The Unintended Consequences of Internet Diffusion: Evidence from Malaysia

Luke Miner[1]

First Draft: November 2011

Current Draft: September 2015

Abstract

Can the introduction of the Internet undermine incumbent power in a semi-authoritarian regime? I examine this question using evidence from Malaysia, where the incumbent coalition lost its 40-year monopoly on power in 2008. I develop a novel methodology for measuring Internet penetration, matching IP addresses with physical locations, and apply it to the 2004 to 2008 period in Malaysia. Using distance to the backbone to instrument for endogenous Internet penetration, I find that Internet exposure accounts for 6.6 points, nearly half the swing against the incumbent party in 2008. I find limited evidence of increased turnover, and no evidence of an effect on turnout.

JEL Classification: D72, O33, L82, R11

Keywords: Internet diffusion, political economy of the media, Malaysian elections.

[1] Crowdpac, 1295 El Camino Real, Menlo Park, CA, USA. Email: lminer@gmail.com

[1] I am extremely grateful to Robin Burgess, Greg Fischer, and Gerard Padró i Miquel for their encouragement and support throughout this project. I would like to thank Jessica Powell for her tireless support and invaluable copy-editing. Jayanath Appudurai, Oriana Bandiera, Tim Besley, Graham Brown, Premesh Chandran, Bill Cheswick, KC Claffy, Edward Clayton, Charif El-Ansari, Ruben Enikolopov, Maitreesh Ghatak, Norizan Ghazali, Edmund Terrence Gomez, Steve Gibbons, Hayati Harudin, Jason Ketola, Brian Knight, Erik Kline, Koay Hock Eng, Valentino Larcinese, Alan Manning, Thomas J. Mather, George Michaelson, Kian Ming Ong, Yeong-Chuan Lim, Henry Overman, Torsten Persson, Maria Petrova, Steve Pischke, Danny Quah, Ramasamy Ramachandran, Yasuyuki Sawada, William Stevenson, Ibrahim Suffian, Dr. Fadhlullah Suhaimi Abdul Malek, Dimitri Szerman, Tan Tze Meng, Hansen Toh, Tom Vest, Georg Weizsäcker, and Dieter Zinnbauer provided valuable advice and comments.

1. INTRODUCTION

Since its inception, much has been made of the Internet's potential as a democratizing force that frees information from the control of governments, implodes the distance between users around the world, and provides access to new viewpoints. Indeed, the Internet's ability to provide unfiltered access to information has caused consternation among many governments. This response has been notable in China, which has invested billions in keeping exposure to the Internet under tight rein. Social media has also been identified as a driving factor behind protests the world over, such as the recent revolutions across the Middle East. Despite a wealth of anecdotal evidence, however, little quantitative work has been conducted to test the ability of the Internet to foster democratization. The primary objective of this paper is to begin to fill this gap.

Malaysia serves as a particularly compelling test case in this regard. First, the ruling coalition, the Barisan Nasional (BN), enjoyed veto-proof control over all branches of government from 1969 to 2008. Although Malaysia holds regular democratic elections, the BN maintained power through strict controls on the judiciary, the police, and, importantly, the mass media.

Second, the BN's hold on power was so secure that it initiated an aggressive information and communications technology (ICT) led development strategy, based on an uncensored Internet. The government has invested heavily in the ICT sector since 1996 as a means to promote growth and enjoys a very high rate of Internet penetration—60% as of 2008. At the same time, to attract foreign direct investment (FDI), the government pledged not to censor the Internet. Third, since the Internet is uncensored, it has become home to a vibrant opposition blogosphere and a number of popular, independent news sites.

In March 2008 the BN lost its two-thirds majority in parliament for the first time since 1969, as well as control of 5 out of 13 states. In the aftermath, commentators argued that the Internet played a leading role in this outcome by providing access to alternative viewpoints. In the rush to promote an information economy, the government overlooked the consequences with regard to political control. This paper tests whether Internet penetration influenced voting behavior in Malaysia, focusing on the 2004 and 2008 elections.

An important contribution of this paper is a novel measure of Internet penetration, which can be applied to almost any country. For most countries there are no geographically disaggregated measures of change in Internet access across time. To address this problem, I use a dataset that maps all of the IP addresses in Malaysia to approximate geographical locations. I aggregate the data up to the yearly period to deal with changes in assignment location across months. Next I use inverse-distance weighting interpolation to convert the data from the city level to the state legislature district level. Finally, I normalize by the number of eligible voters in a district to create the final measure. I find that this measure performs well when tested against census data from 2004.

To address problems of endogenous Internet placement and confounding political trends, I instrument for Internet growth. I calculate the shortest distance from each electoral district to the backbones of Malaysia's main Internet Service Providers (ISPs). An increase in distance to the backbone leads to higher costs of supplying Internet connectivity (e.g., digging new trenches and laying cabling). This provides exogenous variation in Internet supply across districts. I exploit differences across ISPs in terms of geographical constraints on the placement of their backbones to argue that distance to the backbone is unlikely to affect voting outcomes, conditional on covariates. The identifying assumption is that conditional on baseline district characteristics (ethnic distribution, GDP per capita, population density) distance to the backbone does not affect change in vote share independently of growth in Internet access.

Based on the identifying assumption, I show a large effect of Internet growth on election results: that the Internet can explain 6.6 points of the 14 point swing against the BN in the 2008 elections for state legislatures in peninsular Malaysia. I find limited evidence that Internet growth led to increased turnover and no evidence of an effect of Internet exposure on turnout.

To the best of my knowledge, this paper is the first to measure the Internet's effects on elections in a developing country. It relates most directly to literature on the political economy of the Internet: Gentzkow & Shapiro (2011) look at the effects of the growth of new media on ideological segregation in the U.S.A; Czernich (2012) finds a positive effect on turnout; Jaber (2013) finds a positive effect on turnout and presidential vote share; Enikolopov *et al.* (2012) finds an effect of corruption reporting in blogs on stock prices. Falck *et al.* (2014) finds a negative effect of Internet on turnout. Campante *et al.* (2013) also finds a negative effect of broadband on turnout, however, over time this negative effect reverts due to the endogenous response of the supply-side of politics to new media platforms.

This paper also relates to literature on the political economy of mass media technologies. (DellaVigna & Kaplan, 2007; Enikolopov *et al.* , 2011; Gentzkow, 2006; Durante & Knight, 2012; Durante *et al.* , 2014) for television; (Besley & Burgess, 2002; Snyder & Strömberg, 2010; Gentzkow *et al.*

, 2011; Chiang & Knight, 2011; Gerber *et al.* , 2009; Drago *et al.* , 2014) for newspapers; and (Stromberg, 2004; Della Vigna *et al.* , 2014; Adena *et al.* , 2015) for radio.

More broadly, this paper relates to literature on the effects of information technology on development. Jensen (2007) looks at the effects of the introduction of mobile phones on fish markets in Kerala; Goyal (2010) similarly analyzes the effects of Internet kiosks on crop prices in Madhya Pradesh; and Jack & Suri (2014) explore the impact of mobile payment on informal risk sharing.

I start in section 2 by outlining a general theoretical framework to help understand the main mechanism at play. Section 3, shows how the theoretical framework pertains to Malaysia, providing background on politics, media, and the Internet. In section 4, I describe my data sources, before outlining my method for constructing a measure of Internet penetration. Section 5 presents my empirical strategy results. I start by exploring the strong correlation that exists in the data and then move on to the issue of identification. In section 6, I examine additional outcomes. Section 7 provides intuition for the effects and section 8 concludes.

2. THEORETICAL FRAMEWORK

In this section I discuss a simple framework for understanding how the Internet influences voting outcomes when all conventional sources of information are government-controlled. This is primarily accomplished by extending the insights from Besley & Prat (2006) to account for differences between traditional media outlets (e.g., TV, radio, print) and web-based media outlets.[1]

Besley & Prat (2006) presents a two-period retrospective voting model, in which voters cannot directly monitor candidate performance, but must instead get their information on candidate quality from the media. The media faces two sources of profit: commercial profit and profit from collusion with the government. An incumbent can pay to capture each media outlet for a cost. Provided that the payoff to holding office exceeds the cost, an incumbent can "capture" the media, suppressing voter information on candidate quality and assuring reelection. In equilibrium an incumbent must capture all media outlets to assure reelection, since if one outlet deviates, all the electorate is informed.

The Internet can be modeled in this framework as a media outlet that differs from traditional outlets in two ways. First, it is too costly to capture (I provide a justification for this assumption in the context of Malaysia in section 3.2). Second, the Internet is only accessible to a fraction of

[1] See web appendix for complete model.

the population. If this fraction is less than fifty percent, incumbents can continue to win elections by capturing the traditional media, as less than half the population will become informed and incumbents only need a simple majority to win.

An equilibrium with capture can be sustained even if Internet penetration is above fifty percent. Two factors can raise the threshold of Internet penetration under which capture is still possible. First, Internet penetration is not uniformly distributed across districts, but rather rightward skewed so that a majority of districts have less than 50% penetration, but average penetration for the country as a whole is above 50%. Second, if population differs substantially across districts and more populous districts have higher rates of Internet penetration, the majority of the population can have Internet, but the incumbent can still win because populous, Internet connected voters count less than rural voters. In section 4.3, I show that the distribution of Internet in Malaysia is rightward skewed and Internet access is higher in districts with higher fractions of the total population. This suggests the incumbent party can win elections by capturing the media even when Internet access is greater than 50% across the country as a whole.

The main empirical prediction is that an increase in Internet access will cause a decrease in the incumbent party's vote share, in the presence of media capture. Intuitively, the Internet allows voters to circumvent media controls, and thus enables them to receive negative signals on candidate quality. The incumbent party's vote share will shrink as an increasing fraction of the population gains access to negative signals. I test this prediction in detail in section 5.

A secondary implication of Internet growth leading to lower incumbent vote shares is that it should also yield higher turnover. Section 6 finds limited evidence corroborating this prediction. Full details of the model can be found in the web appendix.

3. BACKGROUND

3.1. *Political regime*

The theoretical framework above provides a useful method for thinking about the Internet's effect in Malaysia. Classified variously as "partly free"[2], a "flawed democracy"[3], and a "pseudo-democracy"[4], Malaysia's political regime combines democratic and autocratic elements.

[2]See Freedom House: <www.freedomhouse.org>
[3]See Economist Intelligence Unit: <www.eiu.com>
[4]See Case (2001)

Malaysia is a federation of thirteen states with a parliamentary system of governance. Elections are first-past-the-post and occur for both the national parliament and each state legislature.

Since independence, Malaysia has been ruled by the same coalition in various guises, the Barisan Nasional (BN, or the Alliance prior to 1969). Though the BN includes parties representing minorities, it is effectively run by the United Malays National Organization (UMNO). Malaysia has had opposition parties since independence, but only in recent years have they posed a real threat to the BN's hegemony. Led by ex-Deputy Prime Minister Anwar Ibrahim the opposition is made up of an alliance between a secular, predominantly Malay party (PKR), a Chinese secular party (DAP), and an Islamist party (PAS).

As in the theoretical framework in section 2, the incumbent coalition has captured the print and broadcast media. Media ownership is concentrated in a handful of conglomerates that are controlled by the government, constituent members of the BN, and closely connected businessmen. For example, UMNO founded and controls the Utusan Group, which includes the *Utusan Melayu*, the oldest and most widely distributed Malay daily. Although the traditional media show some variation in their level of bias, in general they tend to under-report on opposition candidates and downplay scandals.[5]

In addition, strict legal restrictions on media outlets prevent the emergence of any mainstream outlet that is overly critical of the government. First, media firms can only operate with a permit and face tightly controlled distribution. Opposition parties are denied permits to publish newspapers, even though constituent members of the BN control multiple media outlets. Second, laws such as the Sedition Act, the Control of Imported Publications Act, and the Official Secrets Act allow the government to censor material with impunity.[6] Finally, the Internal Security Act, enacted in 1960 to fight a Communist insurgency, allows detention without trial for up to two years and can be renewed indefinitely.[7]

[5] See Centre for Independent Journalism Malaysia (2008) or <www.malaysiakini.com/news/168567> for specific examples.

[6] For example, the Home Ministry censored an article in the July 16th, 2011 issue of the Economist on an electoral reform rally.

[7] As of September 2011, an announcement was made to reform these laws. See section 3.2 for details.

3.2. *The Internet and politics*

The second essential parallel to the theoretical framework lies in Malaysia's treatment of the Internet. In stark contrast to print, radio, and television, the Internet has never experienced significant censorship.

By 2008, Malaysia had a very high rate of Internet access: Internet users comprised 56% of the population[8] with 72% of users spending more than four hours online as of 2008.[9] Malaysia's high Internet penetration rate stems from its 1996 decision to invest heavily in ICT infrastructure as a way to foster a knowledge-based economy. To make Malaysia more attractive to FDI, the government signed an Internet "bill of rights", pledging not to censor the Internet.[10]

As a result, the Internet is the only platform available for alternate view points and has become an important source for independent news and opposition news and views. As early as 1999, members of the opposition used blogs and newsgroups to spread their message, and there is some evidence that the Internet influenced the 1999 general elections.[11] These blogs were soon joined by official opposition sites like PAS owned *Harakah Daily* and independent online news sites like Malaysiakini, which were not constrained by strict censorship laws constraining their print counterparts.

In terms of the theoretical framework, the Internet's most salient feature is its ability to provide information about scandals that previously would have been suppressed. The best example, the V.K. Lingam video uploaded to YouTube in late-2007, showed high-level officials engaged in judicial fixing for the Supreme Court. The video received millions of hits in a matter of days and erupted into one of the defining issues of the 2008 elections.

An important assumption in the theoretical framework is that the Internet is much more costly to capture than traditional media. There are a number of factors contributing to the cost of Internet capture. First, the diffuse nature of the Internet makes it costly to regulate. Second, from a purely economic standpoint, in order to engage in censorship, the government would have to go against the Internet bill of rights signed in the nineties, and potentially scare away the FDI needed for Malaysia's ICT-based growth strategy. Efforts could be made to censor the Internet, as in China's "Great Firewall", but the cost would be significant. Such censorship would require substantial

[8] See World Bank Development Indicators: <data.worldbank.org>

[9] See Malaysian Communications and Multimedia Commission (2013)

[10] See MSC Malaysia Bill of Guarantees at <www.mscmalaysia.my.

[11] See Zinnbauer (2003)

investment in not only physical capital but also human capital.[12]

3.3. *Internet placement*

As of 2008 there were three players in the provisioning of Internet infrastructure: Telekom Malaysia (TM), the state-owned incumbent, which is measured with the least accuracy due to lack of detail in public records; Time, a private company, geared toward consumers and businesses; and Fiberail, running the length of Malaysia's major railways, and jointly owned by TM and Malaysia's public railway service.

Annual reports, consultant reports, and interviews concur that cost plays a central role in governing placement. The costs of delivering the Internet to consumers can be divided into several categories. First is the cost of installing the backbone, the trunk lines, nodes, and routers that form the core of an ISP's network. In terms of geography, costs include digging trenches so that the fiber-optic cabling can be laid underground. These trenching costs depend on the terrain: it is much more expensive to lay fiber-optic through a jungle than alongside a road. As a result, all three backbones follow preexisting routes: roads and highways in the case of TM and Time, and railways in the case of Fiberail. Moreover, most land-based trunk cabling runs along federal and not state roads, since it is much less costly and time-consuming to secure a license from the federal government than from state governments.

Once the backbone has been laid, plenty of supplementary costs must be incurred before an ISP can deliver its service to consumers. To serve a new area, an ISP must install a local switch and connect it to the backbone via backhaul cable. This step adds further trenching costs, which increase with distance to the backbone. It also entails the costly and time-consuming process of getting permission from local authorities. Even TM, which owned an extensive telephone network before the advent of the Internet, faces these costs. TM had to upgrade much of its copper wire to carry data signals, and dig up and replace its backhaul cable with fiber-optics to provide the extra capacity and speed needed to delivery Internet.

[12] China has an army of Internet police whose sole job is to peruse forums, blogs, search results, etc. for objectionable content. See <http://www.guardian.co.uk/technology/2005/jun/14/newmedia.china>

4. DATA

4.1. *Political Data*

This paper uses election data at the state legislature district level for the 1986, 1990, 1995, 1999, 2004, and 2008 elections. The data includes candidate names, parties, and votes along with turnout, the number of eligible voters in a district, the number of rejected votes, and the district's ethnic composition. Significant redistricting occurred between 1990 and 1995 and between 1999 and 2004. I have manually entered each set of electoral boundaries into ArcGIS to account for changes in district size and number since 1986. Because of better data availability, I focus on peninsular Malaysia, excluding the states of Sabah and Sarawak. Kuala Lumpur is excluded from the study because it can only vote in Parliamentary elections.[13]

As of 2008, there were 445 state legislature districts in peninsular Malaysia. I drop 6 because of uncontested elections. Table I provides summary statistics covering the 2004-2008 period for state legislature districts in peninsular Malaysia, excluding Kuala Lumpur. The 2008 election is marked by a large drop in vote share for the BN and a modest increase in turnout. The number of eligible voters varies significantly across districts, with a mean of 18,000 and a standard deviation of around 7,000.

Peninsular Malaysia is multi-ethnic: 63% Malay, 28% Chinese, and 8% Indian.

4.2. *Demographic and Geographic Data*

I have complete geospatial data for Malaysia: clutter data (which classifies all land as either urban, semi-urban, plantation, jungle, inland water, or open); elevation data (which allows for the calculation of land-gradients); the locations of all major roads, highways, and railways in Malaysia; and data from the LandScan service, which estimates population distribution at the one square kilometer resolution through a combination of census data and satellite imagery.[14]

Table I helps make sense of the geospatial data. State legislature districts, on average, are 21% urban and 50% farmland (rural), with the remaining 29% classified as jungle. Although jungle covers large swaths of the country, the fairly extensive road network spans more than 80,000 kilometers of roads as of 2007.

[13] If Parliamentary results are used for Kuala Lumpur, my findings are even stronger.

[14] See <http://www.ornl.gov/sci/landscan/> for details on the construction of this dataset.

I have constructed a dataset of controls using the Population and Housing Census of Malaysia for 1980, 1991, and 2000; Malaysia's Household Basic Amenities and Income Survey for 2004; and geographically disaggregated measures of GDP per capita 2005, generated by the consultancy Booz & Company. Unless otherwise stated, this data is available at the level of Malaysia's 927 census districts, called *mukim*.

Mukim-level data does not match up perfectly with state legislative districts. To address this discrepancy, I use the LandScan population data to assign a weight to each one kilometer cell within each mukim. State electoral district values are generated from the weighted sum of these one square kilometer cells.

4.3. *Internet Data*

I use official Internet measures from the Population and Housing Census 2000 and the Household Basic Amenities and Income Survey (HBAIS) for 2004. Both datasets provide the fraction of households with Internet subscriptions at the mukim (census district) level. As explained in section 4.2, I use ArcGIS to aggregate the HBAIC data to the legislative district level, which introduces some measurement error.

In the theoretical framework in section 2, I presented two sufficient conditions for media capture when average Internet access is greater than 50%. The first condition is rightward skew in the distribution of Internet connectivity by district. Figure 1(a) shows an approximation of the PDF for Internet subscription per household variable alongside the PDF of a normal distribution. As can be seen, the distribution is severely rightward skewed. The second condition is a positive correlation between Internet penetration and the fraction of total population. Figure 1(b) graphs a scatter plot of the log of households with Internet subscriptions in 2004 against the fraction of total eligible voters in a district. Showing a strong positive relationship between these two variables, this graph implies that districts with larger populations also have higher Internet penetration per capita.[15]

Since the census data does not cover the 2008 period, I turn to two extra sources of data on Internet connectivity. The first, the GeoIP City database, is produced by the geo-location company MaxMind. GeoIP City is a service that matches IP addresses to geographical locations, allowing web services to tailor advertisements based on visitor location and to detect fraud. The GeoIP City database comprises monthly data from 2004 to the present and covers virtually all IP addresses in

[15]These results hold for alternate measures of Internet penetration for 2004 and 2008 explained in web appendix.

the world.[16] For each IP address assigned to Malaysia, the GeoIP City database provides the name and location of the nearest large city on a monthly basis. There are 782 locations that appear in the data for 2004 and 487 for 2008. Although the 2008 data has fewer locations, it has roughly twice as many IP addresses, reflecting the enormous growth in Malaysia's Internet penetration in the 2004-2008 period.

My second data source comes from APNIC (Asia-Pacific Network Information Center), the regional Internet registry responsible for delegating blocks of IP addresses to national Internet registries, ISPs, and large companies in the Asia-Pacific region. As such it has a complete record of all IP blocks allocated to Malaysia along with the recipient of the block (normally an ISP) and the date of allocation.

The GeoIP City database is used in conjunction with the APNIC dataset, which together identify: the initial date of IP assignment to Malaysia, the ISP managing the IP addresses, and the IP blocks location(s) during the 2004-2008 period. I create the measure $IPperVoter$ by aggregating this data up to the electoral district level and then normalizing by the number of eligible voters. This gives the number of IP addresses per voter in each state legislature district and is expressed in logs. I drop 12 districts in which $IPperVoter$ growth is negative, suggesting large measurement error as Internet penetration nearly doubled across the country during this time. Unless stated otherwise, results are largely unchanged by including these 12 districts. This leaves 427 state legislature districts.

I check $IPperVoter$ for 2004 against official government statistics on the percentage of households with Internet subscriptions in 2004, and find a high correlation of .63. Although this result indicates a strong correlation, it still leaves a large, unexplained difference between the two measures. There are several explanations for why this difference occurs. First, the survey data only counts home subscriptions, whereas $IPperVoter$ would also count work connections.

Second, there are likely a large amount of IP addresses that are counted by $IPperVoter$ but are in fact unused due to inefficient IP address allocation during the early years of the Internet. For example, out of the four million IP addresses assigned to Malaysia as of 2008, the Malaysian government owned oil company, Petronas, a division of Shell based in Sarawak, and UTM, a Malaysian University, each controlled sixty thousand IP addresses, all of them assigned in the nineties.[17] This suggests that $IPperVoter$ is a better measure of Internet growth, which differences

[16] MaxMind does not cover IPv6 addresses. However, IPv6 adoption was infinitesimal in Malaysia at the time.

[17] See APNIC records as of April 2015.

out these long disused blocks, than level of Internet penetration.[18]

5. EMPIRICAL ANALYSIS

5.1. *Basic Correlations: OLS Estimates*

I start by examining the basic relationship between Internet penetration and the BN's share of the vote at the state legislature district level. Figure 1(c) plots change in voting share for the BN and growth in $IPperVoter$ during the 2004 to 2008 period. As can be seen, there is a strong negative relationship in the raw data, implying that areas with more Internet growth are associated with greater negative swings against the BN.

I explore this relationship in more detail by controlling for other characteristics that might affect changes in BN voter share. Let y_{ist} be the BN's vote share for legislative district i in state s at time t and $IPperVoter_{ist}$ be the log of the number of IP addresses per voter:

$$\triangle y_{ist} = \alpha_1 + \alpha_2 \triangle IPperVoter_{ist} + X_{is}\beta + \lambda_s + (\delta_i + \triangle \varepsilon_{ist}) \tag{1}$$

Where δ_i is the district trend, λ_s is the state trend, and ε_{ist} is an idiosyncratic error term. With two periods of data, it is not possible to estimate the legislative district specific trend δ_i. OLS estimation of equation (1) will be biased as long as $\delta_i + \triangle \varepsilon_{ist}$ is correlated with $\triangle IPperVoter_{ist}$, which we would expect if Internet is allocated to areas that are trending for or against the BN for unobservable reasons. As a first pass, I include a vector of legislative district covariates (X_{is}) to control for some factors that might affect δ_i:.

OLS estimates for equation (1) for the 2004-2008 period appear in Table II. The first column, reporting estimates of equation (1) with fixed effects and state trends, indicates a strong negative association between change in BN share and $IPperVoter$ growth. As argued above, ethnicity is a central driver in Malaysian politics, with non-Malays more likely to switch allegiances from 2004 to 2008. Since the Chinese population is wealthier and more urban, it could be that $IPperVoter$ is simply picking up this trend. In column (2) I control for this possibility by adding ethnicity, and although the magnitude of the effect diminishes, it remains strongly significant. In line with

[18]The web appendix provides a full account of the methodology employed and the likely sources of measurement error.

anecdotal evidence, Indians and Chinese swung heavily against the BN relative to Malays.[19]

As the Internet backbone runs along Malaysia's major roads, a major concern is that $IPperVoter$ is proxying for the effect of proximity to roads. Specification (3) includes controls both for road density and distance from the centroid of a district to the closest major road. The magnitude of the effect drops slightly but remains significant.

Another concern is that Internet access simply captures wealth or urbanization; the opposition party PKR, for example, derives much of its support from wealthier Malays in large cities. In column (4) I add a measure of GDP per capita as of 2005 as well as population density. Again the magnitude drops, but the relationship remains very significant. Meanwhile, there is no evidence of any relationship between population density and voting trends after controlling for district fixed effects and state trends.

Specification (5) includes additional controls for urbanization using data from satellite imagery. With clutter data on land usage, I create additional controls for the percent of the district that is urban vs. rural vs. jungle. Last, following Burchfield *et al.* (2006), I control for the effect of physical topography on urbanization, constructing a variable for the standard deviation of the land gradient.[20] As can be seen, the estimated relationship between change in BN share and growth in IP addresses per eligible voter remains unchanged. To give a sense of magnitudes, specification (4) implies that a one standard deviation increase in $IPperVoter$ growth translates to a 1% swing against the BN.

As a further check, I run equation (1) for the 1995-1999 period, when Internet connectivity grew from zero to 15%.[21] As mentioned in section 3.2, the Internet was seen to play a decisive role as early as the December 1999 elections. Significantly, demographic composition of the electoral swing differed in the 1999 election. In 2008, Chinese and Indian voters abandoned the BN in favor of the opposition; whereas, in 1999 minority voters stayed with the BN and the Malay electorate instead split. I create a measure of Internet growth from 1995 to 1999, $InternetHH$, which is the natural log of the percentage of households with an Internet subscription in 2000 (Internet penetration was zero in 1995). I drop six observations corresponding to candidates who won as members of an opposition party, S46, in 1995 but switched to the BN in 1999.

[19] The coefficient on percent Chinese is also negative and significant, when percent Malay is the omitted variable.

[20] All regressions have also been run with average land gradient, ruggedness, and the standard deviation of ruggedness with no significant differences. See web appendix for details of variable construction.

[21] I cannot run a regression for the 1995-2008 period because of redistricting between 1999 and 2004 and because my measures of Internet penetration are different.

Table III provides the results. Column (1) shows that, in the absence of controls, a significant positive relationship exists between Internet growth and change in vote share between 1995 and 1999. I interpret this as picking up the fact that the bulk of the swing occurred among Malay voters rather than the relatively more connected Chinese. Indeed, adding ethnicity controls in specification (2) renders the relationship strongly negative and significant. In specifications (3), (4), and (5) I control for GDP per capita, population density, and geography, and the effect is largely unchanged.

The results from the 1995-1999 period reinforce the initial finding of a negative relationship between Internet growth and BN share of the vote. That this result holds for a completely different measure of Internet growth suggests that the result is not merely artifact of the $IPperVoter$ measure. Moreover, since the relationship holds in the presence of a different demographic shift in the electorate, there is less reason to believe that unobserved state trends are driving the result. Notably, the relationship is larger in magnitude: a one standard deviation increase in percentage of households with an Internet subscription implies a 2% swing against the BN in 1999 (the total swing against the BN in 1999 was 11%). This most likely arises because $InternetHH$ is measured with less error than $IPperVoter$.

Placebo regressions

As an additional check, I test whether Internet growth between 2004 and 2008 is higher in areas that were already predisposed to swing against the BN for unobservable reasons. Running OLS on equation (1), I use change in BN share for earlier elections as the dependent variable, but keep $IPperVoter$ 2004-2008 as the independent variable and use the same set of controls. As explained in section above, this analysis is possible only for two previous elections, 1986-1990 and 1995-1999, with the same limitations to the controls.

The 1986-1990 period is a good test case of whether places that experienced more Internet growth between 2004 and 2008 were already more predisposed to swing against the BN. The year 1990 saw an abortive move toward a multi-party system in Malaysia with the BN suffering its worst setback since 1969.[22] It won only 53% of the vote, but managed to retain its two-thirds majority in parliament thanks to gerrymandering.

Panel A of Table IV shows the results of regressing change in BN share from 1986 to 1990 on

[22] Two years earlier, divisions in the UMNO, the dominant Malay party within the BN, caused a formal split in the party with a large number of UMNO politicians leaving to form the opposition Malay party *Semangat 46*.

Internet growth from 2004 to 2008. As indicated, the coefficient on $IPperVoter$ proves small and insignificant regardless of the specification. This suggests no correlation between support for the BN in the 1986-1990 period and Internet growth in the 2004-2008 period.

Next, I run the equivalent regression for the 1995-1999 period, regressing BN share 1995-1999 on $IPperVoter$ 2004-2008. Panel B shows that the negative relationship between Internet growth and BN share does not hold if 2004-2008 measures are used instead. No sign of a relationship between the 1995-1999 election swing and 2004-2008 Internet growth appears, regardless of controls. This suggests that the areas with the greatest swing in 1995-1999 differ from areas experiencing the greatest relative growth in Internet access in 2004-2008.

5.2. *Identification Strategy*

Although the OLS estimates demonstrate a negative relationship between Internet growth and change in BN share, it remains unclear if the relationship is causal. OLS estimation of equation (1) will not identify the causal effects of Internet growth if $\delta_i + \triangle\varepsilon_{ist}$ is correlated with $\triangle IPperVoter_{ist}$. If Internet connectivity is allocated more heavily to districts that are trending toward the BN for unobservable reasons (e.g., patronage) then $\hat{\alpha}_{2,OLS}$ would be biased upward toward zero. If anything, however, this would lead me to underestimate the negative relationship. A greater concern is that Internet connectivity was allocated to areas that trended against the BN for unobserved reasons, leading to a negative bias in the results.

To deal with these challenges, I use the distances from the centroid of a district to Malaysia's three largest ISP backbones as instruments (Z_{ij}) that are correlated with growth in Internet penetration, but uncorrelated with district level characteristics that influence voting behavior. As argued in section 3.3, cost, which is a major determinant of Internet placement, increases in the distance to the backbone. Since the backbones were being built in the 1995-1999 period, the instruments apply only to the 2004-2008 elections. This produces the following system of equations:

$$\triangle y_{ist} = \alpha_1 + \alpha_2 \triangle IPperVoter_{ist} + X_{is}\beta_2 + \lambda_s + (\delta_i + \triangle\varepsilon_{ist}) \tag{2}$$

$$\triangle IPperVoter_{ist} = \pi_0 + Z_{ij}\pi_1 + X_{is}\pi_2 + \gamma_s + \tau_{ist} \tag{3}$$

The identification assumption is that, conditional on baseline district characteristics—ethnic distribution, GDP per capita, population density—distance to the backbone does not affect change in vote share independently of growth in Internet access. So long as the instruments are also uncorrelated with the bias in $IPperVoter$ toward large cities, they will produce consistent estimates even though $IPperVoter$ is measured with error.

An endogeneity concern is that the backbones for Malaysia's ISPs run through areas more likely to swing against the BN for reasons that the controls do not capture. Since the backbones pass through Malaysia's most populous regions and cities, the instruments could simply be picking up the direct effect of urbanization on voting trends. I supplement my controls for population density (log of eligible voters, log of total area) with variables based on satellite data. As in the OLS specifications, I include controls for geography.

Another concern with the instruments is that they are picking up the direct effect of Malaysia's major roads and railways on district trends (e.g., via increased trade and exposure to outside information). To help address this concern, I include controls for distance to nearest road and road density.

In sum, I am exploiting exogenous variation in Internet supply due to geographical constraints in backbone placement. State fixed effects are included to control for differences across states in terms of Internet connectivity and sociopolitical factors. Thus I am exploiting within state variation.[23]

5.3. *First Stage*

Table V shows the first-stage estimates for Internet penetration growth in state legislative districts, using growth in IP addresses per eligible voter as a proxy for growth in Internet access. Column (1) shows estimates of equation (3) with minimal controls for ethnicity. The coefficient on distance to Time, which is highly significant and in the expected direction, suggests that growth in Internet access decreases with distance from Time's backbone.

For Fiberail, both a linear and square term are included. This is meant to capture a non-linear relationship with $IPperVoter$ growth due to restrictions on Fiberail's geographic area of operation until 2006, as mentioned in section 3.3.[24] The negative coefficient on the linear term can be interpreted in the same way as the coefficient for distance to Time: Internet growth decreases as distance increases. The positive square term captures the geographical limitation effect: the rela-

[23]There are 11 states in the sample and 38.8 legislative districts per state.

[24]The linear term by itself is insignificant. There is no evidence of a non-linear relationship for any of the other instruments

tionship between Internet growth and distance to Fiberail becomes less negative until it reaches a zero threshold.

Distance to TM's backbone remains insignificant regardless of the specification. In fact, even if I run the same set of regressions only including IP addresses assigned to TM, the results are largely the same. There are several explanations for this outcome. First, as mentioned above, the location of TM's backbone is measured with error due to incomplete documentation. Second, this result is consistent with the idea that the government exerted more influence over TM than its competitors, compelling it to build out infrastructure in areas with low demand.

Specification (2), controls for distance to roads and road density and the results are unchanged, helping to mitigate the worry that the instruments are simply picking up the direct effect of roads on elections. Column (3) controls for GDP per capita. The size of the coefficients for the instruments decreases yet remains highly significant. Column (4) shows that including controls for population size and density does not noticeably alter the result.

Column (5), which includes geo-spatial controls for urbanization, is the preferred baseline specification. The coefficients of interest decrease slightly in magnitude but maintain their significance. Internet growth demonstrates a negative relationship with population density, but a positive association with percentage of the district that is urban. The most likely reason for this result is catch-up: ISPs had already brought Internet service to the most densely populated areas by 2004 and thus had the most room to grow in regions that were urban but more sparsely populated. To give some interpretation of the magnitudes here, for every 10 kilometer increase in distance to Time's backbone, $IPperVoter$ growth decreases by 0.18 of a standard deviation.

Finally, specification (6) controls for BN share in 2004. As shown, I find no evidence of political interference on Internet roll-out. Several other results suggest that demand, rather than patronage, was the primary determinant of Internet growth. First, there is a strong positive relationship between GDP and Internet growth regardless of the specification. Second, there is no significant correlation between ethnicity and Internet growth, suggesting a lack of large-scale favoritism to specific ethnicities in ICT investment.

5.4. *Instrumental Variable Results*

The IV estimates appear in Table VI. The specifications for (1)-(5) match their first-stage counterparts from Table V. The coefficient on $IPperVoter$ is negative, significant, and of roughly the same magnitude throughout. The Hansen test does not reject the null hypothesis that the instruments are uncorrelated with the error term, lending credence to the identification assumption. The strong and stable coefficients on ethnicity confirm the importance of race in the 2004-2008 elections.

The effect's magnitude drops in column (2), suggesting that $IPperVoter$ in (1) was picking up some of the effect of proximity to major roads. The result remains large and significant, however, and in (5) the coefficient of interest returns to its previous size once controls for urbanization are also included. GDP per capita loses much of its significance, suggesting that it was proxying for urbanization.

To get a sense of the change in magnitudes, for specification (5) a standard deviation increase in Internet growth translates to a 3.6% swing against the BN. Putting this shift into context, IP addresses per voter doubled in the 2004-2008 period, while share of the vote for the BN dropped from 63.9% to 52.2%. This implies that Internet growth accounted for about a third of the vote swing.

The magnitude is substantially larger than the OLS estimate, which is likely due to measurement error biasing the OLS estimates toward zero.

5.5. *Validity of the Exclusion Restriction*

As a reminder, the identification assumption is that, conditional on baseline district characteristics (ethnic distribution, GDP per capita, population density, road density, percent urban vs. rural vs. jungle), distance to the backbone does not affect change in vote share independently of growth in Internet access. It is impossible to test this assumption directly. However, I perform some additional checks to assess its plausibility.

Pre-Internet Trend Tests

The most basic concern is that unobservable characteristics of areas close to the backbone make those areas more prone to swing against the incumbent party in general. I check for this possibility by examining the reduced form relationship between distance to the backbone and swings in previous elections.

Since Malaysia regularly redraws electoral district boundaries, it is not possible to run this exercise for the complete set of preceding elections. Fortunately, the 1969-2008 period has only two other elections—1986-1990 and 1995-1999—in which there was a sizable swing against the BN, and on both occasions boundaries were fixed. I control for ethnicity and population density using the 1991 and 2000 censuses. I also include controls for population density, road density, distance to roads, land usage based on 2008 estimates, and the standard deviation of the slope. A large expansion in state roads occurred during this time, which introduces error into the road density control. Finally, I control for GDP per capita using a 2005 estimate for 2004-2008, 1996-1999, and 1986-1990. Results do not change if the controls measured with error are dropped.

Table VII shows the results of reduced-form regressions for the 1986-1990, 1995-1999, and 2004-2008 periods. Columns (1), (2), and (3) show a negative and insignificant relationship between vote swing and distance to either backbone in the 1986-1990 period. Columns (4), (5), and (6) the relationship with to distance to backbone shifts to positive but remains insignificant. In terms of the 2004-2008 period, column (7) indicates a positive relationship with distance to Time that is significant at the 5% level, and column (8) shows that the linear and square distance to Fiberail variables are jointly significant at the 1% level.

Table VIII performs the same exercise for the complete IV regression. Again the results are insignificant with the exception of the 2004-2008 period.

Controlling for alternate channels

Since Time, TM, and Fiberail run along Malaysia's roads and railroads, another major concern is that some characteristic particular to the distance to roads and railways (or an omitted variable driving it) affected voting trends through some channel, which switched on only after the 2004 elections. I believe this possibility is unlikely for several reasons.

First, the backbones travel along Malaysia's federal roads, most of which were built before 1980. Thus, the effects would have had to remain dormant for more than twenty years.

Second, I have already controlled for distance to roads in the baseline specifications and there is a minimal effect on the coefficient of interest. As a further exercise, I limit instruments to those that travel along major roads. Also, since these backbones travel along a subset of Malaysia's federal roads in particular, I rerun the regressions looking at distance to federal roads. Regardless of the combination of IVs employed, the coefficient on $IPperVoter$ is significant and of constant magnitude.

Finally, I restrict my set of instruments to those based solely on the railway network. In specifications (4), (5), and (6) we see that the coefficient on *IPperVoter* remains highly significant regardless of the control. Comfortingly, the magnitude stays largely the same as in the case using only road-based instruments.

Since Fiberail travels the length of Malaysia's railroads, it is impossible to include equivalent controls for distance to the railway. However, it is worth noting that the railroad network was completed as early as 1931. Thus, to invalidate the instrument, the effect of proximity to railroads would have to have remained dormant for 75 years and then activate just in time to influence the 2008 elections.

Other issues

Another concern is the possibility of heterogeneous effects of Internet access on voting. If the effect of Internet access on voting is more highly negative for areas closer to the backbone, my identification strategy would lead to an overestimation of the effect. An example of this scenario is if areas closer to the Time backbone are better able to exploit Internet technology through better education. Were that the case, however, we would expect to see a markedly different coefficient on *IPperVoter* when only distance to Fiberail is used as an instrument. This is because many of the districts near to Time's backbone are far from Fiberail's backbone. However, the coefficient on *IPperVoter* is largely the same across specifications regardless of the combination of instruments used.

I have run regressions controlling for change in ethnic distribution, eligible voters, and population density between 2004 and 2008. The results are unchanged suggesting that migration is not driving the effect.

6. ADDITIONAL RESULTS

6.1. *Turnover*

A secondary prediction of the theoretical framework is that higher Internet penetration will yield higher turnover. I test this prediction by looking at turnover in seats defended by the BN during the 2008 elections when Internet penetration was greater than 50% and in the 1999 elections when Internet penetration was below 20%. In contrast to previous specifications the analysis is at the cross-sectional level. I run probit regressions of a BN victory dummy on the level of Internet penetration while limiting the sample to districts won by the BN in the previous election. Table IX reports the results.

First, I examine the Internet's effect on turnover in the 1999 election. Since Internet penetration across the country as a whole had reached only 15%, it isn't clear whether there will be a large enough effect to be noticeable. Turning to the data, I find no turnover in BN-defended seats in the states of Johor and Negeri Sembilan. Since my empirical strategy exploits within state variation, I drop the 68 observations corresponding to these two states. I also drop 7 observations corresponding to the state of Kelantan, where all BN-defended seats fell to the opposition. Specification (1) reports the result of a probit regression for 1999. The effect, positive and insignificant, provides no evidence that low levels of Internet penetration substantially affect turnover.

In 2008 Internet penetration for Malaysia as a whole had surpassed 50%. Column (2) shows that in the presence of higher Internet penetration there is a marginally significant effect on turnover. The BN had less chance of retaining a seat in districts with higher Internet penetration. This specification includes the full set of baseline controls plus distance to federal roads. Logit and linear probability specifications yield commensurate results.

To address endogeneity concerns, in columns (3)-(6), I instrument for $IPperVoter$ 2008 using distance to the backbone. Additionally, I run an IV regression for the equivalent linear probability model and include the F-statistics from the first stage. As can be seen, the F-statistics are small. The most likely explanation for the lower significance is that $\Delta IPperVoter$ is a better measure of Internet change than $IPperVoter$ is of Internet levels for reasons specified in section 4.3.

Given the weak first stage, the IV results are less robust than above, thus the evidence for an effect on turnover is weaker. The coefficients on $IPperVoter$ 2008 are relatively stable across specifications (3)-(6), but much larger in magnitude than the simple probit case. In column (3), I include all instruments and the effect proves significant at the 10% level. Turning to specification (4), I drop my weakest instrument, distance to TM, and the significance jumps to the 5% level. In column (5), I restrict the instruments to distance to Time and again the coefficient on $IPperVoter$ is significant at the 5% level. Finally, in column (6), I use only distance to Fiberail and distance to Fiberail squared. Although the point estimates remain similar, the standard errors are much higher, leading to insignificant results.

Although there is some evidence of an effect of Internet diffusion on turnover, it is based off of weak instruments and only holds for some specifications. Unlike in the case of vote share, there is only limited evidence of an effect of Internet diffusion on turnover.

6.2. *Turnout*

Although turnout is not modeled, there are both theoretical and empirical reasons to believe that access to better information on politician quality yields increased turnout (e.g., Banerjee *et al.* (2010)). I look at the effect of Internet diffusion on turnout, focusing on the 2004-2008 and the 1995-1999 periods.

Table X presents the results. In specification (1), I run equation (1) using change in turnout as the dependent variable. The relationship between Internet growth and turnout is positive but insignificant and the magnitude is low. Column (2) uses all three IVs; the effect is negative and insignificant. Finally column (3) shows results for identical OLS regression run for the 1995-1999 elections where results are again insignificant.

Turnout measures in Malaysia are noisy due to electoral irregularities. Allegations of electoral manipulation range from phantom voters (in which deceased individuals still manage to cast ballots) to vote-buying to multiple votes by the same individual.[25] To address this challenge, I ran an extra set of regressions that drop districts with serious irregularities.[26] For the 2004-2008 elections, 13 out of 427 districts are dropped but a lack of information on specific examples of irregularities in earlier elections makes it impossible for me to do the same for the 1995-1999 period.[27] The results, which have been omitted for reasons of space are not noticeably different.

In sum there does not appear to be any meaningful effect of Internet exposure on turnout.

7. INTERPRETATION

To provide some intuition on the magnitude of the effects, I predict the outcome of the 2008 election had there not been any Internet growth in the 2004-2008 period. This exercise suggests that the BN share would have risen from 50.2% to 56.8%, implying that Internet growth accounts for 6.6 points of the 14% swing away from the BN in peninsular Malaysian state electoral districts. Translated into state legislature level results, out of the four state houses that swung to the opposition in 2008, three of them would have remained under the BN. Thus without Internet growth

[25] See Pepinsky (2007) and Hai (2002) for details.
[26] See web appendix for details of irregularities.
[27] All earlier results are robust to dropping these observations.

between 2004 and 2008, the BN's 2008 election setback would have proven fairly modest. To produce this result, 11.7% of the 56% of the population with Internet access would have to have swung as a result of Internet exposure.[28]

In order to compare the effects of Internet with the effects of other media sources documented in the literature, I calculate the persuasion rate. The persuasion rate is defined as the fraction of the Internet audience convinced to change their behavior as a result of exposure to the Internet. The approach differs from DellaVigna & Kaplan (2007) in several key ways: first, like in Enikolopov *et al.* (2011), the approach focuses on the case of continuous exposure by analyzing the effect of an infinitesimal change in Internet exposure. Second, I look at the effect of exposure to Internet-based media in general rather than the effect of one particular media outlet. As such it is not possible to disentangle the effects of individual new media outlets: the effect of pro-reform sites will be mixed up with that of pro-BN sites. Third, as there is very little evidence of an effect of Internet exposure on turnout, change in turnout as a function of change in Internet exposure is not incorporated in the formula for the persuasion rate. Lastly, due to data limitations, I assume that the percentage of voters who look to the Internet for political information is the same as the percentage of voters with Internet access. This is likely an unrealistic assumption. For example, according to Pew, as of 2008 in the United States only 24% of the population got its news on the presidential elections regularly from the Internet.[29] As a result, I also report estimates assuming a smaller percentage of voters get political information from the Internet.

I assume that the positive effect applies only to those who would have voted for the BN in the absence of Internet exposure. Following Enikolopov *et al.* (2011), the formula for the positive persuasion rate is:

$$f = \frac{1}{1 - v_0 t_0} \left(t \frac{dv}{de} + v \frac{dt}{de} \right) \tag{4}$$

Where v_0 is the predicted vote share of the opposition without Internet; t_0 is the turnout without Internet; and t is turnout. $\frac{dv}{de}$ is constructed as follows: turning back to equation 2, with $\triangle y_{ist}$ as the change in share of the opposition from 2004-2008, $\frac{dv}{de}$ is the product of the coefficient on $\triangle IPperVoter_{ist}$ (.042), and the inverse of the share of voters with Internet access (.56). Since the

[28] See Table E.2 in the web appendix for a breakdown of results.

[29] See http://www.people-press.org/2008/01/11/internets-broader-role-in-campaign-2008/.

change in turnout associated with Internet exposure is negligible, t_0 is simply turnout in 2008, .73, and $\frac{dt}{de}$ is 0. Using equation 2, v_0 is the predicted opposition vote share with no Internet growth: .41. Evaluating f at $t = t_0$ and $v = v_0$ we get the positive persuasion rate equal to 7.8%: $(1/(1 - .41 * .73) * (.73 * (.042/.56) + .41 * 0) = 7.8\%$. This is comparable to rates found elsewhere in the literature: 8.23% in DellaVigna & Kaplan (2007) and 7.7% in Enikolopov *et al.* (2011). If we lower the percentage of Malaysians who use the Internet for political information to the US rate of 24%, the positive persuasion rate climbs to 18.3%. If we set it at half the US level of 12%, the positive persuasion rate becomes 36.6%.

8. CONCLUSION

This paper contributes to our understanding of the effect of Internet diffusion on democratization. Focusing on the context in which the traditional media is government-controlled, I have argued that the Internet can facilitate evolution toward a multiparty system by preventing any single agent from monopolizing information. Malaysia provides a key opportunity to test this idea: ambitious investment in an Internet free of censorship coincided with strict controls on all other forms of media.

This paper's central contribution is to quantify the effects of the Internet on democratic change, in the context of the huge growth in Internet penetration that has accompanied Malaysia's recent electoral upheavals. I find that Internet growth accounts for 6.6 points out of the 14 point swing against the BN in the 2008 state elections.

To put this number in perspective, I predict the outcome of the 2008 elections had there not been any Internet growth during the 2004-2008 period. IV estimates imply the BN would have retained control of three of the four statehouses that switched to the opposition. Thus the BN's ICT-based development strategy had the unintended consequence of weakening its control.

I go on to test a secondary prediction. I find weak evidence that Internet exposure can increase turnover, and no evidence of an effect on turnout.

Another important contribution is a novel measure of Internet growth from 2004 to the present. Such a metric is lacking for most countries in the world, including the U.S.A. My measure of Internet connectivity uses IP geo-location data in conjunction with regional Internet registry records. I smooth the IP address point data into a surface using inverse distance weighting interpolation and then normalize by population. Finally I check the accuracy against an independent measure of Internet diffusion from household census data. This measure is central to the paper's results as it allows me to track Internet growth at the state legislature district level. This measure can also extend to research well outside the ambit of this paper. Equivalent IP geo-location data exists for almost every country in the world and is only becoming more accurate as the technology matures.

This paper presents some of the first evidence of the Internet's quantitative effects on political outcomes. However, there is much scope for future work. First, it is important to get a better understanding of the channels of causation. Anecdotal evidence suggests that the Internet influenced elections via the media market: there was a drop in the popularity of BN-owned newspapers and even a decrease in bias among some media outlets as a means to reestablish credibility. Finally, it would be fruitful to explore the Internet's consequences in terms of economic development. Malaysia invested heavily in Internet infrastructure to promote an information economy. An important next step would be to gauge whether this investment paid off, as it would imply a relationship between political openness and economic growth.

REFERENCES

Adena, Maja, Enikolopov, Ruben, Petrova, Maria, Santarosa, Veronica, & Zhuravskaya, Ekaterina. 2015. Radio and the Rise of the Nazis in Prewar Germany. *Working Paper*.

Banerjee, AV A.V., Kumar, Selvan, Pande, Rohini, & Su, Felix. 2010. *Do Informed Voters Make Better Choices? Experimental Evidence from Urban India*.

Besley, Timothy, & Burgess, Robin. 2002. The Political Economy of Government Responsiveness: Theory and Evidence from India. *The Quarterly Journal of Economics*, **117**(4), 1415–1451.

Besley, Timothy, & Prat, Andrea. 2006. Handcuffs for the Grabbing Hand? Media Capture and Government Accountability. *American Economic Review*, **96**(3), 720–736.

Burchfield, Marcy, Overman, H.G. G., Puga, D., & Turner, M.A. a. 2006. Causes of Sprawl: A Portrait from Space. *The Quarterly Journal of Economics*, **121**(2), 587–633.

Campante, FR, Durante, R, & Sobbrio, F. 2013. Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation. *NBER Working Paper W19029*.

Case, William. 2001. Malaysia's Resilient Pseudodemocracy. *Journal of Democracy*, **12**(1), 43–57.

Centre for Independent Journalism Malaysia. 2008. Report on the Quantitative Analysis of the Media Monitoring Initiative for the 12th General Elections. 1–41.

Chiang, CF, & Knight, B. 2011. Media Bias and Influence: Evidence from Newspaper Endorsements. *The Review of Economic Studies*, **78**(3), 795–820.

Czernich, Nina. 2012. Broadband Internet and Political Participation: Evidence for Germany. *Kyklos*, **65**, 31–52.

Della Vigna, Stefano, Enikolopov, Ruben, Mironova, Vera, Petrova, Maria, & Zhuravskaya, Ekaterina. 2014. Cross-border media and nationalism: Evidence from serbian radio in croatia. *American Economic Journal: Applied Economics*, **6**(3), 103–132.

DellaVigna, Stefano, & Kaplan, Ethan. 2007. The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics*, **122**(3), 1187–1234.

Drago, Francesco, Nannicini, Tommaso, & Sobbrio, Francesco. 2014. Meet the press: How voters and politicians respond to newspaper entry and exit. *American Economic Journal: Applied Economics*, **6**(3), 159–188.

Durante, Ruben, & Knight, Brian. 2012. Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi's Italy. *Journal of the European Economic Association*, **10**(3), 451–481.

Durante, Ruben, Pinotti, Paolo, & Tesei, Andrea. 2014. No News, Big News. The Political Consequences of Entertainment TV. *Working Paper*.

Enikolopov, Ruben, Petrova, Maria, & Zhuravskaya, Ekaterina. 2011. Media and Political Persuasion: Evidence from Russia. *American Economic Review*, **101**(7), 3253–3285.

Enikolopov, Ruben, Petrova, Maria, & Sonin, Konstantin. 2012. Do Political Blogs Matter? Corruption in State-Controlled Companies, Blog Postings, and DDoS Attacks. *SSRN Electronic Journal*, 25–27.

Falck, Oliver, Gold, Robert, & Heblich, Stephan. 2014. E-Lections: Voting behavior and the internet. *American Economic Review*.

Gentzkow, Matthew. 2006. Television and Voter Turnout. *Quarterly Journal of Economics*, **121**(3), 931–972.

Gentzkow, Matthew, & Shapiro, Jesse M. J.M. 2011. Ideological Segregation Online and Offline. *The Quarterly Journal of Economics*, **127**(4), 1–49.

Gentzkow, Matthew, Shapiro, Jesse M, & Sinkinson, Michael. 2011. The Effect of Newspaper Entry and Exit on Electoral Politics. *American Economic Review*, **101**(7), 2980–3018.

Gerber, Alan S, Karlan, Dean, & Bergan, Daniel. 2009. Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics*, **1**(2), 35–52.

Goyal, Aparajita. 2010. Information, Direct Access to Farmers, and Rural Market Performance in Central India. *American Economic Journal: Applied Economics*, **2**(3), 22–45.

Hai, L.H. H. 2002. Electoral Politics in Malaysia: 'managing' Elections in a Plural Society. *Pages 101–148 of:* Croissant, A., Bruns, G., & John, M. (eds), *Electoral Politics in Southeast and East Asia*. Singapore: Friedrich Ebert Stiftung.

Jaber, Ahmed. 2013. Broadband Internet and Political Behavior: Evidence from the United States. *SSRN Electronic Journal*.

Jack, William, & Suri, Tavneet. 2014. Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution. *American Economic Review*, **104**(1), 183–223.

Jensen, Robert. 2007. The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics*, **122**(3), 879–924.

Malaysian Communications and Multimedia Commission. 2013. Communications and Multimedia Pocket Book of Statistics. **2**, 12.

Pepinsky, Thomas B. 2007. Malaysia: Turnover Without Change. *Journal of Democracy*, **18**(1), 113–127.

Snyder, James M., & Strömberg, David. 2010. Press Coverage and Political Accountability. *Journal of Political Economy*, **118**(2), 355–408.

Stromberg, D. 2004. Radio's Impact on Public Spending. *The Quarterly Journal of Economics*, **119**(1), 189–221.

Zinnbauer, Dieter. 2003. *Power and Activism in the Context of a Maturing Internet : The Case of Malaysia*. Ph.D. thesis, London School of Economics and Political Science.

9. TABLES

TABLE I
Summary Statistics

Variable	Mean	Std. Dev.	*N*
$\Delta BNShare$ 2004-2008	-.122	.093	427
$\Delta Turnout$ 2004-2008	.019	.036	427
IPperVoter Growth	1.94	.867	427
% Internet 2000	.167	.163	427
% Malay 2004	.632	.276	427
% Chinese 2004	.280	.239	427
% Indian 2004	.077	.078	427
% Internet 2004	.170	.176	427
GDP per capita 2005	16787.82	7182.87	427
Eligible Voters 2004	17720.98	7215.80	427
Population Density	810.12	1418.96	427
% Urban	.220	.241	427
% Rural	.502	.250	427
Slope Std. Dev.	4.00	2.95	427
Road Density	.625	.632	427
Km to Federal Road	3.48	4.68	427
Km to Major Road	1.34	2.13	427
Km to Time	15.22	18.32	427
Km to Fiberail	22.22	28.32	427
Km to TM	7.06	7.75	427

Notes. The table reports summary statistics for state legislature districts in peninsular Malaysia, excluding Kuala Lumpur and regions with negative IpPerVoter growth. Variables measured in 2008 unless otherwise stated. See web appendix for details on the construction and sources of variables.

TABLE II
Relationship between BN share and Internet growth from 2004 to 2008

	Dependent variable is $\Delta BNShare$ 2004-2008				
	(1)	(2)	(3)	(4)	(5)
IPperVoter Growth	-0.019***	-0.013***	-0.011***	-0.009**	-0.009***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
% Malay		0.164***	0.148***	0.129***	0.121***
		(0.016)	(0.017)	(0.019)	(0.019)
% Indian		-0.348***	-0.368***	-0.384***	-0.404***
		(0.053)	(0.053)	(0.054)	(0.054)
Km to Road			0.017**	0.017**	0.020**
			(0.009)	(0.008)	(0.008)
Road Density			-0.012*	-0.003	-0.005
			(0.006)	(0.008)	(0.010)
GDP per capita				-0.033***	-0.031***
				(0.010)	(0.010)
log Eligible Voters				0.006	0.003
				(0.012)	(0.011)
Population Density				-0.001	-0.001
				(0.003)	(0.003)
% Urban					-0.019
					(0.026)
% Rural					-0.003
					(0.016)
Slope std					-0.003**
					(0.001)
N	427	427	427	427	427
R^2	.441	.695	.702	.712	.718

Notes. The table reports OLS estimates of equation (1). All specifications include 11 state trends. IPperVoter growth is natural log of IP addresses per eligible voter in 2008 divided by IP addresses per voter in 2004. See web appendix for details on the construction and sources of variables. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels.

TABLE III

Relationship between BN share and Internet growth from 1995 to 1999

	Dependent variable is Δ*BNShare* 1995-1999				
	(1)	(2)	(3)	(4)	(5)
InternetHH 1995-1999	0.021***	-0.015***	-0.017***	-0.020***	-0.020***
	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
% Malay 1999		-0.298***	-0.286***	-0.280***	-0.281***
		(0.022)	(0.022)	(0.024)	(0.024)
% Indian 1999		-0.213***	-0.197***	-0.195***	-0.185***
		(0.065)	(0.062)	(0.063)	(0.063)
Km to Road			0.042**	0.043**	0.041**
			(0.019)	(0.020)	(0.021)
Road Density			0.144**	0.132	0.077
			(0.069)	(0.091)	(0.100)
GDP per capita 2005				0.005	0.004
				(0.012)	(0.012)
log Eligible Voters 1995				0.014	0.014
				(0.017)	(0.017)
Population Density 2008				-0.000	-0.001
				(0.002)	(0.002)
% Urban					0.007
					(0.031)
% Rural					-0.028
					(0.021)
Slope std					-0.003*
					(0.001)
N	368	368	368	368	368
R^2	.269	.572	.583	.584	.59

Notes. The table reports OLS estimates of equation (1). All specifications include 11 state trends. 1999 election is from December 1999. Internet growth is the natural log percentage of households with Internet subscriptions in 2000 (Internet access was zero in 1995). Drops six candidates who ran under opposition in 1995, but switched to the BN in 1999. See web appendix for details on the construction and sources of variables. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels.

TABLE IV

Placebo regressions of Δ*BNShare* on Internet growth in a different time period

	(1)	(2)	(3)	(4)	(5)
PANEL A: Dependent variable is Δ*BNShare* 1986-1990					
IPperVoter 2004-2008	-0.001	-0.001	-0.002	-0.004	-0.003
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
PANEL B: Dependent variable is Δ*BNShare* 1995-1999					
IPperVoter 2004-2008	-0.001	-0.005	-0.004	-0.004	-0.003
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Controls					
Ethnicity	N	Y	Y	Y	Y
Road	N	N	Y	Y	Y
GDP per capita	N	N	N	Y	Y
Population	N	N	N	Y	Y
Geographic	N	N	N	N	Y

Notes. The table reports OLS estimates of equation BN share change on Internet growth in a different period. Panel A reports results of BN share 1986-1990 on IPperVoter Growth from 2004 to 2008. Panel B reports results of BN share 1995-1999 on IPperVoter Growth from 2004 to 2008. IPperVoter 2004-2008 is the natural log of IP addresses per eligible voter in 2008 divided by IP addresses per voter in 2004. All specifications include 11 state trends. Ethnicity controls are % Malay and % Chinese, from each respective period. GDP per capita is taken from 2005 for panels A and B. Population controls for population density, and log of eligible voters. Log of eligible voters is for 1986, 1995, and 2004, respectively, while the other controls are from 2008. Road controls for road density and distance to federal roads as of 2008. Geographic controls for standard deviation of slope, % urban, and % rural. Panel B drops six candidates who ran under opposition in 1995, but switched to the BN in 1999. See web appendix for details on the construction and sources of variables. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels. For expository clarity, coefficients on controls are not reported.

TABLE V
First Stage relationship between distance to backbone and Internet growth

	Growth in IPs per eligible voter 2004-2008: $\triangle IPperVoter$					
	(1)	(2)	(3)	(4)	(5)	(6)
Km to Time*10	-0.115***	-0.117***	-0.097***	-0.096***	-0.085***	-0.085***
	(0.027)	(0.027)	(0.027)	(0.028)	(0.028)	(0.029)
Km to Fiberail*10	-0.194***	-0.194***	-0.180***	-0.177***	-0.157***	-0.158***
	(0.043)	(0.044)	(0.044)	(0.044)	(0.045)	(0.046)
Km to Fiberail*10 SQ	0.021***	0.021***	0.020***	0.019***	0.018***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Km to TM*10	-0.015	-0.026	-0.034	-0.032	-0.016	-0.015
	(0.059)	(0.064)	(0.064)	(0.064)	(0.067)	(0.067)
% Malay 2004	-0.001	0.032	0.265	0.192	0.299	0.310
	(0.196)	(0.216)	(0.234)	(0.240)	(0.239)	(0.258)
% Indian 2004	0.134	0.123	0.313	0.039	0.396	0.426
	(0.648)	(0.667)	(0.666)	(0.679)	(0.694)	(0.732)
Km to Road		0.122	0.110	0.129	0.146	0.147
		(0.235)	(0.221)	(0.220)	(0.226)	(0.227)
Road Density		0.031	-0.052	0.065	-0.126	-0.124
		(0.068)	(0.069)	(0.081)	(0.086)	(0.086)
GDP per Capita			0.410***	0.451***	0.402***	0.404***
			(0.120)	(0.123)	(0.124)	(0.126)
log Eligible Voters 2004				-0.004	-0.037	-0.042
				(0.156)	(0.155)	(0.162)
Population Density				-0.093**	-0.134***	-0.134***
				(0.044)	(0.051)	(0.051)
% Urban					1.002***	0.999***
					(0.355)	(0.355)
% Rural					0.021	0.020
					(0.251)	(0.251)
Slope std					0.000	0.001
					(0.019)	(0.019)
BN Share 2004						-0.066
						(0.493)
N	427	427	427	427	427	427
R^2	.307	.308	.326	.335	.352	.352

Notes. The table presents OLS estimates of equation (3). It presents first stage results for the relationship between distance to backbone and growth in IP addresses per voter. All specifications include 11 state trends. IPperVoter is natural log of IP addresses per eligible voter in 2008 divided by IP addresses per voter in 2004. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels. See web appendix for details on the construction and source variables.

TABLE VI

IV Estimates of the relationship between BN share and Internet growth

	Dependent variable is $\Delta BNShare$ 2004-2008					
	IV					OLS
	(1)	(2)	(3)	(4)	(5)	(6)
IPperVoter Growth	-0.036***	-0.029***	-0.024***	-0.026***	-0.033***	-0.009***
	(0.007)	(0.008)	(0.008)	(0.009)	(0.010)	(0.003)
% Malay 2004	0.156***	0.146***	0.133***	0.133***	0.128***	0.121***
	(0.017)	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)
% Indian 2004	-0.331***	-0.349***	-0.359***	-0.366***	-0.379***	-0.404***
	(0.055)	(0.054)	(0.053)	(0.054)	(0.054)	(0.054)
Km to Road		0.013	0.013	0.014*	0.017*	0.020**
		(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
Road Density		-0.009	-0.004	-0.001	-0.007	-0.005
		(0.007)	(0.007)	(0.008)	(0.010)	(0.010)
GDP per capita			-0.023**	-0.022**	-0.019*	-0.031***
			(0.010)	(0.010)	(0.011)	(0.010)
log Eligible Voters 2004				0.009	0.007	0.003
				(0.012)	(0.012)	(0.011)
Population Density				-0.004	-0.005	-0.001
				(0.003)	(0.003)	(0.003)
% Urban					0.007	-0.019
					(0.029)	(0.026)
% Rural					-0.003	-0.003
					(0.016)	(0.016)
Slope std					-0.003**	-0.003**
					(0.001)	(0.001)
N	427	427	427	427	427	427
R^2	.657	.68	.696	.693	.684	.718
F-Stat	18.1	16.0	13.0	11.71	9.2	
Hansen Test (p-value)	.75	.45	.42	.45	.68	

Notes. Specifications (1) through (5) show results of IV regressions of change in BN vote share 2004-2008 on IPperVoter growth 2004-2008. Instruments are distance to Time, distance to Fiberail, distance to Fiberail squared, and distance to TM. Column (6) reports results from an ordinary least squares regression of BN vote share 2004-2008 on IPperVoter growth 2004-2008. F-stat is the f-statistic of the instruments from the first stage. The *p*-value for the Hansen test is for the Sargan-Hansen test of overidentifying restrictions. The joint null is that the instruments are uncorrelated with the error. All specifications include 11 state trends. IPperVoter growth is the natural log of IP addresses per eligible voter in 2008 divided by IP addresses per voter in 2004. See web appendix for details on the construction and sources of variables. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels.

TABLE VII

Reduced form estimates of distance to backbone on elections

	Dependent variable is $\Delta BNShare$								
	1986-1990			1995-1999			2004-2008		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Km to Time*100	-0.040		-0.039	0.027		0.023	0.036**		0.031*
	(0.026)		(0.026)	(0.022)		(0.023)	(0.017)		(0.018)
Km to Fiberail*100		-0.064	-0.066		0.029	0.024		0.032	0.033
		(0.048)	(0.047)		(0.037)	(0.037)		(0.030)	(0.030)
Km to Fiberail*100 SQ		0.046	0.042		-0.025	-0.018		-0.053**	-0.051**
		(0.035)	(0.036)		(0.031)	(0.031)		(0.023)	(0.022)
Km to TM*100			-0.045			0.039			-0.037
			(0.065)			(0.048)			(0.039)
Road Density	0.219*	0.233**	0.227**	0.089	0.081	0.088	-0.004	-0.003	-0.003
	(0.115)	(0.115)	(0.115)	(0.103)	(0.104)	(0.105)	(0.010)	(0.010)	(0.010)
Km to Major Road*10	0.006	0.003	0.012	0.040**	0.043**	0.035**	0.015	0.017**	0.017*
	(0.015)	(0.016)	(0.017)	(0.018)	(0.019)	(0.018)	(0.010)	(0.009)	(0.010)
Fiberail joint significance		.39	.37		.72	.81		.002	.004
N	325	325	325	368	368	368	427	427	427
R^2	.478	.478	.483	.578	.578	.58	.716	.717	.719

Notes. Reduced form regressions of change in BN share on distance to the backbone are reported. Columns (1), (2), and (3) cover the 1986-1990 elections; columns (4), (5), and (6) cover the 1995-1999 elections; and columns (7), (8), and (9) cover the 2004-2008 elections. Fiberail joint significance presents the p-value of a test of the joint significance of Km to Fiberail and Km to Fiberail squared. Reduced form regressions containing only distance to TM have been ommitted for space but are insignificant in all cases. All specifications control for ethnicity, GDP per capita, percent of the district that is urban and rural, standard deviation of the slope, the log of eligible voters, population density, distance to major roads, road density and 11 state trends. GDP per capita is taken from a 2005 estimate in all cases. In columns (4), (5), and (6), six candidates are dropped who ran under opposition in 1995, but switched to the BN in 1999. For all specifications population density, road density, % urban, and % rural are calculated from a 2008 measure. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels.

TABLE VIII

Placebo IV regressions for earlier elections

	Dependent variable is $\Delta BNShare$								
	2004-2008			1995-1999			1986-1990		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IPperVoter Growth 04-08	-0.035**	-0.033***	-0.033***						
	(0.017)	(0.011)	(0.010)						
IPperVoter Growth 95-99				-0.028	-0.010	-0.015			
				(0.024)	(0.015)	(0.014)			
IPperVoter Growth 86-90							0.045	0.013	0.024
							(0.030)	(0.016)	(0.015)
Road Density	-0.007	-0.007	-0.007	0.039	0.072	0.064	0.260**	0.248**	0.253**
	(0.011)	(0.010)	(0.010)	(0.115)	(0.105)	(0.107)	(0.126)	(0.113)	(0.117)
Km to Major Road*10	0.017*	0.017*	0.017*	0.038*	0.042**	0.040**	0.008	0.004	0.006
	(0.010)	(0.009)	(0.009)	(0.021)	(0.019)	(0.019)	(0.020)	(0.017)	(0.018)
N	427	427	427	368	368	368	325	325	325
R^2	.678	.684	.684	.535	.574	.568	.339	.459	.431
Instruments	Time	Fiberail	ALL	Time	Fiberail	ALL	Time	Fiberail	ALL

Notes. IV regressions of change in BN vote share on IPperVoter growth. Columns (1), (2), and (3) cover the 2004-2008 elections; columns (4), (5), and (6) cover the 1995-1999 elections; and columns (7), (8), and (9) cover the 1986-1990 elections. ALL includes Time, Fiberail and TM. IV regressions containing only distance to TM have been omitted for space but are insignificant in all cases. All specifications control for ethnicity, GDP per capita, percent of the district that is urban and rural, standard deviation of the slope, the log of eligible voters, population density, distance to major roads, road density and 11 state trends. GDP per capita is taken from a 2005 estimate in all cases. In columns (4), (5), and (6), six candidates are dropped who ran under opposition in 1995, but switched to the BN in 1999. For all specifications population density, road density, % urban, and % rural are calculated from a 2008 measure. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels.

TABLE IX

Probit estimates of turnover on Internet

	Probit		IV Probit			
	1999	2008	2008	2008	2008	2008
	(1)	(2)	(3)	(4)	(5)	(6)
IPperVoter 2008		-0.233*	-1.022**	-1.046**	-1.171**	-0.961
		(0.139)	(0.517)	(0.481)	(0.531)	(0.649)
InternetHH 1999	0.235					
	(0.225)					
N	253	379	379	379	379	379
Pseudo R^2	.44	.503				
First stage: Dependent variable is *IPperVoter* 2008						
Km to Time*10			-0.048*	-0.047*	-0.056**	
			(0.027)	(0.026)	(0.028)	
Km to Fiberail*10			0.000	0.001		-0.003
			(0.042)	(0.041)		(0.042)
Km to Fiberail*10 SQ			0.004	0.004		0.005
			(0.004)	(0.004)		(0.004)
Km to TM*10			0.010			
			(0.057)			
F-stat			2.75	3.65	4.42	3.67
N			379	379	379	379

Notes. Probit estimates of turnover on Internet connectivity are reported. Specification (1) regresses turnover from December 1999 on log % households with Internet subscription in 1999, and restricts sample to districts won by the BN in 1995. Specifications (2)-(6) regress turnover 2008 on log IPperVoter 2008, and restrict sample to districts that the BN won in 2004. All specifications control for ethnicity, GDP per capita, percent of the district that is urban and rural, standard deviation of slope, the log of eligible voters, population density, road density, distance to major roads, and 11 state trends. GDP per capita is taken from a 2005 estimate. For all specifications distance to major roads, road density, standard deviation of slope, % urban, and % rural are calculated from a 2008 measure. F-stat is the f-statistic of the instruments from the first stage of 2SLS estimate from the equivalent linear probability model. See web appendix for details on the construction and sources of variables. Coefficients are reported with standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels.

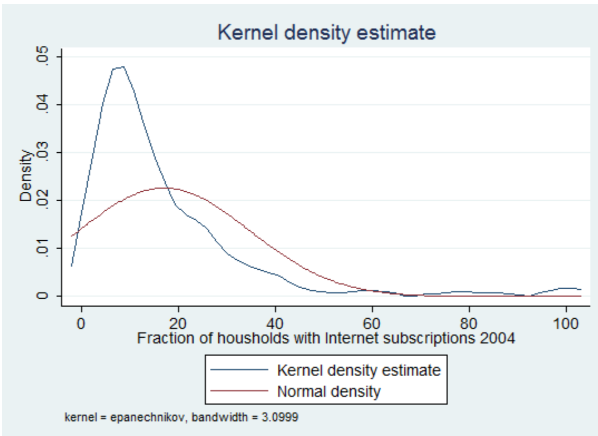
TABLE X

Relationship between turnout and Internet growth

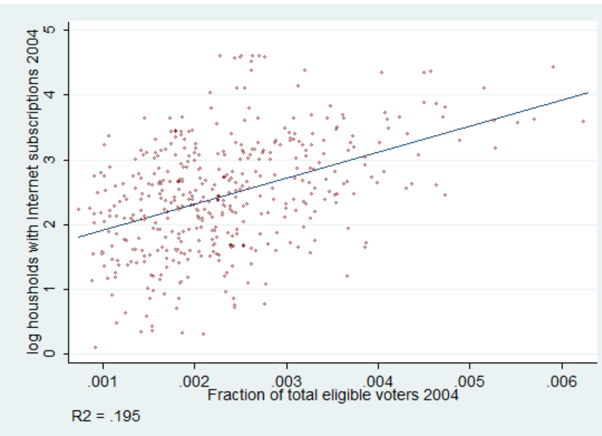
	$\Delta Turnout$ 2004-2008		$\Delta Turnout$ 1995-1999
	OLS	IV	OLS
	(1)	(2)	(3)
IPperVoter growth 04-08	0.005	-0.004	
	(0.003)	(0.005)	
InternetHH 1995-1999			0.002
			(0.003)
N	427	427	368
R^2	.403	.362	.286

Specification (1) shows results of OLS regression of change in turnout 2004-2008 on IPperVoter growth 2004-2008. Column (2) presents results of IV regressions using all three instruments. Specification (3) reports results of regression of change in turnout 1995-1999 on Internet subscription per household growth 1995-1999. InternetHH 1995-1999 is the natural log percentage of households with Internet subscriptions in 2000 (Internet access was zero in 1995). All specifications control for ethnicity, GDP per capita, percent of the district that is urban and rural, standard deviation of slope, the log of eligible voters, population density, road density, distance to major roads, and 11 state trends. For specifications (3) GDP per capita is taken from 2005. Road density, % urban, % rural are from 2008. IPperVoter growth is natural log of IP addresses per eligible voter in 2008 divided by IP addresses per voter in 2004. In specification (3), six candidates are dropped who ran under opposition in 1995, but switched to the BN in 1999. See web appendix for details on the construction and sources of variables. Coefficients are reported with robust standard errors in brackets. ***,**, and * indicate significance at the 1%, 5%, and 10% levels.

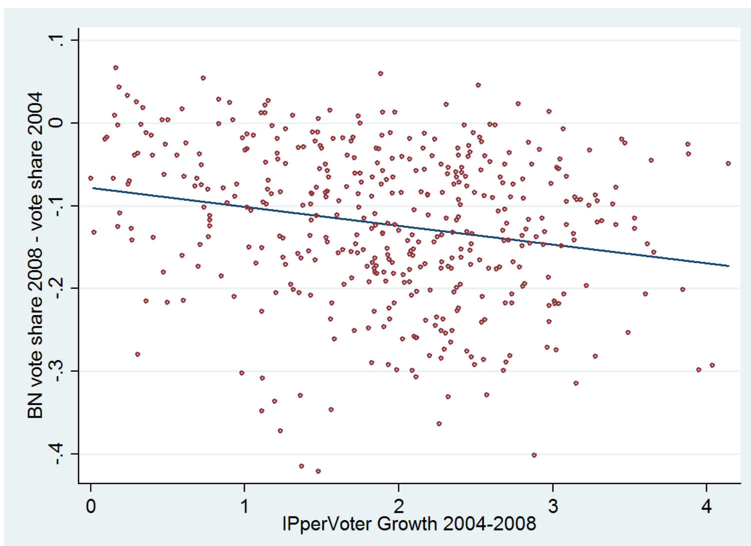
10. FIGURES



(a) Skewness of households with Internet 2004



(b) Internet Penetration vs. Population



(c) Vote share vs. Internet Growth

FIGURE I.— Graphs