Proceedings of The National Conference
On Undergraduate Research (NCUR) 2020
Montana State University, Bozeman MT
March 26-28, 2020

Disruptive Internet: A New Determinate of Education Quality

Denton Knight
International Affairs
United States Military Academy at West Point
West Point, New York 10997 USA

Faculty Advisor: Major Daniel Flynn

Abstract

This research proposes and tests two theories to explain why education quality and test outcomes vary between countries. The first, more traditional view, argues that state capacity and institutional design choices determine outcomes. The second, disruptive theory, argues that state capacity and democratic institutions together determine internet quality. Good internet access could then improve learning through mechanisms such as technology use in school, exposure to English media, and use of educational apps. While the first theory suggests that regime type has little importance or that authoritarian states may even provide better educations, the second mechanism would make give democratic countries a boost in educational performance. I test the link between internet access and education quality by regressing mobile data affordability (cost of 1 GB of data as a percent of GDP Per Capita) on TOEFL (Test Of English as a Foreign Language) score outcomes. The findings provide correlative support for the disruptive theory: democratic institutions are better for establishing an open, fast, and functional internet that then improves learning, at least in terms of English ability.

Keywords: Education Quality, Internet, State Capacity

1. Introduction

Education quality matters. States that possess more effective educational institutions are wealthier, have more intelligent and entrepreneurial populations, and (with some caveats) are more likely to be or become democracies in the future. Yet it is not obvious what determines the quality of a state's educational institutions. One can easily imagine an array of factors that might affect the quality of a state's education system: its economic wealth and prosperity; the government's capacity to plan, execute, and enforce policies; particular decisions regarding the design of the state's school system; or the society's cultural context and attitude towards learning. Out of these, scholars have historically coalesced around two primary determinants of education quality: the state's capacity "to plan and execute policies and enforce laws cleanly and transparently," and the quality of the technocratic design of its educational institutions. Neither of these factors are dependent on whether the state is a democracy. The second point may even favor technocratic autocracies that are better able to make centralized and efficient decisions. Consider more authoritarian East Asian countries' dominance in STEM compared to the United States' internal variations and general dysfunctionality. Compare yet again schools in the US or Europe with those in Sub-Saharan Africa to notice the importance of state capacity.

However, the past two decades have seen the rise of a new technology that is disrupting societies and changing almost everyone's ways of life: the internet. At a micro level, the internet offers new access to online learning resources such as eBooks, massive open online courses (MOOC), and education apps or games such as Khan Academy. Internet can be used in schools to make teaching more effective. Finally, the internet might expose people to English and new ideas through their consumption of entertaining and educational videos, articles, or social media. While extensive research support this evidence on a small scale, a question remains: do these micro-level effects lead to large-scale impacts? Unlike education quality, openness and quality of internet access is dependent on more democratic institutions. This is because state control of media or bans on certain sites directly slow down access and speed. It also

challenges the ethos of the internet: a free exchange of content and ideas. If internet access affects education at a country-wide level, then countries with more inclusive institutions could begin seeing larger improvements in their educational outcomes compared to more authoritarian states. The internet could prove to disrupt the historical factors dictating education quality.

This research incorporates education quality, an outcome of an educational system measured through test scores rather than the more commonly measured inputs of number of years of schooling or funds invested in education. This decision was made because outcomes are much more directly related to the growth of human capital than inputs or 'indicators of education quantity' that may not actually result in better education. Dahlum and Knutsen found that while quantity and quality of education are correlated, a high number of countries deviate from this pattern. Unfortunately, however, international tests such as the PISA test (for math, science, and reading) are not conducted in many developing countries and suffer from several validity issues. For this reason, I choose to sacrifice some generalizability by measuring quality via the limited assessment of TOEFL (Test of English as a Foreign Language) English scores. This test is globally ubiquitous and also relevant to economic development as takers of the TOEFL test tend to be the most educated and business-oriented members of society. Those most likely to take the TOEFL exam are students seeking to establish their mastery of English for application to universities, scholarships, or visas.

In this paper, I first review the literature that suggests that state capacity and institutional design choices determine education quality. Then, I present the theoretical basis for the new disruptive theory: first, states must pass through four junctures that require democratic choices in order to establish internet infrastructure. This is backed by some correlative evidence that connects democratization with internet speed. Micro-level studies have found a linkage between internet/technology and educational outcomes. After reviewing this body of evidence, I present the correlational findings that show a macro-level linkage between internet affordability and English test scores, particularly in reading and listening domains. This finding is not causal nor fully generalizable given that the internet is likely to affect English scores more than other learning outcomes such a science or math. However, it does provide evidence that a connection between internet affordability and education quality is worth exploring further.

1.1. Standard Theory: Education Quality and Institutional Design

Extensive economic research has established the positive impacts of education on wages and human capital development. Furthermore, modernization theory and subsequent analyses argue that education acts as an important prerequisite for democratic transitions. Glaeser argues that this is because education increases socialization and builds up the civil society necessary to sustain a democracy. Of course, some highly-educated states still have autocracies. This is because education may fail to build up needed civil society when educational opportunities are limited to elites. In a semi-authoritarian state, the most educated are often also the most indoctrinated and therefore less likely to agitate for democracy. While literature exploring the political effects of education is quite extensive, less research explores the under-side of this coin: at a macro-level, why are some states better educated than others?

The first explanation for educational differences lies with modernization theory: as a state transitions from a labor-based economy to a service economy it will become more capable of providing better educations. In turn, citizens will begin to demand good education in order to be employed in jobs with newly technical requirements. Increasingly, researchers recognize a third component in this process: health improvements. Together, higher incomes, health improvements, and better education can create virtuous circles of development.[1] Higher incomes ensure families are able to put their children all the way through school without having to work or go absent because they cannot afford fees. Income also helps parents afford the better nutrition and medicines to help their children be healthier. In turn, healthier kids do not miss as much school from being sick.[2] More generally, healthier people are able to work more consistently and end up with a higher income. Education about diseases and healthy living practices can be a cost-effective way to improve health.[3] Finally Education has been shown to contribute to long-term income gains. As seen, each of these factors contributes to the other two; a lack of any one of them can impede the establishment of the rest.

Despite this, no one factor is dependent on the other two. Radelet proposes three inputs that contribute to health and education developments.[4] First, global health has vastly improved over the past few decades through new technologies and vaccines. This has led to a global improvement in health irrespective of incomes.[5] Second, local leadership in developing countries have begun investing in stronger education and health-care systems.[6] These investment decisions are affected by an array of factors, not least of which include a state's cultural, ideological, and political opinion toward education.[7] Finally, international efforts by non-governmental organizations (NGOs) and governments have played an important role in distributing new medicinal technologies and providing resources or pressure for the development of education.

If all this is true, there is no simple explanation for variations in a state's academic performance. The health of the populace, degree of international pressures and number of NGO initiatives, cultural norms towards the value of

education, and a government's spending priorities all affect the quality of an education system. Yet while income alone does not explain it, education quality does appear to improve in parallel with other factors: you cannot have a failed or impoverished state that also provides great educations. In a term, quality of education is contingent upon a state's capacity.[8] However, even a strong state can have a weak academic performance if its institutions are organized poorly. Unlike what one might expect, these are *not* the same institutions that make a state inclusive or extractive.

Extensive literature has established international education production functions using cross-sectional data at both an aggregate and student level.[9] Building off this base, Woessmann developed a production model that finds that family background and the institutional structures of school systems account for most achievement variation at a country level.[10] However, the school's resources and the size of the funding allocated to education by the state only accounts for a small percentage of the variation.[11] Family background represents wider socio-economic factors that fit into the broad modernization argument. Surprisingly, he finds that institutional choices matter at least as much if not more than these background factors. The introduction of external exit exams and competition from private schools increases student achievement. Giving schools greater autonomy works well in developed countries, depending upon where that autonomy is granted, but generally has a negative effect in developing countries.[12] This implies that overall level of economic development and a country's institutional choices matter the most to education quality. The amount of funds a state or system allocates to education (in terms of expenditure per student or class size) has a limited effect on student achievement.[13] These findings support Radelet's argument that local leadership plays an important role in shaping education quality by making intelligent policy and structural decisions. However, these decisions are not reliant upon inclusive institutions – you do not necessarily need to be a democracy to craft better policy.

Several studies have found that democracies provide more education in terms of educational enrollment and years of schooling.[14] However, recent research has found with great robustness that democracies do not provide better *quality* education results globally.[15] Dahlum and Knutsen find that autocracies provide more variable education quality. This fits with the argument that specific policy decisions matter most – autocrats can choose between policies that are good or bad for education while democracies are more consistently mediocre. This may explain why undemocratic East Asian states such as China are far outperforming western liberal democracies academically.

Education quality depends on two factors. First, a state's capacity "to plan and execute policies and enforce laws cleanly and transparently."[16] Those states that are able to do this are generally richer and healthier. Second, it must have expertly designed institutions that facilitate learning outcomes – these can exist in both democratic and autocratic states. All this recognized, there is a new player in town: the internet. How does a state's decision to develop or permit the development of new internet infrastructure affect education performance? Does internet technology on its own aid with learning (like medicinal technologies) or is it dependent upon the institutional structure of how it is employed?

1.2. Disruptive Theory Part One: Democratic Decisions in developing Internet Infrastructure

Before we can address how internet affects learning, one needs to understand how a state develops internet infrastructure. If a newly minted state decided to connect to the global internet connection for the first time, what would they need to do? What factors influence their success? A proper internet adoption means two things for each user: a) reliable and affordable access to the network connection and b) free access within the network to relevant and minimally censored material.[17] Along the path to establishing an internet network, the regime must repeatedly make two important decisions: First, do they create clear regulations and promote a sustainable commercial environment for the private sector that develops the human capital needed to create, sustain, and maintain infrastructure?[18] Second, do they support freedom of internet use or censor offending sites or social media? This not only impacts human rights and freedoms but also has an effect on demand, adoption rates, reliability, speed, and affordability.[19]

There are certain critical junctures where the regime faces important choices regarding internet access. These are: (1) establishing international connections, (2) developing a national backbone for internet infrastructure, (3) stimulating local content development, and (4) promoting last-mile connections.[20]

Before a state can have any internet access, they must organize an access-point to the international network. There are three main ways to achieve this: undersea fiber-optic cables, terrestrial network connections through a neighboring country, or satellite connections. When available, the undersea connection is by far the best approach. Secure submarine lines ensure a quick and affordable connection to the wider internet for an entire country. Unfortunately, some countries are land-locked and must develop terrestrial connections. These lines can face construction challenges, particularly because they require the government to negotiate with neighboring countries in order to connect with foreign lines.[21] While private investors and companies will mostly build these links, the government has an important role in negotiating bilateral and multilateral agreements with neighboring countries to develop permissions and incentives for both physical cables or satellites. Furthermore, they have the power to deny the private introduction of

new connections. Authoritarian governments have also been known to directly block access to content at this point. The decision to block access often further raises prices, slows internet speeds, and decreases competition.[22]

The second component, the national backbone, is the set of high capacity lines that connect regions and cities within the state. They must also connect to the international access point at the border. The state's primary role in developing this infrastructure is to create a good environment for private sector development. For industries to invest in this region, they need to know regulations are consistent fairly enforced. The state may also need to play a role in subsidizing infrastructure investment to connect more remote or mountainous regions.[23] The next element deals more with the demand side of the internet. As mentioned earlier, a free and open internet is important to stimulating demand, though some states, like China, are able to maintain a widely used interface even with their firewall. States need to also develop their own local content. This should be material that is in the local language(s), culturally appropriate and socially relevant to the country. This again means creating a good economic and political environment for human creativity and industry.[24] An important piece of infrastructure to facilitate this is the establishment of Internet Exchange Points (IXPs).[25] IXPs are physical locations where internet infrastructure come together, located on the 'edge' of different networks. Locating an IXP within a country provides a place for local content providers to place their websites. Accessing these sites is faster than accessing internationally located sites since data does not have to travel all the way to an international node before returning to the user.[26] As well as developing laws to protect human creativity and ingenuity, the government can play a role in negotiating the creation of new IXPs in their country.

Finally, last-mile connections are those that connect the national backbone to the individual user. These can include dial-up modems through telephone lines, digital subscriber lines (DSLs), wide area LAN, or line-of-sight antennas.[27] Today, this last-mile connection is increasingly supplied via mobile phones and SIM cards. As many countries develop, they are bypassing the establishment of a physical connection altogether and only offering a mobile data connection – at least for much of the country.[28] This cellular method of mobile data delivery is what I measure in the quantitative portion of this paper.

At each tier of infrastructure, we have seen that the government has a key role in determining how affordable, open, and fast the internet access will be. While internet quality largely depends upon basic economic development and state capacity, democratic institutions appear to be equally important. This is supported statistically as well. Internet download speed (in Megabits per second from 2019) is correlated with V-Dem's Participatory Democracy Index at 0.5312 while 2018 GDP per capita is correlated at 0.7408.[29] Similar to education quality, internet quality appears to be largely determined by two main factors, one of which is state capacity. However, while educational institutional choices are either unaffected by type of governance or biased in favor of authoritarianism, institutional choices around establishing a stable internet infrastructure are biased towards democratic countries. While in theory authoritarian leaders could chooses to support the development of the internet, in most cases they oppose it as a threat to their rule. Their actions of opposing it not only limit the online content available in their countries, but directly slow down the speed of the connection, especially when implemented at the national level. Again, consider the case of China: economically strong and with the 9th highest state capacity in the world (as measured by the ratio of shadow economy to legal economic activity), it boasts some of the world's highest test scores yet maintains the 153rd fastest internet in 2019 at 2.69 Mbps. It also receives a 0.044 on V-Dem's Participatory Democracy Index.

1.3. Disruptive Theory Part Two: From Internet Quality to Educational Outcomes

Since the late 90's and early 2000's, the field of development economics has seen the introduction of an important new technique: field experiments where economists partner with low-income countries or NGOs to conduct randomized controlled trials (RCTs) that assess the effectiveness of programs to assist the poor. Out of these thousands of studies, many have studied the effects of internet and technology use in schools on learning outcomes. Escueta et al. (2017) and J-PAL (2019) conducted two literature reviews of education technology RCTs that included programs that increased basic access to the technology, provided computer-assisted learning, employed technology-enabled behavioral interventions, and assessed online learning. They consistently found evidence that supports the effectiveness of computer-assisted learning and behavioral interventions in increasing student achievement. However, they identified little value in giving a child a device without additional assistance and argue that online-only courses are less effective than face-to-face teaching.[30] Another review has found that mobile technology within the African context increases student engagement and formation of learning communities at schools and universities while still facing infrastructure and access challenges.[31]

There are three primary ways through which open and affordable internet access could improve learning outcomes: 1) increasing access to online learning resources such as eBooks, massive open online courses (MOOC), and education apps or games such as Khan Academy.[32] 2) Internet use combined with technology in schools could improve learning outcomes.[33] This is particularly true if teachers use artificial intelligence to provide personalized learning instruction

to students. 3) Finally, people might learn through passive exposure to English and other ideas by consuming entertaining and educational videos, articles, and social media.[34]

Kim and Lee conducted an econometric analysis of the factors correlated with National Performance on the TOEFL.[35] They built on Snow (1998), the only other study to "investigate factors influencing English skills at cross-country level," and consider two broad categories of factors: linguistic and nonlinguistic.[36] Linguistic factors indicate how similar a language is to English and nonlinguistic include measures like GDP per capita, globalization, or expected years of schooling. They found that linguistic 'closeness' to English and expected years of schooling were most related to a higher level of proficiency. Most pertinently, they found that the internet was strongly correlated with TOEFL scores in 1997/1998 but was no longer statistically significant in 2004/2005.[37] They suggest this is because a higher percentage of the internet was in English in 1998 than in in 2005. I draw heavily from Kim and Lee's paper to help determine which covariates to include in the quantitative evaluation.

2. Quantitative Analysis: Mobile Data Affordability on English Proficiency

Despite the proliferation of international education assessments, most notably OECD's Program for International Student Assessment (PISA),[38] there is still no internationally comparable exam math or science exam that covers a majority of countries in the world. English foreign-language proficiency exams come closest. Since English proficiency is essential for conducting business, governance, and communications in a globalized world, many people try to learn the language in every country. Several authoritative exams tests individuals' ability for jobs, schools, or immigration applications. Of these, the Test of English as a Foreign Language (TOEFL) exam captures results from 168 countries. While these results are only generalizable insofar as we recognize that only a certain subset of students are likely to take the test, they do have truly international representation. If a certain type of country is less likely to have students trying to learn English, this may limit generalizability to overall education quality. Additionally, it is seems likely to the author that English-learning would be more assisted by the internet than other subjects because students may pick up English words through exposure to the English social-media and videos pervasive to the internet. Finally, because of the English language's dominance online, there are naturally more English-learning resources than subject-specific lessons in any one foreign language.

This quantitative analysis tests the hypothesis that mobile data affordability affects TOEFL scores. According to basic supply and demand theories, if mobile data costs a higher percentage of GDP per capita, individuals should consume less data on average. Furthermore, since the two main factors determining mobile data prices are a consumer's abilities to pay and the quality of infrastructure, the data prices should not merely correlate with economic well-being.[39] For example, the country with the highest 1 GB data cost in the world is Zimbabwe ($75.20/1 GB) while the cheapest data is available in India ($0.26/1 GB). Meanwhile, both countries have a low GDP per capita ($2,688.41 PPP in Zimbabwe versus $6,899.20 PPP in India).[40]

2.1. Data & Summary Statistics

TOEFL is one of the oldest and most ubiquitous language tests. It is mostly taken by international secondary school students seeking to study in English-speaking countries. However, English as a Second Language (ESL) students in English-speaking countries also take it. The TOEFL exam is a reasonable choice to operationalize the outcome of interest because of its preeminence and history of use in social science research, its inclusion of 168 countries, and because it has easily accessible data.[41] It is scored out of 120 points made up of a sum total of reading, listening, speaking, and writing subcomponents worth 30 points each.

The measure of mobile data affordability is the cost of one gigabyte of mobile data as a percentage of gross domestic product per capita, accounting for purchasing power parity. A higher percent represents internet that is less affordable and therefore less accessible. The data for the GDP per capita (PPP) comes from the World Bank's Data Bank.[42] The data of mobile-data costs was produced in research funded by cable.co.uk. Their stated goal in producing this free and public data is to "demonstrate the state of mobile data pricing across as much of the globe as is possible" for use by educational research, governmental bodies, NGOs, or as daily news.[43] Cable.co.uk gathered the data from 6,313 mobile data plans in 230 countries during 2018.

Kim & Lee identified which linguistic and non-linguistic factors influence English as a foreign language (EFL) proficiency. They developed three linguistic factors, two of which are related to linguistic distance – how different a language is from English. I tested both these factors, but found they had an insignificant effect on TOEFL scores. Because of this insignificance, their unrelatedness with data affordability, and their limitation on the total number of

observations, I have excluded these factors from the model. In their place, I regress a dummy variable that simply indicates whether or not English is an official language in the country. This variable is collinear with the HAV and Word Order variables but preserves sample size. It is constructed from data found on the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII)'s website.[44] The third linguistic factor Kim & Lee considered was Linguistic Fractionalization – the probability that two individuals from one population belong to different language groups. A higher linguistic fractionalization score means children are exposed to more foreign languages as children and perhaps more likely to score better. Alternatively, its high correlation with ethnic fractionalization may reflect greater conflict in these regions and resultant lower scores. I gather this data from Alesina et. al's linguistic fractionalization index.[45]

This regression also includes several economic and social factors. The economic factors are globalization and the number of fixed broadband subscriptions per 100 people. I follow Kim & Lee's lead in adopting the Konjunkturforschungsstelle (KOF) index of globalization to capture a country's degree of globalization and exposure to English. The KOF index is a measure of the overall degree of globalization based on economic, social, and political factors.[46] To focus the research on the effect of mobile internet, I added broadband rates which indicate access to high-speed internet through a variety of mediums in a fixed location. Finally, this analysis included democratization scores. I draw the democratization data from V-Dem's Participatory Democracy Index.[47] V-Dem is the current authoritative on democracy and participatory democracy reflects "active participation by citizens in all political processes, electoral and non-electoral."[48] This was chosen over other indications of democracy because it is the indicator of democracy with the highest standard of inclusive and full democratization.

Table 1. Summary statistics

	Mean	St.Dev	Min	Median	Max	Count
Total Score	82.515	9.307	62	83	101	167
Reading Score	19.222	2.834	13	20	25	167
Listening Score	20.97	2.851	14	21	26	167
Speaking Score	21.635	1.998	17	22	26	167
Writing Score	20.599	2.018	16	21	25	167
Data Affordability (%)	.1	.181	.002	.034	1.191	179
Log Data Affordablity	-3.32	1.44	-6.342	-3.375	.175	179
Cost 1GB of data	8.242	8.938	.26	5.25	65.83	229
Official English	.335	.473	0	0	1	224
Linguistic Fractionalization	.389	.28	.002	.353	.923	197
KOF Globalization	61.682	14.317	29.974	60.148	91.168	207
Broadband per 100 people	15.172	14.086	.001	11.794	54.146	213
V-Dem Score	.333	.192	.014	.322	.778	178

2.2. Model

I test the connection between mobile internet affordability and English learning by regressing a multivariate linear model of the cost of 1 GB of mobile data on TOEFL score outcomes. As a cross-sectional analysis with a non-random treatment, the results will never be causal. However, testing for potential covariates mitigates omitted variable bias. The cross-country regression is an appropriate technique to assess the macro level effect of mobile data affordability on English learning because it can provide a strong suggestive correlation. As well, since a complete tabulation of mobile data prices only became available in 2018, there is little to no historical data to measure the effect with panel data. The following equation represents the preferred model:

$$Score = \beta_0 + \beta_1 \, log(DataAffordability) + \beta_2 OfficialEnglish + \beta_3 LinguisticFrac + \beta_4 Globalization + \beta_5 FixedBroadbandSubs + \beta_6 VDemScore + \varepsilon$$

Score is the average score on the TOEFL test in a country. It is the outcome of interest, though each subcomponent score is considered separately as different outcomes. β_0 represents the constant and holds little meaning since the equation captures the differences between countries. $log(DataAffordability)$, predicted to be negative, is the factor

of interest. Data affordability is measured by the cost of one GB of mobile data as a percentage of GDP per capita, PPP and logistically transformed because of its visible relationship with test scores. Its correlation coefficient increases from -0.22 to -0.54 when logistically transformed. The variable measures the percent change in the percentage cost.

$OfficialEnglish$ signifies whether one of a country's official languages is English. It is binary and the predicted sign is positive. $LinguisticFr$ac is the probability that two individuals in the same country speak different languages and β_3 should be negative. $FixedBroadbandSubs$ indicates how many fixed or home internet subscriptions there are per 100 people. $VDemScore$ indicates the V-Dem index's Participatory Democracy score, predicted to have a positive sign. Finally, $Globalization$ indicates the KOF globalization index and has a positive sign.

2.3. Results

Data affordability raises average TOEFL scores across countries all else held equal and constant. The table below displays the regression results. Each column, moving from left to right, adds additional control variables. Model 1 is the simple bivariate regression. The main effect is large and statistically significant. Model 2 adds the linguistic factors of whether the official language is English and the linguistic fractionalization score. Both these indicators are statistically significant at the 1% level and cause the magnitude of β_1 to decrease slightly.

Table 2: The Impact of Internet Affordability on TOEFL Scores

	(1)	(2)	(3)	(5)
VARIABLES	Model 1	Model 2	Model 3	Model 4
Log Data Affordability	-4.158***	-3.549***	-1.421**	-1.572***
	(0.435)	(0.387)	(0.543)	(0.473)
English Official Language		7.552***	6.666***	5.896***
		(1.317)	(1.285)	(1.299)
Linguistic Fractionalization		-11.31***	-4.427*	-4.246*
		(2.372)	(2.281)	(2.315)
KOF Globalization Index			0.276***	0.209***
			(0.0626)	(0.0760)
Fixed Broadband per 100			0.162**	0.0989
			(0.0619)	(0.0667)
V-Dem Score				9.478***
				(3.181)
Constant	67.82***	73.19***	59.88***	59.25***
	(1.748)	(1.904)	(3.895)	(4.001)
Observations	147	138	123	122
R-squared	0.339	0.478	0.647	0.659

Notes: Dependent variable is the total average TOEFL scores for each country in 2018. All coefficients are estimated with OLS. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Model 3 adds economic and technological factors. The *KOF* Globalization index has a positive and highly significant sign. The number of fixed broadband subscriptions per 100 people is statistically significant. This variate captures internet access that is not mobile – important since many African countries have developed a mobile infrastructure and skipped over fixed access. Broadband and internet affordability are correlated at -0.53. With these additions, the size of β_1 decreases significantly. Finally, Model 4 adds the V-Dem participatory democracy score. The democratization index has a large and statistically significant effect. In this final specification, data affordability is statistically significant to the 1% level and β_1 is -1.572. This means that all else held equal and constant, across countries, if the cost of 1 GB of data as a percentage of GDP is increased by one percent, TOEFL scores decrease by 0.016 points.

As seen in the discussion above, with the addition of linguistic and economic factors the effect of internet affordability decreases slightly. This might suggest that the effect would eventually disappear if one added enough covariates. However, other factors, particularly social and democracy scores, do not affect the size. From a theoretical perspective, it is unclear what other variables might affect TOEFL scores. The economic significant of a percentage change in a percent is hard to intuitively understand. We can better grasp this effect by considering the point change over a change in percentiles. A change from the most affordable price to the 25th percentile is correlated to an 9-point decline in score. The next 25th percentile to the median corresponds with a 3.2-point decrease, another -3.7-points to the 75th percentile, and finally a 33.7-point decrease predicted from the upper 75th percentile to the least affordable price. While the extremes are inaccurate, this middle 50th percentile reflects a 7-point gap between the 25th and 75th percentiles. On a test of only 120 points with a standard deviation of 9 points, this is significant.

Table 3 displays the final model comparing each subcomponent score as an outcome of interest. The listening and reading scores have the largest coefficients, β_1 is -0.57 for listening and -.447 for reading, and most significant effects, both at the 1% level. Meanwhile, the speaking and writing scores are both only significant to the 5% level with β_1 of 0.307 for writing and -0.266 for speaking. Therefore, most of the impact of data affordability on English scores come from its effect on listening and reading abilities.

Table 3: Subcomponent Scores

	(1)	(2)	(3)	(4)
VARIABLES	Reading Score	Listening Score	Speaking Score	Writing Score
Log Data Affordability	-0.447***	-0.570***	-0.266**	-0.307**
	(0.148)	(0.135)	(0.121)	(0.120)
Controls	Yes	Yes	Yes	Yes
Constant	13.03***	13.98***	16.59***	16.07***
	(1.369)	(1.139)	(1.055)	(0.973)
Observations	122	122	122	122
R-squared	0.664	0.685	0.497	0.568

Notes: Dependent variable is the total average TOEFL scores for each country in 2018. Controls are Covariates in Model 5 above. All coefficients are estimated with OLS. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

3. Discussion

These results are significant across several statistical tests and analyses. Though the data is not a representative sample, since it comes from every country in the world, it includes the entire population. One important caveat to this point is that the generalizability is still limited: not every student takes the TOEFL exam and so, even if scores came from every country in the world, they would only represent those students who have the resources and inclination to study English. Heteroskedasticity was corrected for with a robust variance estimator. Though it cannot be declared causal, the correlation is statistically and economically significant and robust.

The stronger subcomponent effects on reading and writing have two implications. First, this is evidence against reverse causality. One might suggest that countries with better English skills would be able to attract more investment in internet and technology because of the greater ease with which foreign companies could operate if many residents speak English. If this was the case, we would expect to see roughly the same correlation across all four components or maybe even a stronger effect connected to the speaking component. Instead, we find a much stronger effect among the reading and listening subcomponents and the smallest correlation with speaking scores. Second, this supports the third mechanism presented in the literature review for how the internet could affect learning: passive exposure to English and other ideas by consuming entertaining and educational videos, articles, and social media. It is logical that this passive exposure would have the greatest effect on reading and listening. Of course, this does not indicate that online courses or apps have no effect since they may simply have the greatest benefit for reading and listening skills. Or, these tools might already be contributing to the smaller but non-negligible effects on speaking and writing abilities.

Additional research might explore the link between internet speed and education quality through alternate approaches. First, a time-series analysis would enable us to better understand how the path of internet development affects education within a country. Do test scores improve as better internet infrastructure is installed or global

technology improves? Intra-country quantitative analysis comparing similar regions with different levels of internet infrastructure are another critical step in understanding the role of the internet in learning outcomes. Additionally, the theoretical basis for a relationship between democracy and higher quality internet should be explored empirically.

While this paper does not discount nor disprove the previous theory of education development, it has explored and presented a plausible alternative causal logic, based in the emergence of internet technology. This theory has two strong implications. First, the large-n analysis supports the argument that internet technology might have a positive influence on learning outcomes. This suggests governments or NGOs should choose to invest in education technology in hopes of improving outcomes. Second, this disruptive theory implies that states would not need to wait for economic growth or capacity development to begin democratizing. Under the disruptive theory, in contrast to some interpretations of modernization theory, foreign democracies have little reason to discourage democratization in states deemed too poor to democratize yet. They should encourage democratization and the positive feedback loops that come with it in almost any economic situation. Furthermore, the historicity of this claim – that the emergence of the internet might fundamentally flip at least one element of modernization theory – suggests that theories of development theories should be more contextually sensitive than most currently are. Rather than making grandiose or universal claims about development, we should recognize the complex differences of every situation while still striving for parsimonious theories.

4. Endnotes

1 Steven Radelet, *The Great Surge: The Ascent of the Developing World* (New York, NY: Simon and Schuster, 2015).

2 Edward Miguel and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica* 72, no. 1 (2004): 159–217. This famous study is one amongst many that establish a robust connection between medical interventions and increasing school attendance.

3 Abhijit Banerjee et al., "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries," *Science* 348, no. 6236 (May 15, 2015): 1260799, https://doi.org/10.1126/science.1260799. Again, just one of a multitude of studies establishing the link.

4 Radelet, *The Great Surge: The Ascent of the Developing World*, 90.

5 Ibid.

6 Esther Duflo, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," The American Economic Review 91, no. 4 (2001): 795–813. Duflo evaluates one such policy decision and finds that a government decision to build schools resulted in an average increase of 0.12 to 0.19 years of education, as well as a 1.5 to 2.7 percent increase in wages for each primary school constructed per 1,000 children.

7 Richard A. Easterlin, "Why Isn't the Whole World Developed?," The Journal of Economic History 41, no. 1 (1981): 1–19.

8 Francis Fukuyama, "The Imperative of State-Building," *Journal of Democracy* 15, no. 2 (2004): 17–31, https://doi.org/10.1353/jod.2004.0026.

9 For a literature review, see: Eric A. Hanushek and Ludger Woessmann, "The Economics of International Differences in Educational Achievement," in *Handbook of the Economics of Education*, vol. 3 (Elsevier, 2011), 89–200, https://doi.org/10.1016/B978-0-444-53429-3.00002-8.

10 Ludger Woessmann, "The Importance of School Systems: Evidence from International Differences in Student Achievement," *The Journal of Economic Perspectives* 30, no. 3 (2016): 10.

11 Woessmann, "The Importance of School Systems: Evidence from International Differences in Student Achievement."

12 Woessmann finds that giving schools greater autonomy in budgets, performance goals, and decisions of what to teach have a negative effect on achievement. However, schools should receive autonomy in how to teach these things: what teaching styles to implement or resources to use.

13 Ibid., 27.

14 Bruce Bueno de Mesquita et al., "The Logic of Political Survival," September 24, 2002, 962; Daron Acemoglu et al., "Democracy, Redistribution, and Inequality," in *Handbook of Income Distribution*, vol. 2 (Elsevier, 2015), 1885–1966, https://doi.org/10.1016/B978-0-444-59429-7.00022-4.

15 Sirianne Dahlum and Carl Henrik Knutsen, "Do Democracies Provide Better Education? Revisiting the Democracy–Human Capital Link," World Development 94 (June 1, 2017): 186–99, https://doi.org/10.1016/j.worlddev.2017.01.001.

16 Fukuyama, "The Imperative of State-Building."

17 "A Policy Framework for Enabling Internet Access" (Internet Society, April 2017).; "World Development Report 2016: Digital Dividends," *World Development Report*, n.d., 359.

18 "Internet Access and Education: Key Considerations for Policy Makers" (Internet Society, November 2017); Joakim Reiter, "4 Ways Governments Can Develop Digital Infrastructure," World Economic Forum, accessed February 27, 2020, https://www.weforum.org/agenda/2017/09/governments-develop-digital-infrastructure-vodafone/.

19 "World Development Report 2016: Digital Dividends," *World Development Report*, n.d., 221-222.

20 Peter Lovelock, "Digital Economy Study in Central and West Asia: Unleashing the Potential of the Internet in Central Asia, South Asia, the Caucasus and Beyond" (Asian Development Bank, December 2015).

21 "Internet Crossing Borders: Boosting the Internet in Landlocked Developing Countries" (Internet Society, June 20, 2018).

22 "World Development Report 2016: Digital Dividends."

23 "A Policy Framework for Enabling Internet Access."

24 Lovelock, "Digital Economy Study in Central and West Asia: Unleashing the Potential of the Internet in Central Asia, South Asia, the Caucasus and Beyond."

25 "A Policy Framework for Enabling Internet Access."

26 Ronan Shenhav, "Tajikistan Internet Exchange Point Environment Assessment" (Internet Society, CAREN, June 2017).; "A Policy Framework for Enabling Internet Access."

27 Walker, Personal conversation about internet establishment in Tajikistan.

28 "A Policy Framework for Enabling Internet Access."

29 Download speed is taken from Cable. "Worldwide Broadband Speed League 2019 | Cable.Co.Uk." Accessed May 31, 2020. https://www.cable.co.uk/broadband/speed/worldwide-speed-league/.; Democratization scores from "Home | V-Dem." Accessed March, 2020.. https://www.v-dem.net/en/.; and GDP per capita from "World Bank Open Data | Data." Accessed December 13, 2019. https://data.worldbank.org/.

30 Maya Escueta et al., "Education Technology: An Evidence-Based Review" (Cambridge, MA: National Bureau of Economic Research, August 2017), https://doi.org/10.3386/w23744; Sophie Shank, "Will Technology Change Education for the Better?," Evidence Review. (Cambride, MA: Abdul Latif Jameel Poverty Action Lab, 2019).

31 Rogers Kaliisa and Michelle Picard, "A Systematic Review on Mobile Learning in Higher Education: The African Perspective," *The Turkish Online Journal of Educational Technology* 16, no. 1 (2017): 18.

32 Khan Academy is an online source that "offers practice exercises, instructional videos, and a personalized learning dashboard that empower learners to study at their own pace in and outside of the classroom." https://www.khanacademy.org/

33 Shank, "Will Technology Change Education for the Better?"

34 Myung-Hee Kim and Hyun-Hoon Lee, "Linguistic and Nonlinguistic Factors Determining Proficiency of English as a Foreign Language: A Cross-Country Analysis," *Applied Economics* 42, no. 18 (July 2010): 2347–64, https://doi.org/10.1080/00036840701857960.

35 Kim and Lee, "Linguistic and Nonlinguistic Factors Determining Proficiency of English as a Foreign Language."

36 MS Snow, "Economic, statistical, and linguistic factors affecting success on the test of English as a foreign language (TOEFL)" (Information Economics and Policy, 1998)

37 Kim and Lee, "Linguistic and Nonlinguistic Factors Determining Proficiency of English as a Foreign Language."

38 "PISA - PISA." Accessed May 31, 2020. https://www.oecd.org/pisa/.

39 Cable research

40 "World Bank Open Data | Data." Accessed December 13, 2019. https://data.worldbank.org/.

41 "Test and Score Data Summary for the TOEFL IBT ® Tests"

42 "World Bank Open Data | Data." Accessed December 13, 2019. https://data.worldbank.org/.

43 "Worldwide Mobile Data Pricing League: Cost of 1GB in 230 Countries." (Cable.co.uk) This UK company's business model is providing simple and clear online comparisons of different broadband internet service providers. It takes a small commission from each use.

44"CEPII - Database." Accessed December 13, 2019. http://www.cepii.fr/cepii/en/bdd_modele/bdd.asp. The data base can be accessed at: http://www.cepii.fr/distance/geo_cepii.xls

45 A. Alesina, Fractionalization (Journal of Economic Growth, 2003)

46 "KOF Globalisation Index." Accessed December 13, 2019. https://kof.ethz.ch/en/forecasts-and-indicators/indicators/kof-globalisation-index.html.

47 "Home | V-Dem." Accessed May 31, 2020. https://www.v-dem.net/en/.

48 Ibid.