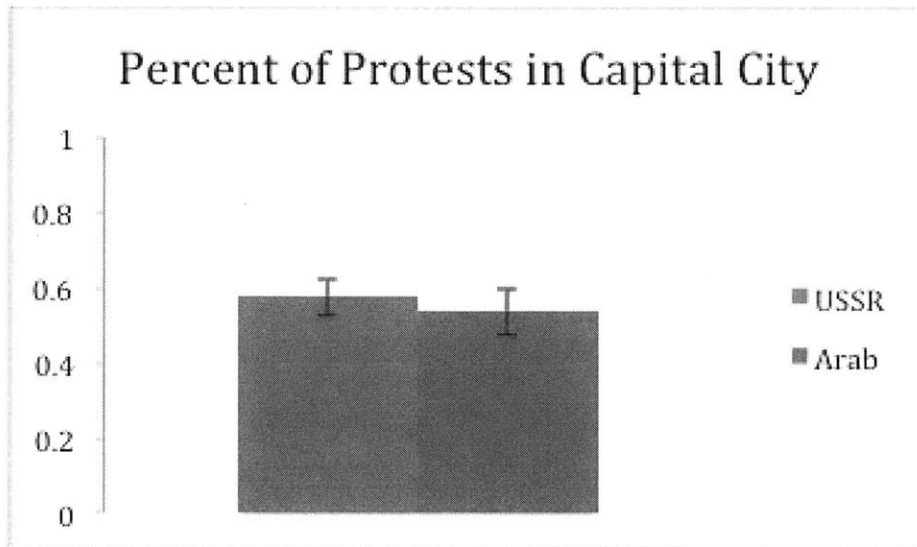
Figure 3: Average percent of protests in capital city
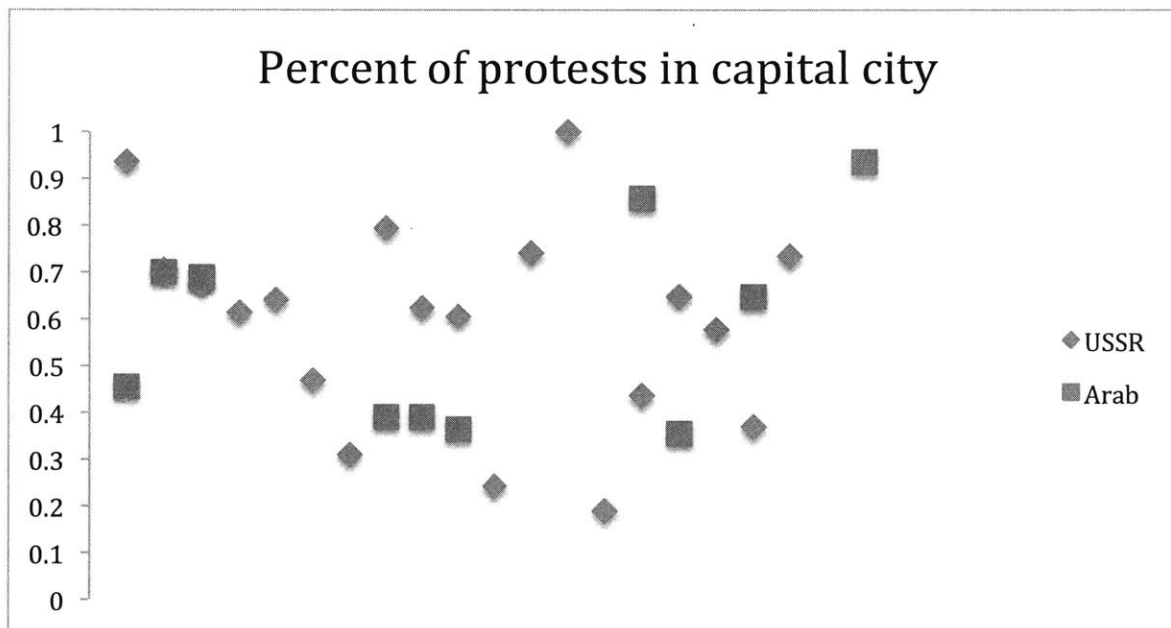
Figure 4: Percent of protests in capital city for each individual case

Lastly, my analysis looks at the effect of ICT on the variance of the day of the week that protests are held. This value was calculated by counting the number of protests with at least 100 participants that occurred each day of the week and then taking the variance of these values. A smaller variance means that most days of the week have a similar number

of protests. I interpret this result to mean that protests are randomly and evenly distributed between all seven days and thus are not organized. Conversely, if the value of the variance is larger, I interpret this result to mean that the protests are better organized, as significantly more protests exist on one or two days and than the rest, which would not occur if protest were randomly distributed. *Figure 5* shows the average variance in the date of protests. After running a t-test, I calculated a t-value of 2.29 with 30 DOF, resulting in a p-value less than 0.05, which means the differences between the two sets of cases is statistically significant. *Figure 6* shows how significant this difference is within the individual cases, with a majority of the USSR cases having equal or less variance than any single Arab Spring case.

Figure 5: Average variance in the date of protest

Figure 6: Variance in date of protest for each individual case

I would have liked to investigate the average frequency of protests, measured by the number of days between protests, because it would have been interesting to know whether ICT accelerates protest organization as is suggested in the lit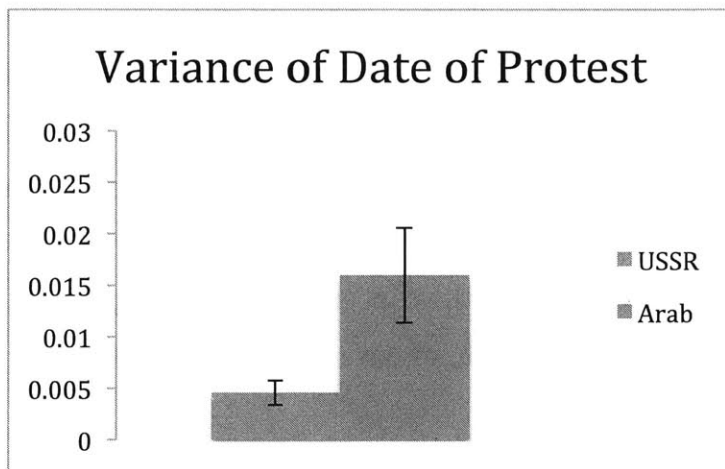erature. However, data availability was an issue. The USSR cases had significantly more protest observations than the Arab Spring cases, perhaps because there has not been enough time to create an extensive protest dataset for all protest events, as has been done in the 1989 cases. For this reason, I could not reliably investigate this topic but it would be an interesting topic to investigate when the data becomes available.

This concern also brings up the issue of data reliability. As previously mentioned, there are some concerns about the reliability protest data because in the totalitarian states studied in this analysis, there is not usually a free press to record protest sizes. This is not to mention that protest sizes are difficult to estimate regardless. Additionally, structural data had to be taken from the most recently available year if otherwise unavailable. Other times, for certain metrics such as the Gini coefficients, models from the literature were used to estimate these values. Despite the concerns about data reliability, I did not want perfect to be the enemy of good enough, so I used as accurate data as was available.

Model and Results

I begin my analysis with a simple correlation test between the ICT variables and the dependent variables. In this test, the correlation between two variables is recorded as one if the variables correlated exactly and positively, negative one if they correlated exactly and negatively, and zero if there is no correlation. Then, the p-values of these correlations are calculated so as to determine statistically significance. *Table 2* and *Table 3* show the results of this analysis.

	Internet	Fixed Phone	Cell Phone	Protest Size	Protest Location	Protest Date
Internet	1.00					
Fixed Phone	0.18	1.00				
Cell Phone	0.85	-0.02	1.00			
Protest Size	-0.40	0.07	-0.41	1.00		
Protest Location	-0.01	0.00	-0.13	-0.05	1.00	
Protest Date	0.42	-0.26	0.54	-0.08	-0.37	1.00

Table 2: Correlation coefficients between variables

	Internet	Fixed Phone	Cell Phone	Protest Size	Protest Location	Protest Date
Internet						
Fixed Phone	0.2617					
Cell Phone	0.0000***	0.9225				
Protest Size	0.0099***	0.6437	0.0073***			
Protest Location	0.9536	0.9978	0.4951	0.7811		
Protest Date	0.0157**	0.1528	0.0015***	0.6565	0.0547*	

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

Table 3: P-values of correlation coefficients between variables

Five sets of variables are found to correlate at statistically significant levels. Two of these correlations are not interesting. Protest location and protest dates correlate, which is not necessarily informative, and Internet usage correlates positively with cell phone subscriptions, which is to be expected.

More interestingly, internet usage and cell phones both correlate negatively with protest size. While this is counterintuitive, recall that *Figure 1* and *Figure 2* show that the Arab Spring protests were on average smaller than the USSR protests. Thus, these correlations are merely reflecting the fact that internet and cell phones are only present in the Arab Spring cases and that the Arab Spring protests are smaller on average than the USSR protests. We can see in *Table 4* that when accounting for the Arab Spring with a binary control variable, neither Internet usage nor cell phone subscriptions correlates at a statistically significant level with protest size.

	Protest Size		
	(1)	(2)	(3)
Internet Usage	−0.006 (0.009)		
Fixed Phone Subscriptions		−0.002 (0.022)	
Cell Phone Subscriptions			−0.0005 (0.006)
Arab Spring	−0.838* (0.490)	−1.064*** (0.338)	−1.008 (0.718)
Intercept	3.810*** (0.231)	3.834*** (0.389)	3.810*** (0.232)
Observations	41	41	41
R²	0.219	0.211	0.211
Adjusted R²	0.178	0.170	0.170

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

Table 4: Initial regression involving ICT variables including a control variable for the Arab Spring and protest size as the dependent variable

In *Table 2* and *Table 3*, Internet and cell phones also correlate with protest date. This time they have positive coefficients, meaning countries with higher ICT usage also saw more variance in the date of protests. This would seem to support the notion that ICT acts as an organizer because more variance means protests are not evenly distributed. But once again, when I include a control variable, I find no statistically significant correlation between Internet or cell phones and protest location.

	Protest Location		
	(1)	(2)	(3)
Internet Usage	0.001 (0.003)		
Fixed Phone Subscriptions		−0.001 (0.005)	
Cell Phone Subscriptions			−0.001 (0.002)
Arab Spring	−0.066 (0.119)	−0.041 (0.089)	0.107 (0.219)
Intercept	0.576*** (0.047)	0.586*** (0.085)	0.576*** (0.046)
Observations	30	30	30
R²	0.011	0.008	0.025
Adjusted R²	−0.062	−0.066	−0.047
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5: Initial regression involving ICT variables including a control variable for the Arab Spring, and protest date as the dependent variable

Two other points worth mentioning from my initial analysis were that fixed phones were not found to correlate at a statistically significant level with any of the dependent variables and that none of the ICT variables correlate with protest location, either in the correlation table or in the regression shown in *Table 6*.

	Protest Date		
	(1)	(2)	(3)
Internet Usage	0.0001 (0.0001)		
Fixed Phone Subscriptions		−0.0002 (0.0003)	
Cell Phone Subscriptions			0.0001 (0.0001)
Arab Spring	0.009 (0.006)	0.011** (0.004)	−0.004 (0.011)
Intercept	0.004* (0.002)	0.007 (0.004)	0.004* (0.002)
Observations	32	32	32
R²	0.243	0.245	0.294
Adjusted R²	0.190	0.193	0.246
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 6: Initial regression involving ICT variables including a control variable for the Arab Spring, and protest location as the dependent variable

These simple correlation tests and regressions comprised my initial analysis. They were intended to give initial insight into the relationship between some of my dependent variables and ICT. Not having found a robust relationship between protest location and protest date and ICT and based on their significantly lower r-squared values, I decided to only continue investigating protest size. Although in my initial model I found no relationship with protest size, the literature provided me with a template to create a better model for protest size.

My initial regression for protest size with only my ICT variable and a binary control variable for the Arab Spring had an r-squared of 0.21. In order to improve the fit of my model, I began creating models with different variations of the principal structural variables in the literature. The results of this process are shown in the appendix. My first eight models included one variable for education, unemployment, and youth in addition to variables for GDP and the binary control for the Arab Spring. In some of these models, GDP per capita was found to be statistically significant but no other variables were.

My next four models included an interaction variable between GDP and education, which acts as a proxy for opportunity costs. The more wealth and more educated an individual is, the less likely they are to organize and protest. As such, at the macroscopic level, I expect that countries with more education and more wealth would be less likely to experience large-scale protests, which is exactly what I find. My GDP and education interaction variable was found to be statistically significant and negative in all of these four models.

Out of my last four models, two models include an interaction variable between unemployment and tertiary education, which attempts to capture the "waithood" phenomenon previously discussed. My last two models include an interaction variable between youth and education, in order to capture the "youth bulge" phenomenon of politically aware youths. However, none of these variables in the last four models was found to be statistically significant.

As mentioned, models 9-12 all had GDP x tertiary education, my magnified opportunity cost variable, correlate negatively and statistically significantly with protest size. Model 12 was found to have an r-squared of 0.349, meaning that it could account for 35% of the variation in protest size among my cases. The adjusted r-squared was 0.277,

which is larger than the initial regression model. For this reason, I decided to use Model 12 as my final regression model with which to test my variables of interest. My sixteen models can be seen in *Table A3* in the appendix.

My final regression models are shown in *Table 7*. These models include the Arab Spring control variable and the four structural variables from Model 12: GDP x tertiary education, youth unemployment, and youth under 14. I then add the following variables: (1) inequality, (2) fuel exports, (3) monarchy, (4) fuel exports x monarchy, (5) Internet usage, (6) fixed phone subscriptions, and (7) cell phone subscriptions. Although all of these variables are found to be statistically significant in the literature, only one is found to correlate with protest size in my analysis.

	Protest Size			
	(1)	(2)	(3)	(4)
GDP x Tertiary Education	−0.00000** (0.00000)	−0.00000** (0.00000)	−0.00000** (0.00000)	−0.00000** (0.00000)
Youth Unemployment	0.017 (0.016)	0.018 (0.016)	0.023 (0.016)	0.017 (0.016)
Youth Under 14	−0.012 (0.021)	−0.015 (0.022)	−0.008 (0.022)	−0.011 (0.022)
Arab Spring	−0.588 (0.431)	−0.838* (0.415)	−0.994** (0.441)	−0.687 (0.418)
Inequality	−0.011 (0.011)			
Fuel Exports		0.003 (0.006)		
Monarchy			0.685 (0.621)	
Fuel Exports x Inequality				−0.0001 (0.0001)
Intercept	4.249*** (0.728)	3.981*** (0.677)	3.768*** (0.690)	3.942*** (0.677)
Observations	41	41	41	41
R²	0.367	0.354	0.371	0.355
Adjusted R²	0.276	0.262	0.281	0.263

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

	Protest Size		
	(5)	(6)	(7)
GDP x Tertiary Education	−0.00000** (0.00000)	−0.00000** (0.00000)	−0.00000*** (0.00000)
Youth Unemployment	0.024 (0.016)	0.017 (0.015)	0.006 (0.015)
Youth Under 14	0.002 (0.024)	0.008 (0.027)	0.006 (0.021)
Arab Spring	−1.443** (0.632)	−0.700* (0.393)	−2.882*** (0.837)
Internet	0.017 (0.013)		
Fixed Phone		0.036 (0.030)	
Cell Phone			0.023*** (0.008)
Intercept	3.474*** (0.755)	2.920** (1.099)	3.839*** (0.616)
Observations	41	41	41
R²	0.381	0.374	0.468
Adjusted R²	0.293	0.285	0.392

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

Table 7: Final regression analysis with introduction of ICT variables

The one variable that is found to correlate at a statistically significant level with protest size is cell phone subscriptions, which does so at a p-value less than 0.01. The coefficient is positive, meaning that more cell phone subscriptions correlate with larger protests,

controlling for all other variables. Further, the r-squared value of this model increases significantly to 0.468 and the adjusted r-squared value jumps to 0.392, meaning my model explains 47% of the variation in protest size. This is an extremely interesting result and will be elaborated on in the discussion section.

Lastly, as a robustness check, I performed a sensitivity analysis on model 12.7. This analysis is shown in *Table 8*. In this sensitivity analysis, I removed one of the four structural variables and re-ran the regression. Unsurprisingly, when I removed the GDP and education interaction variable and the Arab Spring control variable, the r-squared value dropped significantly and cell phones no longer correlate at a statistically significant level. This is most likely because the removal of either of these two variables fundamentally alters my model. When I remove the unemployment and youth variables, cell phones continue to correlate positively and statistically significantly with protest size. This is because the removal either of these variables does not fundamentally alter the model. Based on these results, I believe that the correlation between cell phone subscriptions and protest size is reasonably robust.

	Protest Size			
	(1)	(2)	(3)	(4)
GDP x Tertiary Education		−0.00000*** (0.00000)	−0.00000*** (0.00000)	−0.00000** (0.00000)
Youth Unemployment	0.019 (0.017)		0.007 (0.014)	0.013 (0.017)
Youth Under 14	0.013 (0.024)	0.008 (0.020)		−0.025 (0.021)
Arab Spring	−1.434 (0.863)	−2.929*** (0.819)	−2.772*** (0.745)	
Cell Phone	0.002 (0.007)	0.024*** (0.008)	0.022*** (0.008)	−0.003 (0.004)
Intercept	3.068*** (0.675)	3.899*** (0.588)	3.988*** (0.365)	4.239*** (0.690)
Observations	41	41	41	41
R^2	0.257	0.465	0.466	0.287
Adjusted R^2	0.175	0.406	0.407	0.208

Note: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 8: Sensitivity analysis of final regression model 12.7

Discussion

Both the non-results and the results in this analysis are of analytical interest. I found that my ICT variables did not correlate with more widespread or more organized protests. This result supports the "substitution" hypothesis – that ICT merely replaced older means of organization. In fact, Tufecki (2012) corroborates this result by finding that 48% of

demonstrators first heard about the demonstration through face-to-face contact. Given these results, it appears that in many instances the Internet and cell phones merely act as a different tool for communication and organization, with no discernable results on protest location or date.

As far as the size of protests, Internet usage did not have a statistically significant role while the number of cell phones subscriptions did. This result also corroborates much of the literature. First, Chebib (2011) argues that the low barriers to entry of social media make it more accessible and thus increase its potential impact. While Internet is easily accessible in liberal democratic countries, it is entirely possible that cell phones are more accessible and less regulated in the authoritarian countries studied here. In fact, *Table A1* shows that the average number of cell phone subscriptions per 100 people is 112, meaning on average individuals have more than one cell phone. This number is especially large when compared to an average of 40 internet subscriptions per 100 people in the same sample. Thus, both quantitatively and qualitatively, cell phones are more accessible to individuals in the countries of this study, which potentially increases impact on protest size.

Next, El Tantawy (2011) argues that the decentralization of ICT is important for organizing and sustaining social movements, especially in authoritarian countries. El Tantaway notes that ICT allows for "weak ties" that would otherwise have been restricted by financial, temporal or spatial constraints. Cell phones can potentially be more decentralized in authoritarian states as they are, as their name suggests, mobile. Further, even if Internet usage is blocked, it is not as easy for states to regulate cell phone usage. In this way, phone calls and texts can still be used to organize and amplify, making cell phones more decentralized and potentially more impactful.

According to Stepanova (2011) and Hussain (2013), the Internet and social media is meant to amplify protests to other potential participants. However, by analyzing Twitter links, Aday (2012) found that over 75% of clicks on popular Twitter hashtags were from viewers outside of the region, while only about 10% of clicks were within the country where the protests were taking place. My results compliment this analysis exactly. While the Internet and cell phones can both amplify information, the Internet is more naturally suited to amplify information abroad – websites and hashtags can be clicked on by anyone,

even people you don't know and who are not in your country. Cell phones, conversely, are generally used locally and with people whom one has some form of a connection. If you were interested in increasing the size of a protest, it would not make sense to call or text groups of people outside of your country, and it would probably be quite expensive as well. Instead, you would probably call your family and text your friends – people who are local and who would actually stand a chance of attending the protest. For this reason, it makes intuitive sense that while both technologies amplify protests, cell phones actually increase protest size while the Internet does not.

As far as the other variables in my analysis, the only two structural variables that were found to correlate at a statistically significant level with protest size were GDP and the interaction variable between GDP and tertiary education. These variables should be interpreted as reflecting the effects of opportunity costs as both of their coefficients are negative, meaning the wealthier a country, the smaller protest size would be expected. Interestingly, the GDP per capita variable was statistically significant in six out of eight models whereas the GDP and education interaction variable, which was meant to serve as a magnified opportunity cost variable, was statistically significant in four out of four models. Both of these results are evidence in favor of the opportunity costs hypothesis.

Conclusion

My analysis appears to fall somewhere in the middle of the cyberfile-cyberskeptic spectrum, in contrast to my proposed null hypothesis. I find that ICT does not effect protest location or date and that fixed phones and the Internet have no discernable effect on protest size. But, I do find that cell phones have a statistically significant effect on protest size, perhaps because of their ability to organize and amplify to potential participants. The fact that this result is supported empirically and theoretically makes it an extremely interesting finding and opens the door for significant amounts of future work on the topic. As such, I conclude that cell phones appear not to be a substitution, but a revolution in large-scale social mobilization.

Acknowledgements

Bibliography
- Aday, S., et al. (2012) "New media and conflict after the Arab Spring." *United States Institute of Peace*
- Agre, P, E. (2002) "Real-time politics: The Internet and the political process." *The Information Society*, *18*(5): 311-331.
- Beissinger, M. R. (2002). *Nationalist mobilization and the collapse of the Soviet State.* Cambridge University Press.
- Bennett, W. (2003) "Communicating global activism." *Information, Communication & Society* 6(2): 143-168.
- Campante, F. R., & Chor, D. (2012). "Why was the Arab world poised for revolution? Schooling, economic opportunities, and the Arab Spring." *The Journal of Economic Perspectives*, 167-187.
- Chebib, N. K., & Sohail, R. M. (2011). "The reasons social media contributed to the 2011 Egyptian revolution." *International journal of business research and management (IJBRM)*, *2*(3): 139-162.
- El Tantawy, N., & Wiest, J. B. (2011). "Social media in the Egyptian revolution: reconsidering resource mobilization theory." *International Journal of Communication*, *5*: 18.
- Fisher, D., et al. (2005). "How do organizations matter? Mobilization and support for participants at five globalization protests." *Social problems*, *52*(1): 102-121.
- Gelvin, J. L. (2015). *The Arab uprisings: what everyone needs to know.* Oxford University Press.
- Hussain, M. M., & Howard, P. N. (2013). "What best explains successful protest cascades? ICTs and the fuzzy causes of the Arab Spring." *International Studies Review*, *15*(1): 48-66.
- Menaldo, V. (2012). "The Middle East and North Africa's Resilient Monarchs." *The Journal of Politics*, *74*(3): 707-722.
- Muasher, M. (2014). *The second Arab awakening: and the battle for pluralism.* Yale University Press.
- Özekin, M. K., & Akkaş, H. H. (2014). An Empirical Look To The Arab Spring: Causes And Consequences. *Alternatives: Turkish Journal of International Relations*, *13*.
- Stepanova, E. (2011). "The role of information communication technologies in the "Arab Spring". *Ponars Eurasia*, *15*: 1-6.
- Tarrow, S., & Tollefson, J. (1994). *Power in movement: Social movements, collective action and politics.* Cambridge: Cambridge University Press.
- Tufekci, Z., & Wilson, C. (2012). "Social media and the decision to participate in political protest: Observations from Tahrir Square." *Journal of Communication*, *62*(2): 363-379.

Appendix

	Mean	Stdev	Min	Max
GDP Per Capita (PPP)	26,300.30	33,826.28	145.06	136,687.70
Tertiary Education	27.68	16.83	2.78	60.88
Secondary Education	77.30	29.90	7.35	122.90
Youth Unemployment	26.76	11.59	1.53	49.03
Unemployment	11.68	7.27	0.53	31.10
Inequality	42.17	20.53	11	87.35
Youth Under 14	30.72	9.74	13.67	47.33
Youth Under 24	35.58	7.91	22.33	47.67
Fuel Exports	53.52	40.43	0	100
Monarchy	0.40	0.50	0	1
Arab Spring	1	0	1	1
Internet Usage	39.79	26.68	1.38	85.00
Fixed Phone Subscriptions	11.19	6.81	0.61	22.00
Cell Phone Subscriptions	112.40	43.07	30	188.67
Protest Size	2.75	1.16	1	5
Protest Location	0.54	0.19	0.36	0.86
Protest Date Variance	0.02	0.02	0.004	0.05

Table A1: Summary statistics of twenty Arab Spring country-cases

	Mean	Stdev	Min	Max
GDP Per Capita (PPP)	6,563.21	4,212.62	1,965.74	19,799.57
Tertiary Education	28.13	12.18	7.92	54.15
Secondary Education	94.89	5.05	85.37	103.45
Youth Unemployment	19.42	9.58	2.80	40.40
Unemployment	10.85	7.92	1.05	40.40
Inequality	27.78	5.47	19.75	38.37
Youth Under 14	27.29	7.87	16	44
Youth Under 24	27.52	7.48	17.50	43
Fuel Exports	15.53	19.57	0	66.67
Monarchy	0	0	0	0
Arab Spring	0	0	0	0
Internet Usage	0	0	0	0
Fixed Phone Subscriptions	14.01	8.36	4.58	40.52
Cell Phone Subscriptions	0	0	0	0
Protest Size	3.81	0.93	1	5
Protest Location	0.58	0.22	0.19	1
Protest Date Variance	0.004	0.01	0.0003	0.02

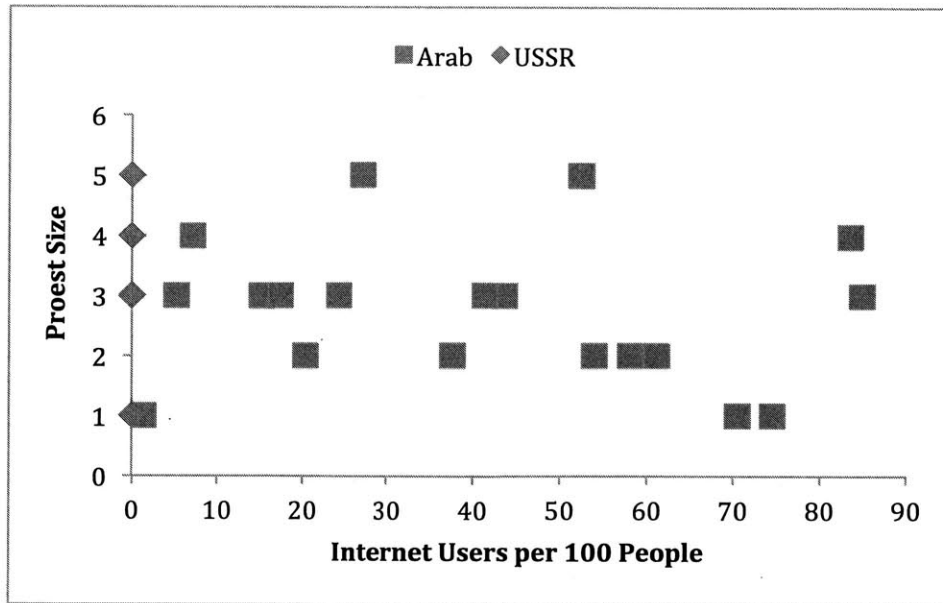*Table A2: Summary statistics of twenty-one USSR country-cases*

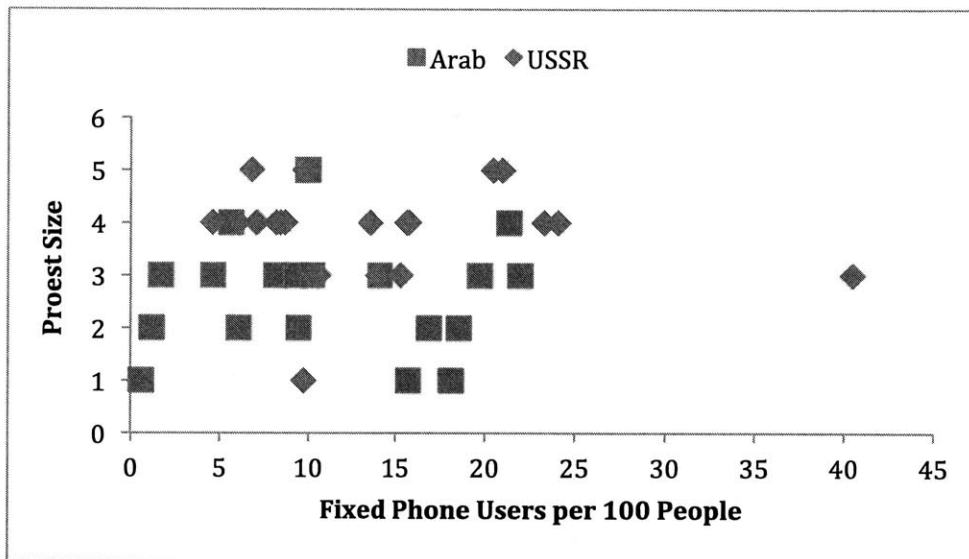*Figure A1: Internet users versus protest size divided by country-case*

Figure A2: Fixed phone users versus protest size divided by country-case
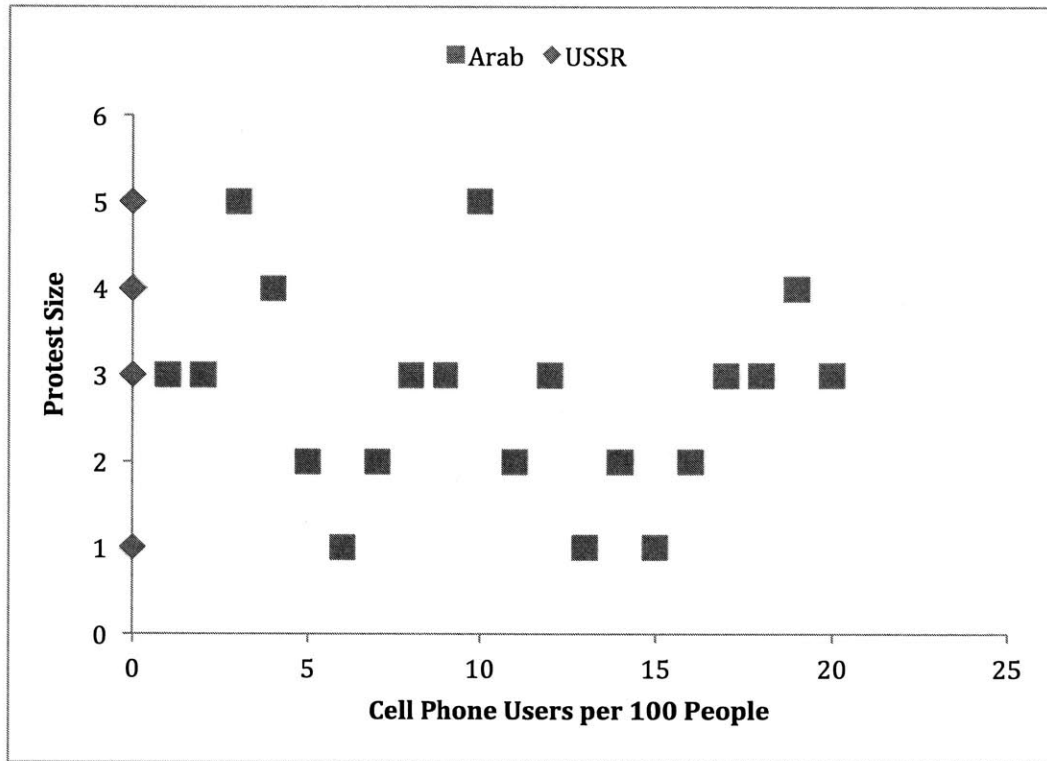
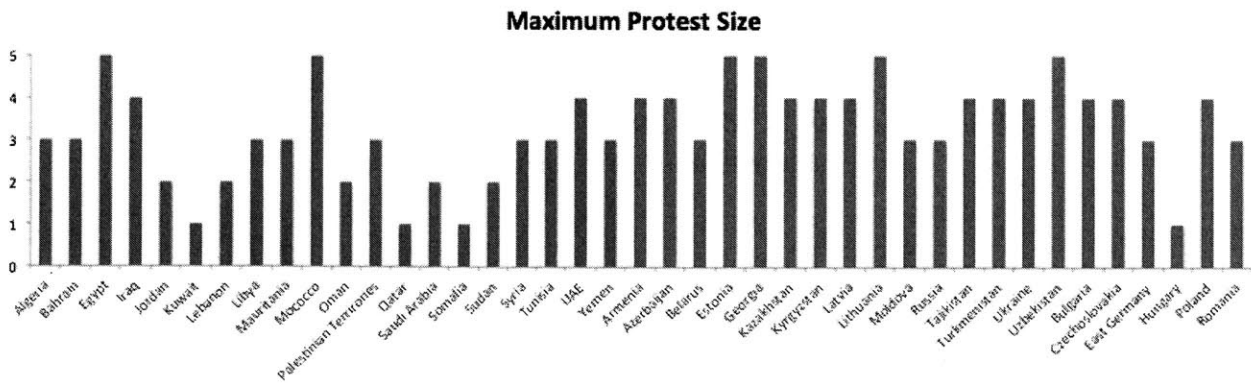*Figure A3: Cell phone users versus protest size divided by country-case*

Figure A4: Protest size divided by country-case. Arab Spring cases are in blue and USSR cases are in Red.

	(1)	(2)	(3)	(4)
GDP Per Capita (PPP)	−0.00002 (0.00001)	−0.00002** (0.00001)	−0.00002* (0.00001)	−0.00002** (0.00001)
Tertiary Education	−0.008 (0.012)		−0.006 (0.012)	
Secondary Education		0.013 (0.010)		0.013 (0.010)
Youth Unemployment	0.012 (0.018)	0.0003 (0.019)		
Unemployment			0.013 (0.025)	0.013 (0.024)
Youth Under 14	−0.016 (0.025)	−0.002 (0.024)	−0.017 (0.025)	−0.004 (0.024)
Arab Spring	−0.778* (0.448)	−0.402 (0.532)	−0.684 (0.405)	−0.430 (0.444)
Intercept	4.344*** (0.928)	2.767** (1.299)	4.399*** (0.916)	2.652* (1.309)
Observations	41	41	41	41
R^2	0.316	0.338	0.314	0.344
Adjusted R^2	0.218	0.243	0.215	0.250

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

	(5)	(6)	(7)	(8)
GDP Per Capita (PPP)	−0.00002 (0.00001)	−0.00002** (0.00001)	−0.00002* (0.00001)	−0.00002** (0.00001)
Tertiary Education	−0.008 (0.012)		−0.006 (0.012)	
Secondary Education		0.013 (0.010)		0.013 (0.010)
Youth Unemployment	0.011 (0.018)	0.0002 (0.018)		
Unemployment			0.013 (0.025)	0.013 (0.024)
Youth Under 24	−0.016 (0.027)	−0.003 (0.025)	−0.017 (0.027)	−0.004 (0.025)
Arab Spring	−0.718 (0.505)	−0.386 (0.568)	−0.621 (0.466)	−0.409 (0.483)
Intercept	4.353*** (0.991)	2.789** (1.290)	4.413*** (0.977)	2.667** (1.301)
Observations	41	41	41	41
R^2	0.315	0.338	0.313	0.344
Adjusted R^2	0.217	0.243	0.214	0.250

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

	(9)	(10)	(11)	(12)
GDP x Tertiary Education	−0.00000** (0.00000)	−0.00000** (0.00000)	−0.00000** (0.00000)	−0.00000** (0.00000)
Unemployment	0.014 (0.023)	0.014 (0.023)		
Youth Unemployment			0.017 (0.015)	0.018 (0.016)
Youth Under 24	−0.013 (0.024)		−0.013 (0.023)	
Youth Under 14		−0.011 (0.022)		−0.013 (0.021)
Arab Spring	−0.606 (0.436)	−0.669* (0.383)	−0.707 (0.443)	−0.768* (0.391)
Intercept	4.155*** (0.727)	4.100*** (0.665)	3.994*** (0.738)	3.963*** (0.669)
Observations	41	41	41	41
R^2	0.332	0.332	0.348	0.349
Adjusted R^2	0.258	0.257	0.276	0.277

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

	(13)	(14)	(15)	(16)
GDP Per Capita (PPP)	−0.00002* (0.00001)	−0.00002* (0.00001)	−0.00001 (0.00001)	−0.00001* (0.00001)
Unemployment x Tertiary Education	0.0002 (0.001)	0.0002 (0.001)		
Youth Under 14	−0.011 (0.023)			
Youth Under 24		−0.010 (0.025)		
Unemployment			0.011 (0.018)	
Youth Unemployment				0.011 (0.024)
Youth x Tertiary Eduation			−0.0001 (0.0004)	−0.0001 (0.0004)
Arab Spring	−0.702* (0.401)	−0.670 (0.459)	−0.852** (0.405)	−0.790** (0.374)
Intercept	4.156*** (0.761)	4.139*** (0.808)	3.787*** (0.461)	3.828*** (0.450)
Observations	41	41	41	41
R^2	0.305	0.304	0.306	0.303
Adjusted R^2	0.228	0.227	0.229	0.226

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

Table A3: Model testing and selection process

Figure A5: Code from R

```
#import dataset
schuman2thu <- read.delim("~/Desktop/schuman2thu.txt", header=FALSE)
schumaninteraction <- read.delim("~/Desktop/schumaninteraction.txt", header=FALSE)

#descriptive statistics
#install.packages("psych")
library(psych)
df = describe(schuman2thu)
mean = df$mean
stdev = df$sd
min = df$min
max = df$max
#create vector with variable names
var_names = c("GDP Per Capita (PPP)", "Tertiary Education", "Secondary Education",
        "Youth Unemployment", "Unemployment", "Inequality", "Youth Under 14",
        "Youth Under 24", "Fuel Exports", "Monarchy","Arab Spring", "Internet Usage",
        "Fixed Phone Subscriptions", "Cell Phone Subscriptions", "Protest Size",
        "Protest Location", "Protest Date Variance")
summary_table = data.frame(row.names = var_names, Mean=mean, Stdev=stdev,
            Min=min, Max=max)

#export to LaTeX
library(stargazer)
stargazer(summary_table, summary = FALSE, digits = 2)

#descriptive statistics (Arab Spring)
#install.packages("psych")
library(psych)
dfa = describe(arab)
meana = dfa$mean
stdeva = dfa$sd
mina = dfa$min
maxa = dfa$max
#create vector with variable names
summary_table = data.frame(row.names = var_names, Mean=meana, Stdev=stdeva,
            Min=mina, Max=maxa)

#export to LaTeX
library(stargazer)
stargazer(summary_table, summary = FALSE, digits = 2)

#descriptive statistics (USSR)
#install.packages("psych")
library(psych)
dfb = describe(ussr)
meanb = dfb$mean
stdevb = dfb$sd
minb = dfb$min
maxb = dfb$max
#create vector with variable names
summary_table = data.frame(row.names = var_names, Mean=meanb, Stdev=stdevb,
            Min=minb, Max=maxb)
```

31

```
#export to LaTeX
library(stargazer)
stargazer(summary_table, summary = FALSE, digits = 2)


#naming columns
gdp = schuman2thu[[1]]
edu_t = schuman2thu[[2]]
edu_s = schuman2thu[[3]]
unemp_y = schuman2thu[[4]]
unemp = schuman2thu[[5]]
ineq = schuman2thu[[6]]
youth_14 = schuman2thu[[7]]
youth_24 = schuman2thu[[8]]
fuel = schuman2thu[[9]]
monarchy = schuman2thu[[10]]
arab = schuman2thu[[11]]
internet= schuman2thu[[12]]
fixed_phone = schuman2thu[[13]]
cell_phone = schuman2thu[[14]]
protest_size = schuman2thu[[15]]
protest_loc = schuman2thu[[16]]
protest_date = schuman2thu[[17]]

gdp_edu_t = schumaninteraction[[1]]
unemp_edu_t  = schumaninteraction[[2]]
fuel_ineq  = schumaninteraction[[3]]
youth24_edu_t  = schumaninteraction[[4]]

#correlation table just IVs
initial_correlation = data.frame(internet, fixed_phone, cell_phone, protest_size, protest_loc, protest_date)

#install.packages("Hmisc")
library(Hmisc)
rcorr1 = rcorr(as.matrix(initial_correlation))

#generate table
# http://www.tablesgenerator.com/

#model with just IVs and Arab Spring binary
modelAA = lm(protest_size ~ internet + arab)
summary(modelAA)
modelBB = lm(protest_size ~ fixed_phone + arab)
summary(modelBB)
modelCC = lm(protest_size ~ cell_phone + arab)
summary(modelCC)
modelDD = lm(protest_loc ~ internet + arab)
summary(modelDD)
modelEE = lm(protest_loc ~ fixed_phone + arab)
summary(modelEE)
modelFF = lm(protest_loc ~ cell_phone + arab)
summary(modelFF)
modelGG = lm(protest_date ~ internet + arab)
summary(modelGG)
```

```
modelHH = lm(protest_date ~ fixed_phone + arab)
summary(modelHH)
modelII = lm(protest_date ~ cell_phone + arab)
summary(modelII)

output_namesAA = c("Internet Usage","Fixed Phone Subscriptions",
        "Cell Phone Subscriptions","Arab Spring", "Intercept")

library(stargazer)
stargazer(modelAA, modelBB, modelCC, covariate.labels = output_namesAA,
    omit.stat = c("ser", "f"), single.row = TRUE)
stargazer(modelDD, modelEE, modelFF, covariate.labels = output_namesAA,
    omit.stat = c("ser", "f"), single.row = TRUE)
stargazer(modelGG, modelHH, modelII, covariate.labels = output_namesAA,
    omit.stat = c("ser", "f"), single.row = TRUE)

#optimizing structural model
model1 = lm(protest_size ~ gdp + edu_t + unemp_y + youth_14 + arab)
summary(model1)
model2 = lm(protest_size ~ gdp + edu_s + unemp_y + youth_14 + arab)
summary(model2)
model3 = lm(protest_size ~ gdp + edu_t + unemp + youth_14 + arab)
summary(model3)
model4 = lm(protest_size ~ gdp + edu_s + unemp + youth_14 + arab)
summary(model4)
model5 = lm(protest_size ~ gdp + edu_t + unemp_y + youth_24 + arab)
summary(model5)
model6 = lm(protest_size ~ gdp + edu_s + unemp_y + youth_24 + arab)
summary(model6)
model7 = lm(protest_size ~ gdp + edu_t + unemp + youth_24 + arab)
summary(model7)
model8 = lm(protest_size ~ gdp + edu_s + unemp + youth_24 + arab)
summary(model8)

#models with interaction variables
model9 = lm(protest_size ~ gdp_edu_t + unemp + youth_24 + arab)
summary(model9)
model10 = lm(protest_size ~ gdp_edu_t + unemp + youth_14 + arab)
summary(model10)
model11 = lm(protest_size ~ gdp_edu_t + unemp_y + youth_24 + arab)
summary(model11)
model12 = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab)
summary(model12)
model13 = lm(protest_size ~ gdp + unemp_edu_t + youth_14 + arab)
summary(model13)
model14 = lm(protest_size ~ gdp + unemp_edu_t + youth_24 + arab)
summary(model14)
model15 = lm(protest_size ~ gdp + unemp_y + youth24_edu_t + arab)
summary(model15)
model16 = lm(protest_size ~ gdp + unemp + youth24_edu_t + arab)
summary(model16)

output_names1 = c("GDP Per Capita (PPP)", "Tertiary Education", "Secondary Education",
        "Youth Unemployment", "Unemployment", "Youth Under 14","Arab Spring", "Intercept")
```

```
output_names2 = c("GDP Per Capita (PPP)", "Tertiary Education", "Secondary Education",
        "Youth Unemployment", "Unemployment", "Youth Under 24","Arab Spring", "Intercept")

output_names3 = c("GDP x Tertiary Education", "Unemployment", "Youth Unemployment",
        "Youth Under 24","Youth Under 14", "Arab Spring", "Intercept")

output_names4 = c("GDP Per Capita (PPP)", "Unemployment x Tertiary Education",
        "Youth Under 14", "Youth Under 24", "Unemployment",
        "Youth Unemployment", "Youth x Tertiary Eduation",
        "Arab Spring", "Intercept")

# use to set margin witdh \usepackage[margin=0.5in]{geometry} BEFORE \begin{document}
library(stargazer)
stargazer(model1, model2, model3, model4, covariate.labels = output_names1,
      omit.stat = c("ser", "f"), single.row = TRUE)
stargazer(model5, model6, model7, model8, covariate.labels = output_names2,
      omit.stat = c("ser", "f"), single.row = TRUE)
stargazer(model9, model10, model11, model12, covariate.labels = output_names3,
      omit.stat = c("ser", "f"), single.row = TRUE)
stargazer(model13, model14, model15, model16, covariate.labels = output_names4,
      omit.stat = c("ser", "f"), single.row = TRUE)

#model 12 selected as best fitting model
# use to set margin witdh \usepackage[margin=0.5in]{geometry}
#test three additional variables (inequality, fuel exports, monarchy, fuel x inequality)
model12A = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab + ineq)
summary(model12A)
model12B = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab + fuel)
summary(model12B)
model12C = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab + monarchy)
summary(model12C)
model12D = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab + fuel_ineq)
summary(model12D)

output_namesX = c("GDP x Tertiary Education", "Youth Unemployment",
        "Youth Under 14", "Arab Spring", "Inequality", "Fuel Exports",
        "Monarchy", "Fuel Exports x Inequality", "Intercept")

library(stargazer)
stargazer(model12A, model12B, model12C, model12D, covariate.labels = output_namesX,
      omit.stat = c("ser", "f"), single.row = TRUE)

#communication variables (interent, fixed phone, cell phone)
model12E = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab + internet)
summary(model12E)
model12F = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab + fixed_phone)
summary(model12F)
model12G = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + arab + cell_phone)
summary(model12G)

output_namesY = c("GDP x Tertiary Education", "Youth Unemployment",
        "Youth Under 14", "Arab Spring", "Internet","Fixed Phone","Cell Phone", "Intercept")

library(stargazer)
stargazer(model12E, model12F, model12G, covariate.labels = output_namesY,
```

```
       omit.stat = c("ser", "f"), single.row = TRUE)

#Sensitivity Analysis
models1 = lm(protest_size ~ unemp_y + youth_14 + arab + cell_phone)
summary(models1)
models2 = lm(protest_size ~ gdp_edu_t + youth_14 + arab + cell_phone)
summary(models2)
models3 = lm(protest_size ~ gdp_edu_t + unemp_y + arab + cell_phone)
summary(models3)
models4 = lm(protest_size ~ gdp_edu_t + unemp_y + youth_14 + cell_phone)
summary(models4)

output_namess1 = c("GDP x Tertiary Education", "Youth Unemployment",
        "Youth Under 14", "Arab Spring", "Cell Phone", "Intercept")

library(stargazer)
stargazer(models1, models2, models3, models4, covariate.labels = output_namess1,
        omit.stat = c("ser",."f"), single.row = TRUE)
```