INTERNATIONAL MONETARY FUND

Determinants and Social Dividends of Digital Adoption

Utkarsh Kumar, David Amaglobeli, and Mariano Moszoro

WP/23/65

2023
MAR

WORKING PAPER

 WP/23/65

IMF Working Paper
Fiscal Affairs Department

Determinants and Social Dividends of Digital Adoption
Prepared by Utkarsh Kumar, David Amaglobeli, and Mariano Moszoro*

Authorized for distribution by Rodrigo Cerda
March 2023

ABSTRACT: We identify key drivers of digital adoption, estimate fiscal costs to provide internet subsidies to households, and calculate social dividends from digital adoption. Using cross-country panel regressions and machine learning, we find that digital infrastructure coverage, internet price, and usability are the most statistically robust predictors of internet use in the short run. Based on estimates from a model of demand for internet, we find that demand is most price responsive in low-income developing countries and almost unresponsive in advanced economies. We estimate that moving low-income developing and emerging market economies to the levels of digital adoption in emerging and advanced economies, respectively, will require annual targeted subsidies of 1.8 and 0.05 percent of GDP, respectively. To aid with subsidy targeting, we use microdata from over 150 countries and document a digital divide by gender, socio-economic status, and demographics. Finally, we find substantial aggregate and distributional gains from digital adoption for education quality, time spent doing unpaid work, and labor force participation by gender.

RECOMMENDED CITATION: Kumar, U., D. Amaglobeli, and M. Moszoro, 2023, "Determinants and Social Dividends of Digital Adoption", IMF Working Paper 23/65.

JEL Classification Numbers:	L96, H23, H50
Keywords:	Social Dividends; Digitalization; GovTech
Author's E-Mail Address:	uk2154@columbia.edu; damaglobeli@imf.org; mmoszoro@imf.org

* We are grateful for generous comments and feedback from Marwa Alnasaa, Angham Al Shami, Maureen Burke, Rodrigo Cerda, Ignatius de Bidegain, Ruud De Mooij, Wala'a El Barasse, Noha El-Gebaly, Wandja Fondja, Vitor Gaspar, Nikolay Gueorguiev, Laura Jaramillo Mayor, Safari Kasiyanto, Michael Kremer, Jeong Dae Lee, Albert Nyikuli, Rumit Pancholi, Mauricio Soto, Tao Sun, and Irene Yackovlev.

WORKING PAPERS

Determinants and Social Dividends of Digital Adoption

Prepared by Utkarsh Kumar, David Amaglobeli, and Mariano Moszoro [1]

[1] We are grateful for generous comments and feedback from Marwa Alnasaa, Angham Al Shami, Maureen Burke, Rodrigo Cerda, Ignatius de Bidegain, Ruud De Mooij, Wala'a El Barasse, Noha El-Gebaly, Wandja Fondja, Vitor Gaspar, Nikolay Gueorguiev, Laura Jaramillo Mayor, Safari Kasiyanto, Michael Kremer, Jeong Dae Lee, Albert Nyikuli, Rumit Pancholi, Mauricio Soto, Tao Sun, and Irene Yackovlev.

Contents

I. Introduction

Innovations in Information and Communications Technologies (ICT) over the last two decades have had transformative impact on societies. ICT has changed how individuals acquire and share information and receive services, how firms compete in the marketplace and how governments interact with citizens and businesses. These transformative changes, however, have not benefited all equally and has led to a large "digital divide" within and between countries. The large digital divide must be a concern for policymakers not only from equity but also from efficiency perspective. The digital divide results from two main economic forces: supply-side forces, which determine availability of ICT, and demand-side forces, which determine the pace of technology adoption. In this paper, we focus on the demand side forces and try to find answers to three policy-relevant questions: What are key determinants of digital adoption? What is the role of public policy in encouraging digital adoption and closing the digital divide? And what does the society gain from accelerating the pace of digital adoption? Our emphasis on the demand-side makes this paper complementary to the World Bank's World Development Report 2016 that extensively covers supply-side policy.

We first develop a list of determinants of digital adoption. This list includes past and current investments in ICT, complementary public infrastructure for ICT use (e.g., electricity access), features of the ICT market (e.g., prices, coverage, usability), different levels of skills and income, sectoral composition of the economy, labor market participation and several structural variables (e.g., inequality, women empowerment, level of urbanization). Using within-country variation and conditional on GDP per capita, we find that digital adoption—measured by the proportion of population who report to have used Internet in past three months—is positively correlated with ICT investments, the level of complementary public infrastructure, coverage of ICT, usability, share of services sector in the economy, female empowerment, and urbanization and negatively correlated with ICT costs, inequality, and poverty.

Next, we use machine learning to select the most robust statistical predictors of digital adoption from our compiled list of determinants. This exercise provides us with a short list of variables strongest predictors of digital adoption that we classify into two kinds of policy handles: "Fast Moving" policy variables include price, mobile phone ownership, coverage as well as use-cases and "Slow Moving" policy variables include skills, labor force participation, inequality, urbanization, and the services' share in the GDP.

Further, we develop a model of demand for internet use and estimate it using aggregate country level data on prices and demand for internet. We find substantial heterogeneity in demand elasticities across countries in different income groups with consumers in low-income developing countries (LIDCs) being the most price sensitive. Using our estimated price elasticities of demand, we calculate subsidies that will be required to increase internet use in LIDCs and emerging market economies (EMEs) to the level of advanced economies (AEs). Assuming subsidies are untargeted (i.e., all users receive discounts on internet fees) fiscal costs are considerable at 0.3 percent and 3.2 percent of annual GDP for EMEs and LIDCs, respectively. When subsidies are targeted to those deprived of internet fiscal costs drop to 0.1 and 1.8 percent of annual GDP for EMEs and LIDCs, respectively.

Having established the most statistically robust predictors of digital adoption along with the importance of targeting subsidies, we explore the composition of digital divide using two important “Fast Moving” policy variables—mobile phone ownership and use-case for the internet. Combining a representative micro-dataset from over 150 countries for 2017 with panel regression techniques and controlling for income, we find strong evidence of digital divide across multiple dimensions: gender, education, age, and labor force participation. Strikingly, women with tertiary education in LIDCs are on average less likely to own a mobile phone than (less than) primary-level educated men, individuals over 65 years of age are less likely to own mobile phones than teenagers and being in labor force is associated with a 5–10 percent increase in likelihood of mobile phone ownership across the world even after controlling for education, gender, age, and income. Similar disparities exist for various uses of internet that we observe in the data: online bill payment, online purchases, online banking and receiving government transfers on phone.

Lastly, we assess potential benefits that societies can receive from faster digital adoption. To this end, we provide estimates of gains from internet use in terms of education quality, time use and labor force participation. We estimate the effect of internet use on our three outcomes using two econometric strategies: difference in differences and instrumental variables. We find that on average, increasing internet use from 10 to 90 percent increases labor force participation by 6 percentage points. The effect is larger for women (7.4 percentage points) than men (4.8 percentage points). A similar increase in internet use increases secondary school test scores by 16 percent on base average. This increase in test scores implies substantial improvements in learning outcomes with potential gains in the future, especially for LIDCs. We find weak evidence that internet use reduces time spent doing unpaid work for women.

Other studies have also found positive impact of ICT adoption on economies and societies. For example, ICT adoption was found to improve market coordination (Jensen 2007, Aker and Mbiti 2010, Aker and Ksoll 2016), firm outcomes (Aker and Ksoll 2016, Jensen and Miller 2018), households access to finance (Jack and Suri 2014, Blumenstock and others 2016), last-mile service delivery (Haushofer and Shapiro 2016, Muralidharan and others 2021), household savings (Jack and Suri 2014, Haushofer and Shapiro 2016, Björkegren and others 2022), financial management (Karlan and others 2016, Kast and others 2012), health (Dammert and others 2014, Flax and others 2014), trade (Akerman and others 2022) as well as human capital and labor market opportunities (Muralidharan and others 2019, Aker and others 2012, Dammert and others 2015, Bettinger and others 2020). The recent literature has also highlighted adverse effects of internet use like digital addiction (Allcott and others 2022).

Using most conservative point estimates from the regression analyses we perform back-of-the-envelope calculations of potential gains from digital adoption. Specifically, we assume an increase in internet use in

LIDCs to the level of EMEs and in EMEs to the level of AEs. We follow Schoellman 2012 to calculate monetized benefits from improved education scores (a proxy for education quality). Using data from IPUMS International, we first back out returns to education for male immigrants in the US by country of birth and education. Assuming that differences in returns to schooling represent differences in education quality, we construct a cross-country log-log relationship between estimated returns to schooling and most recent average test scores for countries in our sample. Finally, we use the derived relationship along with increases in test scores from internet use to estimate benefits. We find average annual gains of 1.1 and 2.3 percent of GDP for EMEs and LIDCs, respectively.

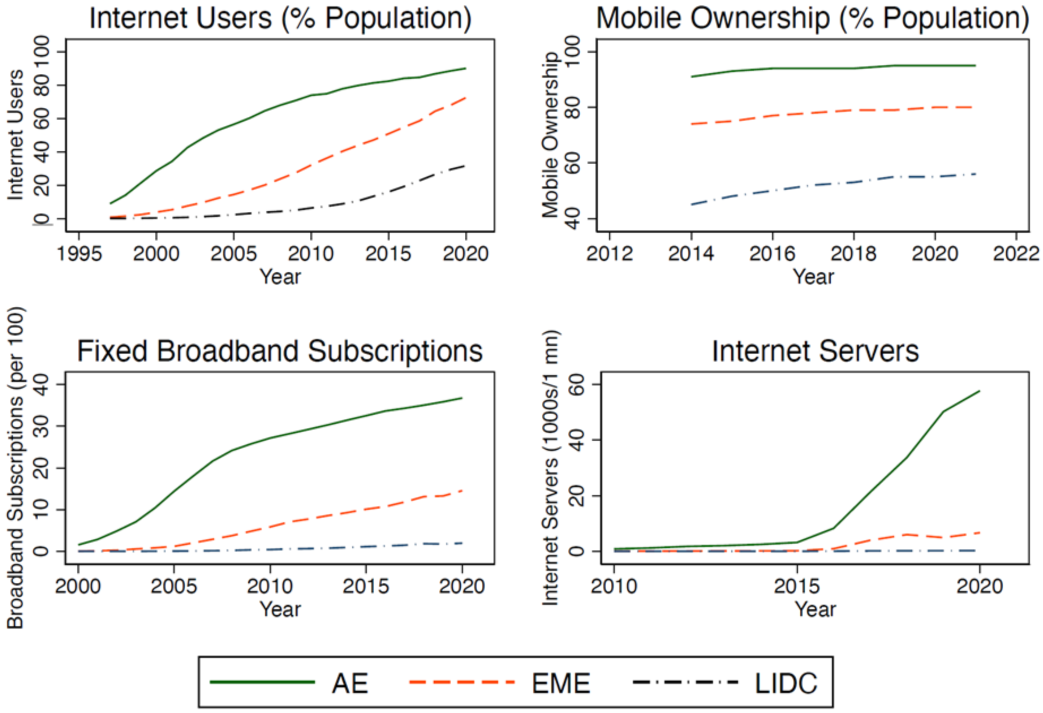
The benefits from increased labor force participation are large and redistributive across gender with women accounting for majority of the gains (in line with Chiplunkar and Goldberg 2022). We find a gain from increased labor force participation of 1.8 and 0.8 percent GDP for EMEs and LIDCs, respectively. The increase in female labor force participation accounts for most of these gains.

II. Determinants of Digital Adoption

The cross-country data reveal large disparities in digital adoption.[1] Despite gains in recent years the share of population using internet is substantially below in LIDCs compared with AEs and EMEs (Figure 1). For example, at the end of 2020, 90 percent of population in AEs reported using internet compared with 72 percent for EMEs and 32 percent for LIDCs. The numbers for fixed broad subscriptions are substantially lower for all country income groups with 36 percent in AEs against 14.5 percent in EMEs and only 2 percent for LIDCs. Mobile phone ownership, which is relatively high in most countries, also shows strong disparities with about 95 percent in AEs, 80 percent in EMEs, and 56 percent in LIDCs. The large differences in mobile adoption versus adoption of broadband subscriptions in LIDCs suggest that mobile internet will likely be more important for faster digital adoption in poorer countries. The number of internet servers, which captures the extent of availability of digital infrastructure, shows less disparity between LIDCs and EMEs but large disparity between AEs and the other countries. In terms of geographic regions, digital adoption is the lowest in Sub-Saharan Africa (Figure 2).

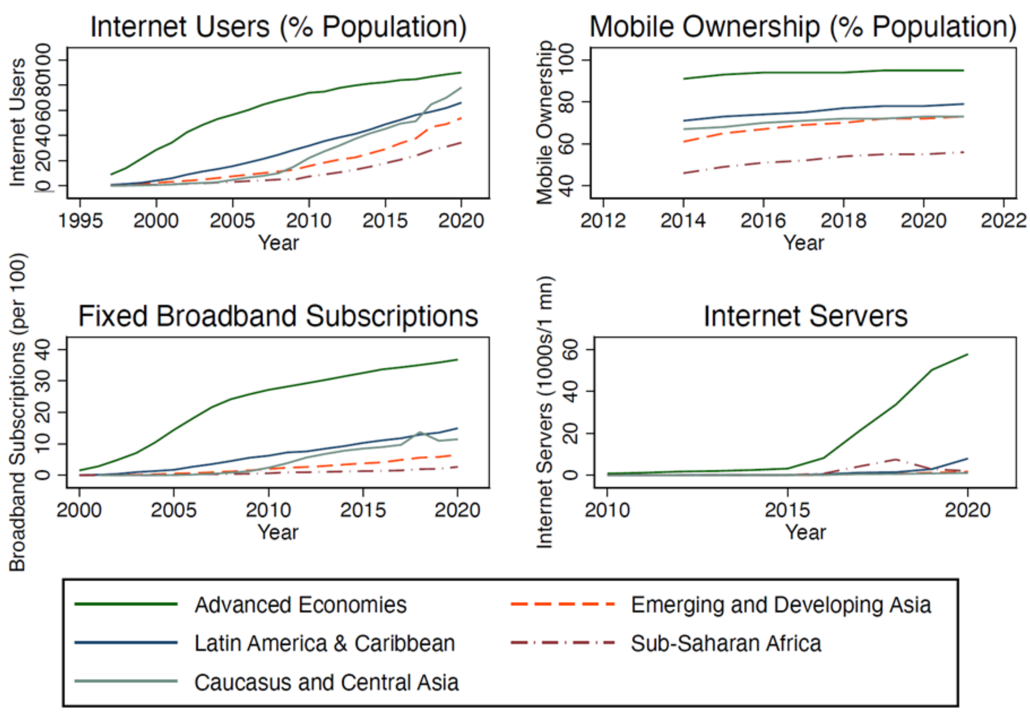
[1] The data comes from four major sources: World Bank's World Development Indicators (WDI), International Telecommunications Union (ITU), World Connectivity Index (GSMA) and World Bank's FINDEX Micro-Data.

Figure 1. Trends in ICT Expansion by Country Income Groups



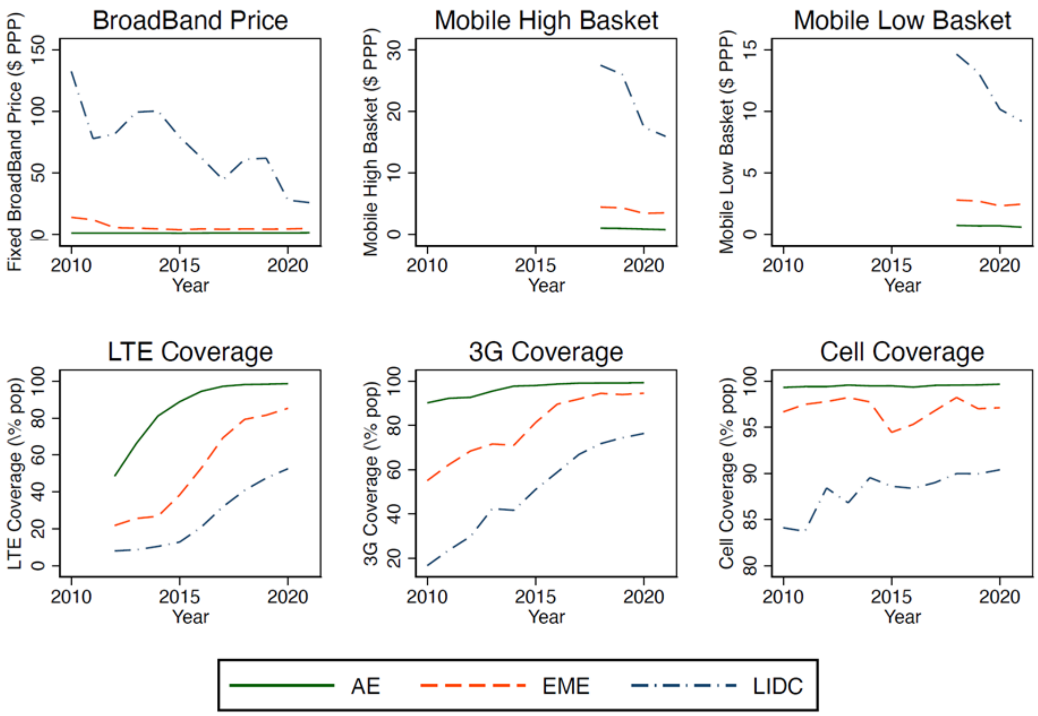
Sources: World Bank, ITU and authors' calculations.

Figure 2. Trends in ICT Expansion by Geographic Regions



Sources: World Bank, ITU and authors' calculations.

Figure 3. Trends in Internet Coverage and Prices



Source: International Telecommunication Union (ITU).

Although AEs are significantly ahead of the other countries in terms of the number of internet servers the differences in signal coverage are much smaller. For example, in terms of 3G signal coverage and cell phone connectivity EMEs and LIDCs are close to AEs (Figure 3). However, differences in 4G connectivity, which allows for much faster video information transmission compared with 3G, are large, especially between AEs and LIDCs. Effective prices of using the internet (that account for purchasing power parity PPP) are the highest in LIDCs despite a sharp decline in recent years. Overall, it appears that LIDCs lag in internet use, have higher effective costs of using the internet, and have slightly lower coverage but poorer quality of service.

We begin our analysis with identifying a broad range of potential determinants of digital adoption from the literature and calculate simple correlations. We use annual cross-country data for all identified variables (Annex I provides the definitions and sources of these data along with their units of measurement). We fill in missing values using linear interpolation and extrapolation within a reasonable range and arrive at a total sample of 4,124 country-year observations comprising 169 countries from 1997-2021.[2] Figure 4 shows simple correlations—coefficient estimates of β_1 in equation (1)—between the internet use (a measure of internet adoption) and each of the potential determinant taking one determinant at a time. Specifically, we estimate the following specification:

[2] For more details on interpolation and extrapolation, see Annex II.

$$Internet\ Use_{ct} = \beta_0 + \beta_1 Determinant_{ct} + \beta_2 GDP\_PC_{ct} + \gamma_c + \epsilon_{ct} \quad (1)$$

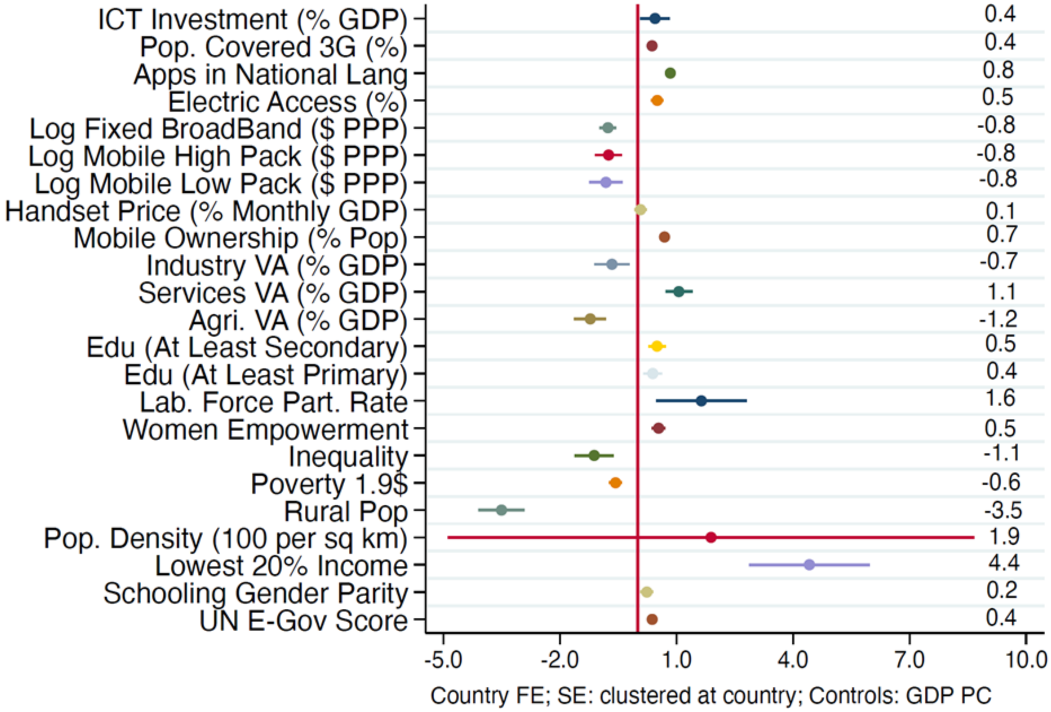
where *Internet Use*$_{ct}$ is the share of population using internet in country *c* in year *t* and *Determinant*$_{ct}$ is the value corresponding to each potential determinant observed in country *c* in year *t*. Each regression controls for contemporaneous level of GDP per capita in each country and only exploits within country variation. Standard errors are clustered at the country level to account for within country correlation in the error term.

The directions on these correlations provide important and intuitive foundation for the analysis.[3] We observe that while holding average income and time in-varying country characteristics constant, internet use is positively correlated with new and existing investments in ICT infrastructure captured by ICT Investment and Coverage of 3G, respectively (possibly suggesting the importance of accessibility and quality), access to complementary infrastructure like electricity, share of the services sector in the economy, average skills as captured by primary and secondary education completion rates,[4] structural variables like labor force participation and level of urbanization, gender parity (norms) as captured by women empowerment, and finally the regulatory environment and governments' emphasis on ICT. Internet use is negatively correlated with prices of operation/subscription, suggesting that operating costs are important frictions to digital adoption. Finally, inequality and poverty are negatively correlated with internet use. Since the regressions control for the level of GDP per capita, both these variables suggest that frictions for digital adoption are likely most binding for the poorest individuals within countries. Appendix Figures AIII.1 and AIII.2 confirm these findings using within income group—year variation as well as an independently collected measure of internet adoption—access to internet at home, as collected by the ITU.

[3] The magnitudes on these correlations across determinants are not comparable and do not have a causal interpretation. As discussed later, multivariate regressions have a somewhat stronger claim on causality, albeit not perfectly due to likelihood of having omitted factors that might be correlated with determinants.

[4] Note that secondary education completion rate correlates more strongly with internet use than primary education completion, suggesting an increasing role of skills in using the internet.

Figure 4. Correlates of Internet Adoption



Source: Authors' estimates.
Note: Country fixed effects and standard errors clustered at country level. Controls include GDP per capita.

While these simple correlations are indicative it is possible that they mask an indirect correlation coming from an omitted variable. To test how robust directions of these correlations are, following the specification in (2) we run multivariate regressions increasingly controlling for variables in Annex I.

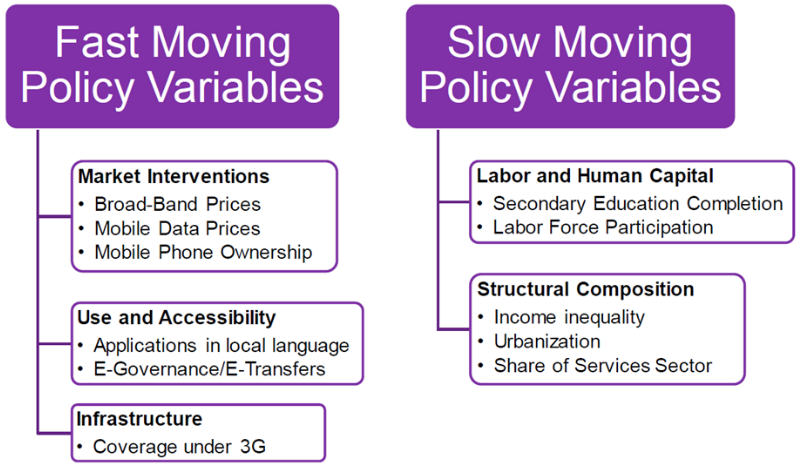
$$Internet\ Use_{ct} = \beta_0 + \beta_1 \mathbf{X}_{ct} + \beta_2 GDP\_PC_{ct} + \gamma_c + \epsilon_{ct} \qquad (2)$$

where β_1 denotes a vector of coefficients on a vector of determinants $\mathbf{X}_{ct}$, γ_c denotes country fixed effects, GDP_PC_{ct} is the GDP per capita in country *c* in year *t* and ϵ_{ct} is the error term clustered at the country level. We find that the direction of correlations between the internet use and its various determinants remain robust to multiple specifications, although, the magnitudes of coefficient estimates are still likely biased due to endogeneity in explanatory variables. Columns 1–4 in Table 1 present regression results for baseline variables that include current ICT investments, existing coverage, fixed and variable costs of using the internet, mobile ownership, and existence of complementary infrastructure like electricity. Controlling for GDP per capita and using only within country variation in the data, internet use is positively correlated with existing ICT coverage, ICT investments and mobile phone ownership and negatively correlated with variable costs (price of broadband and mobile data). Surprisingly, handset prices (a proxy for fixed costs) and electricity access are not statistically significantly correlated with internet use after conditioning for other baseline variables.

In columns 5–8 of Table 1, we find this pattern to be robust to the inclusion of more determinants. Moreover, secondary education completion rate appears to be more important than primary education completion rate and internet use increases with the number of mobile applications in national languages, suggesting the importance of higher skills and usability. Further, internet use increases with the share of the services sector in the economy; and, consequently, falls with higher agricultural and manufacturing activity. Finally, column 9 reports a full specification using all the potential determinants from appendix Annex I. We see that in addition to previous variables remaining remarkably stable, we find that conditional on all previous variables, internet use is positively correlated with the share of income held by the poorest 20 percent of people (suggesting a role for redistribution), labor force participation, and UN E-Governance score which captures the regulatory environment in an economy as well as the level of e-governance.

With these correlations in hand, we try to create a short list of most powerful determinants of digital adoption (after conditioning for GDP per capita and country fixed effects). We use the Least Absolute Shrinkage and Selection Operator (LASSO) estimator from Machine Learning (Chernozhukov and others 2021, Ahrens and others 2020) that adds a penalty term—the sum of the absolute values of the regression coefficients—to the usual least-squares minimization problem. LASSO drops variables that fail to contribute much to predicting the dependent variable (see Annex IV for more details on the methodology). Columns 10 and 11 in Table 1 present regression results from the regression specification in (2) using only the variables that the LASSO estimator selected to be most predictive. Column 11 has fewer variables because of a stricter penalty than the LASSO estimation in column 10. Conditional on GDP per capita and using only within-country variation in data, several determinants survive the LASSO penalization, and all selected variables preserve their earlier direction of correlation with internet use. While these are statistically the strongest predictors of internet use, they are not necessarily the variables that can be immediately affected by policies. For example, incomes at the lowest bin of income distribution are strongly correlated with internet use but may be very hard to change with policy. With this caveat in mind, we categorize the selected variables from the column 10 into two groups: fast-moving (i.e., variables that can be affected in the short run through policy intervention) and *slow-moving* (i.e., variables that generally take a long time to move; see Figure 5). Such grouping allows us to identify those variables that governments could target using policy instruments.

Figure 5. Strongest Predictors of Digital Adoption



Fast-moving variables associated with the cost of internet use (broadband and mobile data prices and smart phone ownership), which are important determinants of digital adoption, can be affected through policy interventions. Specifically, governments can use fiscal policy to provide targeted subsidies to lower the cost of internet connection and hence encourage digital adoption. To this end, in Section III we estimate a model of demand for internet use at the country level to obtain credible estimates of price elasticities of demand and use the elasticities to compute the magnitude of subsidies to achieve a near-full internet adoption.

Whereas price subsidies are generally known to be distortionary, they might be helpful when pursuing certain social and economic objectives. As documented below, digital adoption can bring substantial social and economic dividends that outweigh the costs of subsidization.[5] Further, internet is a unique consumer good that has several advantageous technological features that might aid with designing subsidies to lower costs. First, targeting subsidies could be simpler. Since internet adoption is likely to happen via mobile internet mobile operators normally have information about the users that would allow them to establish eligibility. For example, mobile operators possess information on a user's length and type of subscription, the location, and the amount spent. Second, we expect positive spillovers from subsidizing internet adoption as adoption by everyone will likely induce subsequent adoption within their network.[6] Given these positive network externalities, the current market prices likely lead to suboptimal levels of internet adoption relative to the social optimum.

Country experiences provide some support for subsidizing internet adoption. In Colombia the internet subsidy program 'Vive Digital' let to an increase in internet adoption by poorer individuals (Hidalgo and Sovinsky 2022).

[5] Other popular policy tools include command and control policies where the government fixes the price of private provision. We believe that such command-and-control policies are often weak to bypass especially in presence of weak institutions as found in LIDCs and EMEs.

[6] Literature shows that that adoption of mobile technology depends on the take-up within each individual's network (Björkegren 2019).

Interestingly, as part of its digital access initiative, Colombia also instituted numerous centers which provided internet training in addition to access to hardware. More recently, the US Affordable Connectivity Program, which was adopted as part of Bipartisan Infrastructure Law in 2021, provides to eligible households a monthly discount on their internet bills in addition to significant infrastructure development towards expanding high-speed broadband. Other examples of internet subsidies include Costa Rica's 'Connected Homes' program and the UK's 'Gigabit Voucher Scheme.'

Table 1. Correlates of Internet Use

	Baseline regressions				Baseline + controls					Machine learning	
	1	2	3	4	5	6	7	8	9	10	11
GDP Per Capita $1,000	2.635***	2.685***	2.745***	2.398***	2.298***	2.359***	2.389***	2.176***	1.750***	2.120***	2.438***
	[0.689]	[0.693]	[0.697]	[0.670]	[0.624]	[0.625]	[0.626]	[0.622]	[0.538]	[0.519]	[0.592]
Population Covered 3G (%)	0.192***	0.210***	0.202***	0.180***	0.138***	0.150***	0.145***	0.134***	0.118***	0.120***	0.120***
	[0.029]	[0.030]	[0.031]	[0.029]	[0.026]	[0.029]	[0.029]	[0.026]	[0.024]	[0.026]	[0.028]
ICT Investment (% GDP)	0.296***	0.272***	0.289***	0.307***	0.233***	0.222***	0.227***	0.245***	0.080*		
	[0.083]	[0.074]	[0.075]	[0.079]	[0.043]	[0.047]	[0.041]	[0.040]	[0.046]		
Log Fixed Broadband (PPP)	-3.546***			-3.182***	-2.460***			-2.312***	-1.597**	-1.498**	
	[0.890]			[0.826]	[0.686]			[0.650]	[0.616]	[0.603]	
Log Mobile High Package (PPP)		-5.270***		-3.817***		-3.425***		-2.756**	-2.130*	-1.854**	
		[1.211]		[1.309]		[0.954]		[1.239]	[1.143]	[0.776]	
Log Mobile Low Package (PPP)			-5.165***	-1.558			-3.047***	-0.652	-0.163		
			[1.499]	[1.752]			[1.072]	[1.500]	[1.432]		
Handset Price (% Monthly GDP)	0.099*	0.074	0.086	0.083	0.084*	0.072	0.078*	0.076*	0.059*		
	[0.055]	[0.057]	[0.058]	[0.053]	[0.044]	[0.045]	[0.045]	[0.042]	[0.035]		
Mobile Ownership (% Pop)	0.365***	0.400***	0.430***	0.348***	0.168***	0.194***	0.210***	0.168***	0.135***	0.144***	0.179***
	[0.062]	[0.062]	[0.064]	[0.058]	[0.042]	[0.042]	[0.044]	[0.040]	[0.037]	[0.040]	[0.043]
Electricity Access (%)	0.054	0.119*	0.104	0.028	0.06	0.104	0.094	0.046	-0.119*		
	[0.070]	[0.068]	[0.067]	[0.068]	[0.062]	[0.063]	[0.062]	[0.061]	[0.061]		
Education (At Least Primary)					0.109*	0.115*	0.107*	0.114*	0.097		
					[0.062]	[0.063]	[0.064]	[0.061]	[0.060]		
Education (At Least Secondary)					0.171***	0.146**	0.147**	0.162**	0.088	0.183***	
					[0.062]	[0.063]	[0.062]	[0.062]	[0.054]	[0.052]	
Apps in National Language					0.513***	0.509***	0.523***	0.478***	0.289***	0.356***	0.421***
					[0.059]	[0.058]	[0.059]	[0.058]	[0.061]	[0.061]	[0.062]
Services Value Added					0.078	0.033	0.062	0.065	-0.014	0.221**	
					[0.159]	[0.159]	[0.166]	[0.155]	[0.139]	[0.092]	
Agriculture Value Added					-0.453**	-0.519***	-0.508**	-0.446**	-0.255		
					[0.193]	[0.196]	[0.206]	[0.185]	[0.165]		
Industry Value Added					-0.415**	-0.503***	-0.455**	-0.437**	-0.318*		
					[0.187]	[0.187]	[0.199]	[0.180]	[0.162]		
Poverty at $1.9									-0.058		
									[0.055]		
Inequality									-0.147		
									[0.146]		
Lowest 20% Income									1.676***	2.102***	
									[0.599]	[0.442]	
Rural Population									-1.211***	-1.255***	-1.576***
									[0.223]	[0.230]	[0.281]
Gender Equality									0.111*		
									[0.059]		
Labor Force Participation Rate									0.696**		
									[0.292]		
Regulatory Environment									0.614		
									[1.156]		
UN E-Gov Score									0.120***	0.140***	0.182***
									[0.030]	[0.034]	[0.038]
Population Density (100 per sq km)									1.276		
									[1.434]		
R^2	0.765	0.764	0.76	0.78	0.834	0.832	0.83	0.839	0.866	0.852	0.83
λ										5370	7466
Number of Countries	169	169	169	169	169	169	169	169	169	169	169
Country FE	X	X	X	X	X	X	X	X	X	X	X
Observations	4124	4125	4125	4124	4124	4125	4125	4124	4124	4124	4125

Note: Internet users are individuals who have used the Internet (from any location) in the last three months. The Internet can be used via a computer, mobile phone, personal digital assistant, games machine, digital TV etc. Within R^2 reported.

III. A Model of Demand for Internet Use

The set of countries in the world is denoted by C and is further divided into three types $b \in \{L,E,A\}$ where $\{L,E,A\}$ stand for different income groups LIDC, EME and AE, respectively. Set of consumers in each country $c \in C$ is denoted by I. A market is denoted by a country-year pair. Consumer $i \in I$ in country $c \in C$ can either choose to use the internet or not. Normalizing utility from not using the internet to zero, consumer i in country c in year t derives the following utility from using the internet:

$$U_{ict} = \delta_{ict} - \alpha_i P_{ct} \tag{3}$$

$$\text{with } \delta_{ict} = \omega_c + \gamma_t + \xi_{ct} + \epsilon_{ict} \tag{4}$$

where δ_{ict} is the individual's utility from using the internet: the individual derives disutility from paying for internet and α_i is individual's sensitivity to prices in the market P_{ct}. We parametrize mean utility δ_{ict} using a country fixed effect that captures the average utility from using the internet in a country, year fixed effect which captures the level of technology or quality improvements over time, a market specific demand shock ξ_{ct}, and an individual-level independent and identically distributed (i.i.d.) error term that follows an extreme value type I distribution. Further, we specify price sensitivities differently for the three income groups. This gives us the following utility specification for internet users:

$$U_{ict} = \omega_c + \gamma_t - \alpha_b P_{ct} + \xi_{ct} + \epsilon_{ict}$$

Using the distribution of the individual level error term, we get the probability of an individual using the internet in country c, income group b and year t:

$$\Pi_{cbt} = \frac{\exp(\omega_c + \gamma_t - \alpha_b P_{ct} + \ \xi_{ct})}{1 + \exp(\omega_c + \gamma_t - \alpha_b P_{ct} + \ \xi_{ct})} \tag{5}$$

Assuming that the share of individuals using the internet as a proxy for probabilities, equation 5 can be estimated using a maximum likelihood estimation to recover price sensitivities for each income group. However, one might be concerned that the unobserved demand shock ξ_{ct} is correlated with prices in the corresponding market P_{ct} giving rise to endogeneity concerns. We get around this issue using a control function approach where we create a variable that shifts supply but not the demand. We use a Hausman-style supply shifter where we take leave-one-out averages of prices in other countries in the region of country c in year t denoted P_{-ct}. If supply of internet is spatially correlated, we should expect a correlation between prices in a country c in year t, P_{ct} and the supply shifter P_{-ct} but the supply shifter should not directly be able to affect demand in country *c*. We make an identifying assumption that this holds and conduct robustness checks using lagged value of our supply shifter for each country-year pair.

Our control function approach takes two steps. We first estimate the following equation:

$$P_{ct} = \beta_0 + \beta_1 P_{-ct} + \eta_{ct}$$

where η_{ct} is the market level error term. Next, we use the predicted value of our error term, $\hat{\eta}_{ct}$, as our control function (CF) in equation 5

$$\Pi_{cbt} = \frac{\exp(\omega_c + \gamma_t - \alpha_b P_{ct} + CF + \ \xi_{ct})}{1 + \exp(\omega_c + \gamma_t - \alpha_b P_{ct} + CF + \ \xi_{ct})} \qquad (6)$$

By control function CF we absorb all the variation in P_{ct} that could not be explained by our exogenous supply shifter, thereby effectively controlling for the correlation between the unobserved demand shock ξ_{ct} and price P_{ct}. We estimate (6) using maximum likelihood to identify price sensitivities separately for each income group. Table 2 shows the results from this estimation and Table 3 presents implied average demand elasticities using the robustness result with two lags of the IV.

Table 2. Demand Estimation

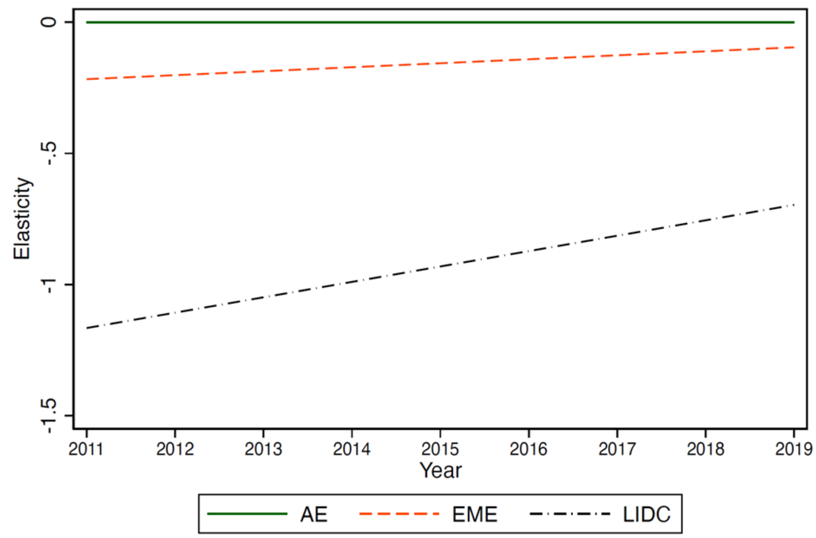
	Baseline		Robustness	
	Logit	Logit-IV	IV Lag1	IV Lag2
	1	2	3	4
α_{AE}	0.04***	0.015	0.009	-0.002
	[0.01]	[0.013]	[0.015]	[0.017]
α_{EME}	-0.011	-0.034***	-0.038***	-0.047***
	[0.008]	[0.011]	[0.013]	[0.015]
α_{LIDC}	-0.004***	-0.023***	-0.028***	-0.037***
	[0.001]	[0.006]	[0.008]	[0.011]
1st St. F-stat		409	418	428

Table 3. Estimates of Demand Elasticities

	AE	EME	LIDC
Elasticity	-0.004	-0.26***	-1.22***

Average demand elasticity in LIDCs is found to be much higher than EMEs while demand in AE is almost non-sensitive to prices (in line with Internet as a necessary item). Combining average price responsiveness from Table 2 with yearly data on internet use and prices we find that analogously to Moore’s law, demand elasticity decreases with income levels and over time (Figure 6).

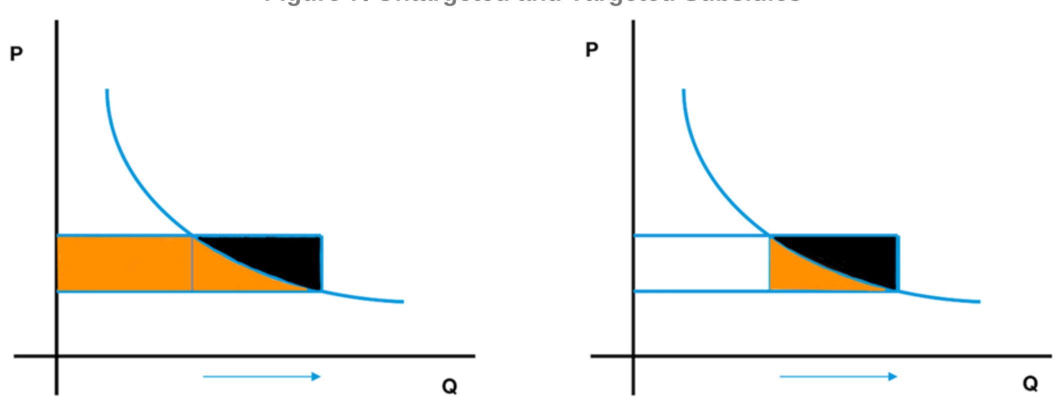
Figure 6. Trends in Demand Elasticities



Source: Authors' calculations.

These estimates of price elasticities of demand can inform the size of subsidies required to move demand to higher levels in LIDCs and EMEs. Figure 7 motivates our subsidy calculations. To increase demand, as depicted by the arrow, untargeted subsidies will support every individual until the desired level of digital adoption is reached. This will require public expense equal to the area of the shaded region (left graph). The graph on the right shows the case of perfect targeting, i.e., only individuals who do not yet have internet subscription, are given subsidies. Since some individuals value internet less than the price, subsidies inevitably induce a deadweight loss (black part of the shaded region). Targeting helps reduce public expenditure but requires incurring some additional administrative cost.[7] Section IV provides guidance on targeting subsidies.

Figure 7. Untargeted and Targeted Subsidies

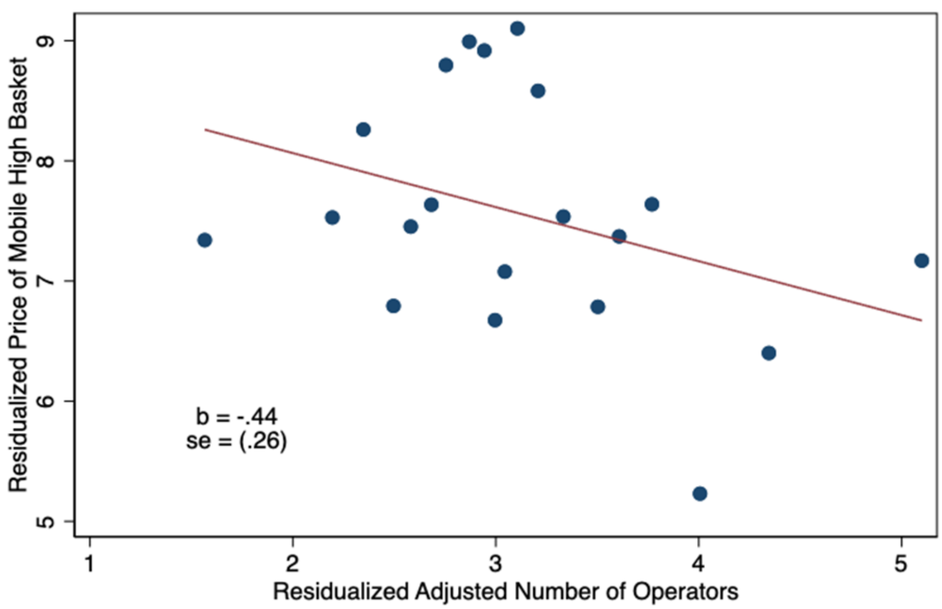


[7] Using the definitions of different kinds of price discrimination (Varian 1998), untargeted subsidies are analogous to no price discrimination (i.e., one price for all), while (our version of) targeted subsidies are analogous to the third-degree price discrimination (i.e., subsidized price for non-adopters) and perfectly targeted subsidies (i.e., prices equal to the willingness to pay) are analogous to first-degree price discrimination with government subsidies making up for the difference between the willingness to pay (i.e., the demand curve) and the actual price (so there is no deadweight loss).

In terms of prices, increasing internet use for EMEs from 72 to 90 percent (the current level in AEs) will require a 96 percent subsidy (18/72/0.26) and for LIDCs from 32 to 70 percent (the current level in EMEs) a 97 percent subsidy (38/32/1.22). If untargeted, other things equal, these amount to 0.34 percent and 3.2 percent of annual GDP for EMEs and LIDCs, respectively.[8] If targeted, the corresponding amounts decline to 0.07 percent and 1¾ percent of annual GDP for EMEs and LIDCs, respectively.[9]

So far, we have been agnostic about the industrial conduct, and it might be possible that a large subsidy program can induce a supply-side equilibrium response which, in turn, can either increase or reduce internet prices and quality. While a rigorous evaluation of the market structure in each country is beyond the scope of this study regressing the number of operators on price of mobile internet (while controlling for country income-group fixed effects and the logarithm of population) suggests a negative but weak relationship (Figure 8). This implies that market concentration likely affects equilibrium outcomes in the case of mobile internet. This may suggest that a price subsidy may lead to a price response from the suppliers especially in highly concentrated markets. For this reason, we shall proceed conservatively in our analysis to err on the side of caution.[10]

Figure 8. Relationship Between Number of Operators and Mobile Data Prices



Sources: OpenCelliD (2020) and Authors' estimates.
Note: Regression of price of mobile high basket on number of operators, income group dummies and log population.

[8] These calculations use an average monthly price of mobile internet as $4 in EMEs and $16 in LIDCs measured in terms of PPP (ITU). Population numbers are averages over income group from year 2020: 62.5 million in EMEs and 29.7 million in LIDCs. GDP PPP in constant 2017 dollars is $122 billion for an average LIDC and $827 billion for an average EME. To calculate subsidies, we take the product of number of people that will be given subsidies, the level of subsidy in percentage terms, and twelve times the monthly cost of using mobile internet.

[9] It is worth noting that if other determinants from figure 5 improve, subsidy requirements will go down.

[10] A point to note here is that prices in LIDCs are still very high and further research is needed to examine whether this is a result mainly of anti-competitive practices or high fixed costs.

IV. Digital Divide

In this section we provide some evidence on who is on the other side of the digital divide and who could be targeted through government subsidies. The analysis uses micro-data from the World Bank's FINDEX survey of over 150 countries in the year 2017.[11] The data samples approximately 1,000 individuals in each of the 150 countries and provides sampling weights. Table 4 reports descriptive statistics from this dataset. On average, women are over-represented for countries of all income levels. Broadly reflecting their demographic picture, LIDCs have the largest fraction of individuals below 20 years of age and smallest fraction of individuals over 65 years of age. This pattern switches for AEs. Education levels in LIDCs are unsurprisingly the lowest. Interestingly, simple averages in this table provide clear insights on the digital divide across countries. We can see that LIDCs have lower mobile phone ownership and debit card owners than EMEs and AEs. Further, LIDCs are behind on various uses of the internet: internet banking, online bill payments, e-commerce, and ability to get emergency funds, providing further evidence of a digital divide across countries.

Despite the digital divide, adoption of digital technologies is promising in multiple avenues. Sending and receiving domestic remittances are highest in LIDCs and almost half of these transactions happen via phones, suggesting ease of transacting on mobile phones. We suspect that sending/receiving international remittances might be even easier than domestic remittances. Moreover, about 16 percent of government transfers in LIDCs are received via phones higher than EMEs and AEs. Overall, despite the cross-country digital divide, digital adoption has a real potential in LIDCs at least in the financial aspects of peoples' lives.

We estimate effects of different demographic and socio-economic characteristics on digital adoption. Due to limited data, we use mobile phone ownership and measures for select internet-based activities to proxy for overall level of digital adoption.[12] Table 5 presents regression results from using the specification below:

$$Owns\ Mobile_{ict} = \beta_0 + \beta_1 \mathbf{X}_{ict} + \gamma_c + \delta_t + \epsilon_{ict} \tag{7}$$

where *Owns Mobile*$_{ict}$ is a dummy that takes a value equal to 1 if individual *i* in country *c* and year *t* owns a mobile phone, $\mathbf{X}_{ict}$ is a vector of individual level characteristics, and γ_c and δ_t are country and year fixed effects, respectively. The fixed effects help control for all time invariant variation at the level of a country and flexibly partial out average time trends. We cluster the standard errors at the country level to control for average correlation in error terms for individuals within a country.

[11] This data is also available for 2011 and 2014; however, these rounds did not ask the respondents about mobile phone ownership, a proxy for digital adoption.

[12] Unfortunately, we do not know whether the individual owns an internet enabled phone or not.

Strikingly, we find that—conditional on age, income, labor market outcomes—women, regardless of educational attainment, are less likely to own mobile phones than men in LIDCs. For example, less than primary educated women in LIDCs are 10 percent less likely to own a mobile phone than men with comparable education level. Even women with tertiary level of education remain 1.2 percent less likely to own a mobile phone than a less than primary educated men. Significant gaps in female mobile phone ownership persist in EMEs and AEs. Research has shown non-economic barriers to female mobile phone ownership. For example, Barboni and others (2018) provide several pieces of anecdotal evidence for why women do not own mobile phones in India, one of them being social norms against female phone ownership.

Further, mobile ownership monotonically increases with the level of education both for men and women. Another margin for within-country digital divide is age. We see that across the world, individuals over 65 years of age are less likely to own mobile phones than individuals below 20 years of age. Older people likely face high barriers to digital adoption potentially due to habit, skill, or ease of use of the internet, but are vulnerable to exclusion from several important services. Lastly and perhaps unsurprisingly, richer individuals and labor force participants have a higher likelihood of owning a mobile phone.

Finally, we use data on four purposes that individuals in the sample responded that they use internet for. The survey asks whether an individual paid bills online, shopped online, banked online, and received government transfers on a phone. Using each of these variables as our dependent variable, we run the regression specification in (7) separately for LIDCs, EMEs, and AEs. Results are reported in Table 6. Mobile ownership is strongly correlated with online activity. For example, mobile phone owners in LIDCs are 35 percent more likely than national average to pay bills online and approximately 40 percent more likely than national average to shop and bank online or receive government transfers using a phone. The remainder of the results are similar to mobile phone ownership with respect to gender, age, education, income, and labor force participation being important margins for digital divide.

Table 4. Summary Descriptive Statistics of Individuals Covered by FINDEX Survey

	LIDC	EME	AE
Female	0.514	0.506	0.518
	(0.500)	(0.500)	(0.500)
Age ≤ 20	0.218	0.137	0.086
	(0.413)	(0.344)	(0.280)
Age between 20 and 65	0.730	0.773	0.725
	(0.444)	(0.419)	(0.447)
Age ≥ 65	0.052	0.090	0.189
	(0.221)	(0.285)	(0.392)
Education(1/2/3)	1.420	1.769	2.060
	(0.568)	(0.662)	(0.624)
Owns Mobile Phone	0.660	0.847	0.935
	(0.474)	(0.360)	(0.247)
Has Debit Card	0.125	0.431	0.848
	(0.331)	(0.495)	(0.359)
Used Phone/Internet for Banking	0.062	0.181	0.588
	(0.241)	(0.385)	(0.492)
Online Bill Payments	0.054	0.150	0.559
	(0.225)	(0.357)	(0.497)
Bought Online	0.036	0.144	0.544
	(0.186)	(0.351)	(0.498)
Online Payments	0.011	0.066	0.070
	(0.106)	(0.249)	(0.256)
Can Get Emergency Funds	0.490	0.495	0.742
	(0.500)	(0.500)	(0.438)
Sent Domestic Remittance	0.226	0.145	0.042
	(0.418)	(0.352)	(0.201)
Sent Remittance by Phone	0.112	0.033	0.005
	(0.315)	(0.179)	(0.070)
Received Domestic Remittance	0.272	0.172	0.040
	(0.445)	(0.378)	(0.196)
If Recieved Remittance By Phone	0.127	0.035	0.002
	(0.333)	(0.185)	(0.048)
Paid Utility Bills by Phone	0.047	0.052	0.130
	(0.213)	(0.222)	(0.337)
Received Wages by Phone	0.025	0.026	0.027
	(0.157)	(0.159)	(0.163)
Received Government Transfers	0.066	0.132	0.277
	(0.248)	(0.338)	(0.448)
Received Government Transfers by Phone	0.011	0.008	0.012
	(0.103)	(0.088)	(0.111)
Received Agricultural Payments	0.282	0.080	0.008
	(0.450)	(0.271)	(0.091)
Received Agricultural Payments by Phone	0.026	0.004	0.000
	(0.160)	(0.063)	(0.022)
Received Self-Employment Payments	0.103	0.089	0.069
	(0.304)	(0.285)	(0.253)
Received Self-Employment Payments by Phone	0.014	0.006	0.005
	(0.117)	(0.078)	(0.069)
Number of Countries	51	67	40
Observations	42643	72120	34056

Note: Standard Deviations in parentheses. Weights used as given in the data. Variable Education has three levels (1, 2, 3). Data from 2017.

Table 5. Determinants of Mobile Phone Ownership

	Individual Owns a Mobile		
	LIDC	EME	AE
	1	2	3
Female	-0.090***	-0.062***	-0.062***
	[0.005]	[0.004]	[0.006]
Secondary Education	0.168***	0.083***	0.019***
	[0.007]	[0.004]	[0.005]
Tertiary Education	0.214***	0.091***	0.035***
	[0.016]	[0.006]	[0.006]
Female × Secondary Education	0.043***	0.042***	0.056***
	[0.009]	[0.005]	[0.007]
Female × Tertiary Education	0.077***	0.068***	0.062***
	[0.023]	[0.008]	[0.008]
Age ≤ 20	0.056***	0.186***	0.162***
	[0.011]	[0.006]	[0.005]
Age between 20 and 65	0.167***	0.208***	0.125***
	[0.010]	[0.005]	[0.004]
2nd Income Quintile	0.031***	0.027***	0.011***
	[0.007]	[0.004]	[0.004]
3rd Income Quintile	0.065***	0.044***	0.018***
	[0.007]	[0.004]	[0.004]
4th Income Quintile	0.108***	0.066***	0.026***
	[0.007]	[0.004]	[0.004]
5th Income Quintile	0.162***	0.089***	0.031***
	[0.007]	[0.004]	[0.004]
Received Government Transfers	-0.003	0.019***	0.005*
	[0.009]	[0.004]	[0.003]
In Workforce	0.069***	0.058***	0.033***
	[0.005]	[0.003]	[0.003]
R^2	0.181	0.161	0.104
Dependent Variable Mean	0.69	0.85	0.95
Country FE	X	X	X
Year FE	X	X	X
Number of Countries	49	64	39
Observations	41375	71756	33809

Note: The data is from 2017 since mobile ownership is only available for this year.

Table 6. Uses of ICT for Various Purposes

	LIDC				EME				AE			
	Using ICT for?											
	Bills	Shopping	Banking	Govt.	Bills	Shopping	Banking	Govt.	Bills	Shopping	Banking	Govt.
	1	2	3	4	5	6	7	8	9	10	11	12
Mobile Owner	0.017***	0.013***	0.032***	0.004***	0.016***	0.028***	0.044***	-0.001	0.166***	0.149***	0.167***	0.013***
	[0.002]	[0.002]	[0.003]	[0.001]	[0.003]	[0.003]	[0.004]	[0.001]	[0.010]	[0.010]	[0.010]	[0.002]
Female	-0.009***	-0.000	-0.005***	-0.002	0.001	0.012***	0.007***	-0.002***	-0.008	-0.015	-0.030***	-0.000
	[0.003]	[0.002]	[0.003]	[0.001]	[0.004]	[0.004]	[0.004]	[0.001]	[0.011]	[0.012]	[0.011]	[0.002]
Secondary Education	0.043***	0.036***	0.068***	0.002	0.069***	0.076***	0.103***	0.002	0.146***	0.140***	0.153***	0.036***
	[0.003]	[0.003]	[0.004]	[0.002]	[0.004]	[0.004]	[0.004]	[0.001]	[0.010]	[0.010]	[0.009]	[0.003]
Tertiary Education	0.118***	0.100***	0.212***	0.002	0.234***	0.241***	0.309***	0.001	0.282***	0.296***	0.283***	0.100***
	[0.007]	[0.006]	[0.008]	[0.004]	[0.006]	[0.006]	[0.006]	[0.002]	[0.011]	[0.012]	[0.011]	[0.006]
Female × Secondary Education	-0.021***	-0.013***	-0.026***	-0.000	-0.011***	-0.018***	-0.028***	0.001	-0.019	-0.002	0.015	-0.013***
	[0.004]	[0.004]	[0.005]	[0.002]	[0.005]	[0.005]	[0.005]	[0.001]	[0.013]	[0.013]	[0.013]	[0.004]
Female × Tertiary Education	-0.020***	-0.008	-0.033***	-0.004	-0.029***	-0.033***	-0.066***	0.002	-0.033***	-0.005	0.005	-0.008
	[0.011]	[0.009]	[0.012]	[0.005]	[0.007]	[0.007]	[0.008]	[0.002]	[0.015]	[0.015]	[0.015]	[0.009]
Age ≤ 20	0.014***	0.026***	-0.007	0.007***	0.072***	0.166***	0.059***	0.004***	-0.011	0.300***	0.100***	0.026***
	[0.005]	[0.004]	[0.006]	[0.002]	[0.005]	[0.005]	[0.005]	[0.001]	[0.010]	[0.010]	[0.010]	[0.004]
Age between 20 and 65	0.020***	0.011***	0.014***	0.004***	0.079***	0.095***	0.089***	0.004***	0.175***	0.255***	0.193***	0.011***
	[0.005]	[0.004]	[0.005]	[0.002]	[0.004]	[0.004]	[0.005]	[0.001]	[0.007]	[0.007]	[0.007]	[0.004]
2nd Income Quintile	0.006***	0.003	-0.004	0.002	0.014***	0.015***	0.023***	-0.001	0.050***	0.041***	0.036***	0.003
	[0.003]	[0.003]	[0.003]	[0.002]	[0.004]	[0.004]	[0.004]	[0.001]	[0.007]	[0.007]	[0.007]	[0.003]
3rd Income Quintile	0.014***	0.009***	0.008***	0.005***	0.029***	0.031***	0.044***	0.002***	0.093***	0.090***	0.075***	0.009***
	[0.003]	[0.003]	[0.003]	[0.002]	[0.004]	[0.004]	[0.004]	[0.001]	[0.007]	[0.007]	[0.007]	[0.003]
4th Income Quintile	0.025***	0.016***	0.023***	0.003***	0.055***	0.053***	0.071***	0.003***	0.118***	0.115***	0.098***	0.016***
	[0.003]	[0.003]	[0.004]	[0.002]	[0.004]	[0.004]	[0.004]	[0.001]	[0.007]	[0.008]	[0.007]	[0.003]
5th Income Quintile	0.042***	0.028***	0.054***	0.003***	0.100***	0.101***	0.121***	0.005***	0.163***	0.158***	0.134***	0.028***
	[0.003]	[0.003]	[0.004]	[0.002]	[0.004]	[0.004]	[0.004]	[0.001]	[0.008]	[0.008]	[0.007]	[0.003]
Received Government Transfers	0.051***	0.033***	0.045***	0.131***	0.017***	0.014***	0.036***	0.044***	0.030***	0.041***	0.021***	0.033***
	[0.004]	[0.003]	[0.005]	[0.002]	[0.004]	[0.004]	[0.004]	[0.001]	[0.005]	[0.005]	[0.005]	[0.003]
In Workforce	0.012***	0.009***	0.020***	0.002***	0.048***	0.037***	0.069***	0.001	0.114***	0.117***	0.138***	0.009***
	[0.002]	[0.002]	[0.003]	[0.001]	[0.003]	[0.003]	[0.003]	[0.001]	[0.006]	[0.006]	[0.006]	[0.002]
R^2	0.070	0.061	0.124	0.109	0.220	0.212	0.246	0.043	0.279	0.263	0.290	0.061
Dep. Var. Mean	.05	.03	.07	.01	.15	.14	.18	.01	.58	.55	.61	.03
Country FE	X	X	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X	X	X
Observations	41376	41376	41376	41376	75731	75731	75731	75731	33809	33809	33809	41376

V. Social Dividends

To calculate aggregate societal gains from digital adoption we consider three key economic outcomes—education quality, time use (as measured by the fraction of a day—percent of 24 hours—spent doing unpaid work), and labor force participation (LFP).[14] Data comes from the World Bank.[15] To get at causal effects of internet use on our economic outcomes, we use two econometric strategies—difference-in-differences (diff-in-diff) regression and diff-in-diff regression with an instrumental variable. To implement our strategies, we discretize our independent variable—internet use—into 5 bins *Q* = (0-20%), (20-40%), (40-60%), (60-80%), (80-100%) and estimate the following equation:

$$Y_{ct} = \beta_0 + \sum_{i\epsilon\{1,2,3,4,5\}} \beta_{1i}\, 1\{Internet\ Use_{ct} \epsilon Q_i\} + \beta_2 \mathbf{X}_{ct} + \gamma_{Region} + \tau_{year} + \epsilon_{ct} \quad (8)$$

where Y_{ct} is the economic outcome of interest, β_{1i} is the regression coefficient on a dummy variable if internet use in country-year *ct* belongs to bin Q_i, $\mathbf{X}_{ct}$ is a vector of country-year level controls that include the GDP per capita in the country along with all other slow-moving determinants of internet use from section II, and γ_{Region} and τ_{year} are *Region* and *Year* fixed effects, respectively. The *Region* fixed effect controls for all time invariant unobservables at the regional level and *year* fixed effects control for aggregates over time. We use robust standard errors to account for heterogeneity in errors.

The coefficients recovered from running the specification in (8) will have a causal interpretation only if the following parallel trends assumption holds: absence the treatment our dependent variables in the treatment and control units must not evolve differently. We present support for this assumption in Annex V. Ideally, we must verify that prior to internet adoption, countries in each region did not trend differentially. Due to a small sample size, we are unable to verify this. Alternatively, we verify parallel pre-trends for the two lowest bins of internet use (0-20%) *and* (20-40%) bunching our data into four time periods at roughly 5 years intervals from 1997 to 2021 and running the same regression specification as in (8). The parallel trends assumption holds for time use and labor force participation but not for education outcomes. For both primary and secondary education test scores, the regression coefficient for period 4 (2015-2021) is negative and statistically different from zero. This negative trend continues even in our instrumental variables approach below and is a limitation of our identification strategy.

[14] Note that labor force participation is also listed as one of the drivers of internet use. We believe the relationship between labor force participation and internet use goes both ways. Intuitively, labor force participation increases income and therefore likely affects internet use. In this section, we show the less intuitive direction of causality i.e., that internet use improves labor force participation using an instrument for internet use to isolate exogenous variation.

[15] Time use and labor force participation is available for women and men separately. We use these to disaggregate dividends for men and women. We use test scores in math, reading, and science for primary and secondary school students from the Harmonized Education Dataset to proxy quality of education.

For the second strategy with the instrumental variable, we use the same specification as in (8) but instrument for $Internet\ Use_{ct}$ using leave-one-out average of internet use in the region of country *c*. The identifying assumption is that internet adoption happens in regional waves and average internet use in the region of country *c* leaving country *c* out should not affect, for instance, test scores in country *c* independently of internet use in country *c* (similar to an IV for democratization in Acemoglu and others 2019). Instead of creating five instrumental variables for each bin, we instead use a control function approach and control for the predicted error term from a “first stage” regression of internet use in the home country on the leave-one-out average internet use in the region. We find that internet use in a country is highly correlated with average internet use in the region.

Across panels A and B in Table 7, our analysis shows that there is a robust positive effect of internet use on secondary school test scores. Relative to internet use in the first bin, increasing internet use up to the fifth bin increases average secondary education test scores between 28 to 44 points. To put this effect in context, the average test score in our sample is 480 and the minimum is 300. Assuming that 300 is the lowest possible score, the effective average score in our sample is 180. Using this as a reference, an increase of 28 points yields a 16 percent increase in test scores over the effective baseline average of 180 points. Secondly, internet use increases LFP anywhere between 6 to 10 percentage points over an average of 68 percent as internet use increases from the first to fifth bin. This increase in labor force participation is mainly driven by internet use in higher bins and is larger for women than men. This is consistent with a recent study by Chiplunkar and Goldberg (2022) who find that the expansion of internet in India led to an increase in the likelihood of working in the services sector for both men and women. We do not find a robust effect on time use patterns, but the coefficients in panel A suggest that time spent doing unpaid work (e.g., household chores) decreases for women but remains unchanged for men.

Next, we conduct back-of-the-envelope calculations of monetized social dividends using our most conservative results from table 7, panel B. Table 8 presents estimates from the back-of-the-envelope calculations when we increase internet use in LIDCs and EMEs from the second and fourth bins to the fourth and fifth bins, respectively.

Table 7. Effects of Internet Use on Education Quality, Labor Force Participation and Unpaid Work

	Education		Labor Force Participation			Time Spent on Unpaid Work	
	Primary	Secondary	All	Female	Male	Female	Male
	1	2	3	4	5	6	7
Panel A: Difference-in-Differences							
Internet Use 20-40%	6.776	18.010**	0.946	1.047	0.501	-1.354*	0.012
	[10.314]	[7.019]	[0.868]	[1.172]	[0.788]	[0.757]	[0.625]
Internet Use 40-60%	12.276	26.768***	2.124*	2.764*	1.775*	-2.029*	-0.39
	[12.370]	[8.407]	[1.102]	[1.504]	[0.917]	[1.035]	[0.629]
Internet Use 60-80%	30.003	36.606***	5.876***	7.227***	3.980***	-1.965	-0.35
	[19.248]	[10.817]	[1.379]	[1.964]	[1.091]	[1.275]	[0.826]
Internet Use 80-100%	45.725	44.713***	10.184***	13.074***	6.362***	-2.754*	0.01
	[29.215]	[13.560]	[2.079]	[2.746]	[1.604]	[1.631]	[1.060]
GDP PC	-0.696**	-0.015	0.186**	0.166*	0.125**	0.009	0.056*
	[0.324]	[0.190]	[0.084]	[0.087]	[0.058]	[0.028]	[0.032]
Panel B: IV Difference-in-Differences							
Internet Use 20-40%	-2.523	12.936***	-0.3	-0.608	0.043	-0.444	0.144
	[7.666]	[4.073]	[0.643]	[0.818]	[0.619]	[0.755]	[0.639]
Internet Use 40-60%	-3.863	18.198***	-0.123	-0.205	0.954	-0.484	-0.165
	[11.689]	[6.592]	[0.793]	[0.974]	[0.690]	[0.941]	[0.619]
Internet Use 60-80%	7.132	24.416**	2.670***	2.949**	2.796***	0.181	-0.038
	[17.712]	[9.966]	[0.980]	[1.313]	[0.919]	[0.995]	[0.781]
Internet Use 80-100%	15.558	28.688*	5.955***	7.419***	4.797***	-0.014	0.409
	[20.220]	[14.496]	[1.389]	[1.862]	[1.243]	[1.318]	[0.876]
GDP Per Capita	-0.781**	-0.061	0.172**	0.147*	0.119**	0.019	0.057*
	[0.379]	[0.207]	[0.083]	[0.084]	[0.058]	[0.029]	[0.034]
Dep. Var. Mean	478.21	478.16	67.65	56.77	78	18.1	7.83
Region, Year FE	X	X	X	X	X	X	X
Slow-Mov. Cont.	X	X	X	X	X	X	X
Observations	1044	1680	4200	4100	4100	1091	1090

Table 8. Monetized Average Annual Social Dividends

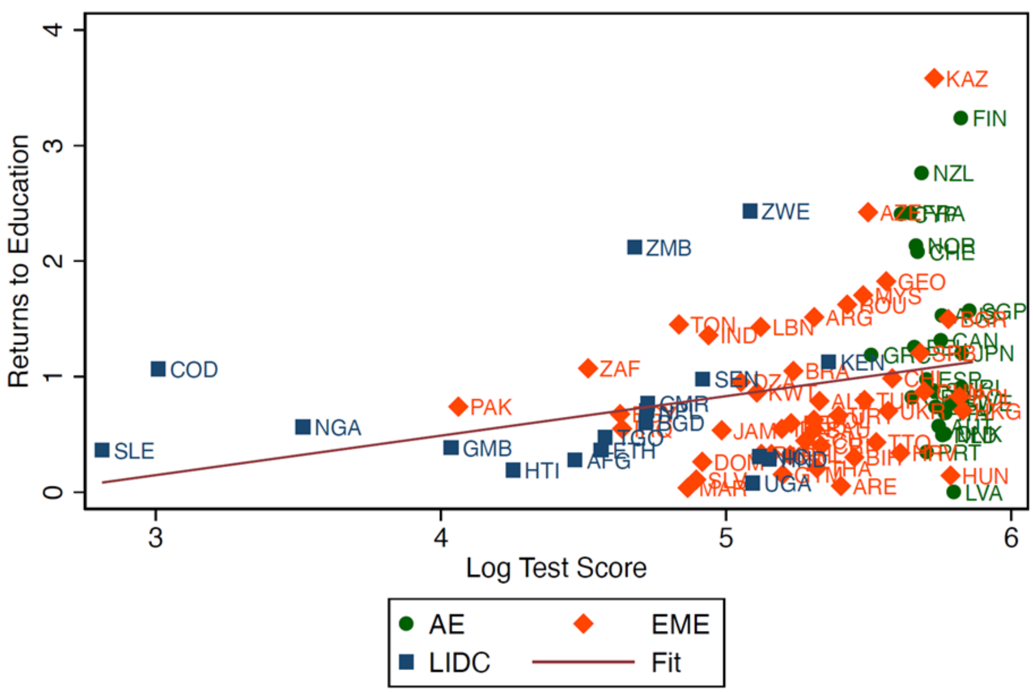
(in billions of PPP constant 2017 US dollars)

Income Group	Average GDP	Education Quality	Female LFP	Male LFP	Total LFP
EME	827	9.0	8.8	6.3	15.2
LIDC	122	2.8	0.5	0.4	1.0

We follow Schoellman (2012) to calculate monetized benefits from improved education scores (a proxy for education quality). Using data from IPUMS International, we first back out returns to education for male immigrants in the US by country of birth and education. Assuming that differences in returns to schooling for a given level of education represent differences in education quality in home countries, we construct a cross-country log-log relationship between estimated returns to schooling and most recent average test scores for countries in our sample (as shown on Figure 9) and find an elasticity of 0.34. Finally, we use the derived elasticity along with average education attainment from IPUMS international and returns to secondary test scores from internet use from panel B of Table 7 to estimate benefits. We find annual gains of 9 billion (1.1 percent of GDP) and 2.79 billion (2.3 percent of GDP) in PPP constant 2017 US dollars for an average EME and LIDC, respectively.

Gains from increased overall LFP are largely driven by women for EMEs as well as LIDCs and are large in magnitude. These gains are calculated in several steps. Using the World Bank's Compensation of Employees data, we first calculate the average annual wage of a labor force participant in the economy assuming equal wages for men and women. These are found to be $16,918 for EMEs and $3,093 for LIDCs in PPP constant 2017 US dollars. We then use the effect of moving a LIDC from the second bin of internet use to the fifth bin of internet use and an EME from the fourth bin to the fifth bin and calculate how many more workers will be added to the workforce. Multiplying the number of new labor force participants with average wage of a labor force participant, we find annual gains of $15 billion in PPP for an average EME or 1.8 percent of GDP PPP in constant 2017 US dollars and $0.95 billion in PPP for an average LIDC or 0.8 percent GDP PPP in constant 2017 US dollars. The increase in female labor force participation accounts for about 58 percent of these gains for EMEs and 57 percent of gains for LIDCs.

Figure 9. Returns to Education Quality



VI. Cost of Subsidies and Benefits from Internet Adoption

The cost-benefit analysis suggests that on average EMEs and LIDCs can expect net benefits from internet adoption. We present cost-benefit estimates for an increase in internet usage from 72 percent to 90 percent in EMEs and from 32 percent to 72 percent for LIDCs. The calculations are done for 81 EMEs and 51 LIDCs. Summing gains from education quality improvements and increased LFP, LIDCs and EMEs on average can expect annual gains of approximately $3.7 billion and $24 billion in PPP constant 2017 US dollars, respectively

(Table 9). Using a real discount rate of 5 percent, the present value of these benefits is $3.8 trillion and $39 trillion for LIDCs and EMEs, respectively. Taking a more conservative approach, we assume that returns start arriving only after 10 years of using the internet. In that case, the discounted value of returns becomes $2.3 trillion and $23.4 trillion in PPP constant 2017 US dollars for LIDCs and EMEs, respectively.

To calculate costs, we conservatively assume that the subsidies are offered for all future periods and are untargeted. In that case, the present discounted costs amount to $4.6 trillion and $4 trillion in PPP constant 2017 US dollars for EMEs and LIDCs, respectively. It appears that the case for internet subsidies is clear for EMEs. For LIDCs there is a case only for targeted subsidies which in terms of present value amount to $2.2 trillion in PPP constant 2017 US dollars.

Three points are worth noting related to the cost-benefit calculations discussed above. First, subsidies may not be required perpetually as other determinants of internet adoption evolve over time and drive internet adoption independently of prices. Second, internet adoption is likely to lead to network effects. Both these mechanisms can potentially reduce the costs of subsidizing the internet. Lastly, our calculations of subsidies use the current high price of $16, but this price is expected to come down steeply given the trends presented in Figure 3.

Table 9. Total Discounted Costs and Benefits in LIDCs and EMEs

(in trillions of constant PPP US dollars)

Income Group	Discount rate (%)	Present Discounted Value ($T = \infty$) in $ Trillion			
		Benefits		Costs	
		Immediate	10-year lag	Untargeted	Targeted
EME	5	39.1	23.4	4.6	0.9
LIDC	5	3.8	2.3	4.0	2.2

VII. Conclusions

This paper considers how to increase digital adoption by systematically answering three pertaining questions: (i) What are key determinants of digital adoption? (ii) What is the role of public policy in encouraging digital adoption and closing the digital divide? and (iii) what does the society gain from accelerating the pace of digital adoption?

We started with selecting the strongest predictors of internet use and categorized them as “Slow-Moving Policy Variables” and “Fast-Moving Policy Variables.” The former includes factors like prices and mobile phone ownership while the latter includes more structural variables that are usually difficult to move in the short run (e.g., labor market outcomes, human capital, inequality, sectoral composition, and level of urbanization). We note that these are statistically the most powerful predictors of internet adoption but may not be economically the most powerful tools. Using prices as a potential policy tool, we develop a structural model of demand for internet use and estimate demand elasticities using our aggregate data. These elasticities provide us estimates of subsidy requirements if the policy makers used prices to increase demand.

Since targeting subsidies requires much less public expenditure, we explored the composition of the digital divide. We found a significant digital divide across as well as within countries. Low-income countries are significantly behind emerging and advanced economies in terms of adoption. Within countries, we found a strong divide across gender, age, education, and labor force participation. These patterns hold stronger on LIDCs but nevertheless existed across the world.

We estimate the social value from internet adoption using three economically relevant variables: education quality, time use, and LFP. We find that not only do these variables improve with internet use but also women tend to benefit more in terms of better time use and higher labor force participation.[16] Reassuringly, the potential benefits on these three economic outcomes seem large and meaningful enough to warrant a policy push toward digital adoption. Since social dividends were calculated with country-level aggregate data, more rigorous econometric work using within-country data might help support these conclusions.

[16] It is worth noting that there are potential negative effects of excess internet usage like obesity and addiction that we have not accounted for in our analysis.

Annex I. Potential Determinants of Digital Adoption

Name of the Variable	Source	Unit
ICT Investment	World Bank	% GDP
Population Density	World Bank	100 per sq km
Electricity Access	World Bank	%
Log Fixed Broadband	ITU	$ PPP
Log Mobile High Package	ITU	$ PPP
Log Mobile Low Package	ITU	$ PPP
Industry Value Added	World Bank	% GDP
Services Value Added	World Bank	% GDP
Agriculture Value Added	World Bank	% GDP
GDP per Capita	World Bank	$1,000
Education Completion (Secondary)	World Bank	%
Education Completion (Primary)	World Bank	%
Women Empowerment	World Bank	%
Inequality	World Bank	GINI
Poverty at $1.9	World Bank	Head-Count Ratio ($1.9)
Rural Population	World Bank	%
Lowest 20% Income	World Bank	%
2nd Lowest 20% Income	World Bank	%
3rd Lowest 20% Income	World Bank	%
2nd Highest 20% Income	World Bank	%
Highest 20% Income	World Bank	%
Labor Force Participation Rate	World Bank	%
Population Covered 3G (%)	ITU	%
Handset Price	GSMA	% Monthly GDP
Gender Parity	GSMA	%
UN E-Government Score	UN	Score (1-100)
Apps in National Language(s)	GSMA	%
Mobile Ownership	GSMA	% Population

Annex II. Data Interpolation, Extrapolation, and Imputation

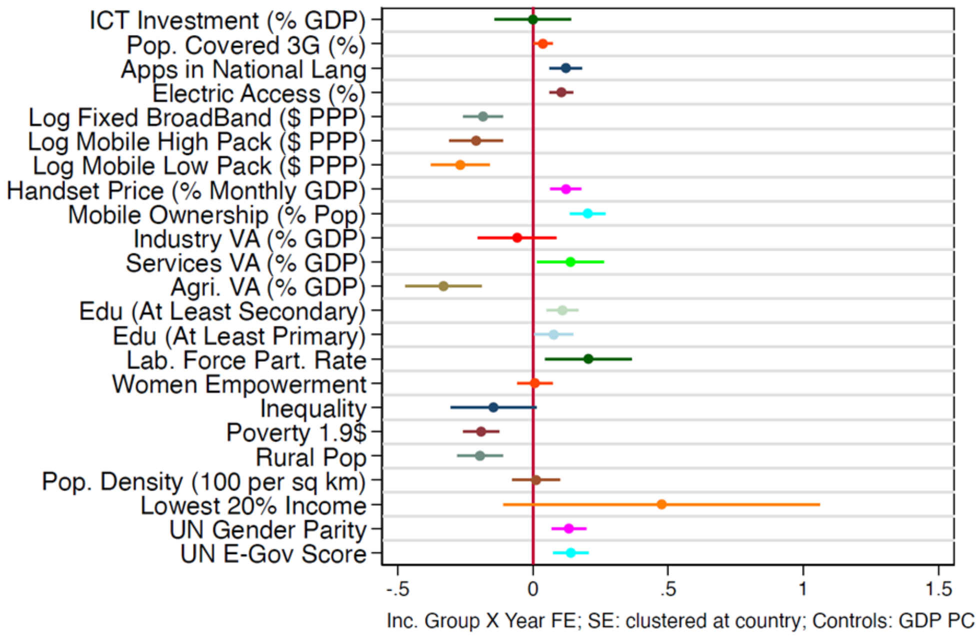
Most variables at the country level in World Bank's World Development Indicators are not available for all years of the study period 1997-2021. We fill in missing values of the variables using multiple methods. For each of our variables, we linearly interpolate and extrapolate using the gradient from two most recent years that we have the data for. For example, if a variable X has values in 2011, 2014 and 2017, we use data from 2014 and 2017 to linearly interpolate values for 2015 and 2016 and extrapolate linearly for 2018 and later. For missing data before 2014, we use information from 2011 and 2014 to interpolate for 2012 and 2013 and extrapolate for years before 2011. Naturally, some of the extrapolations generate negative values for our variables. All negative values have been replaced with missing values. Lastly, for variables that are measured in percentages, we replace all extrapolated values above 90 percent with missing values since it is unlikely that variables grow linearly at very high levels.

This process produces missing values which hinder running multivariate regressions in section 3. To deal with this, we replace missing values with average values of each variable and generate a corresponding dummy variable that take a value of one for values that have been imputed using means. In all our regression results, we control for these dummy variables to partial out all the imputed values.

Annex III. Correlates of Internet Adoption

Figure AIII.1 Correlates of Internet Use

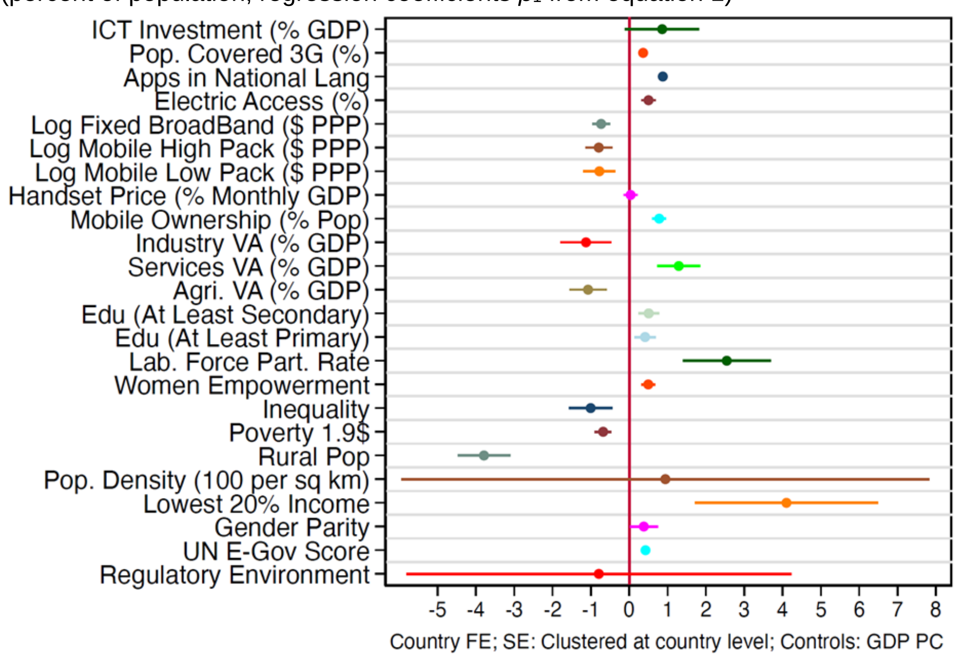
(percent of population; regression coefficients β_1 from equation 1)



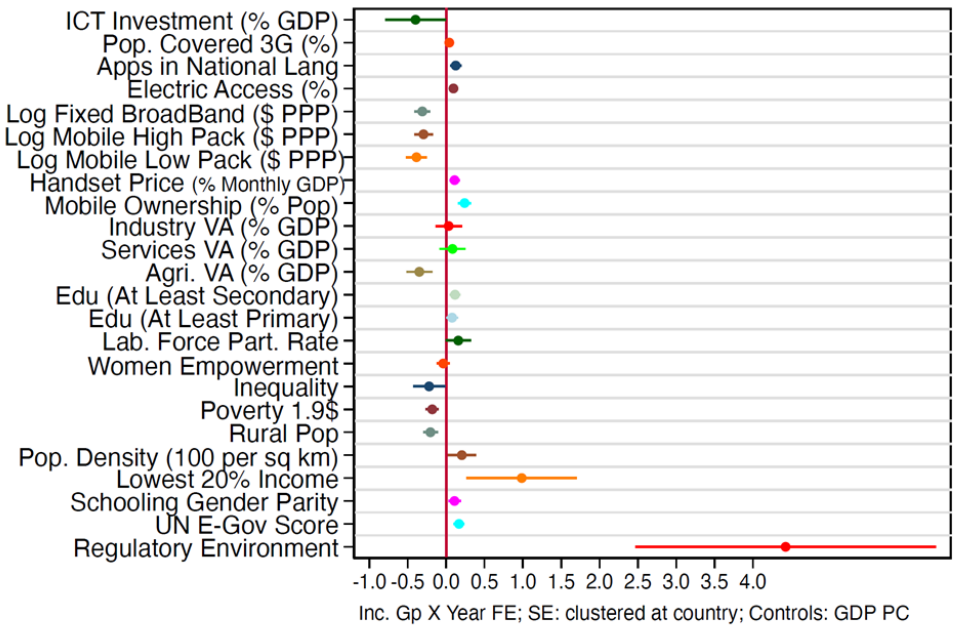
Note: Income Group x Year fixed effects and standard errors clustered at country level. Controls include GDP per capita.

Figure AIII.2 Correlates of Internet Access at Home

(percent of population; regression coefficients β_1 from equation 1)



Note: Country fixed effects and standard errors clustered at country level. Controls include GDP per capita.



Note: Income Group x Year fixed effects and standard errors clustered at country level. Controls include GDP per capita.

Annex IV. Details of the Lasso Procedures

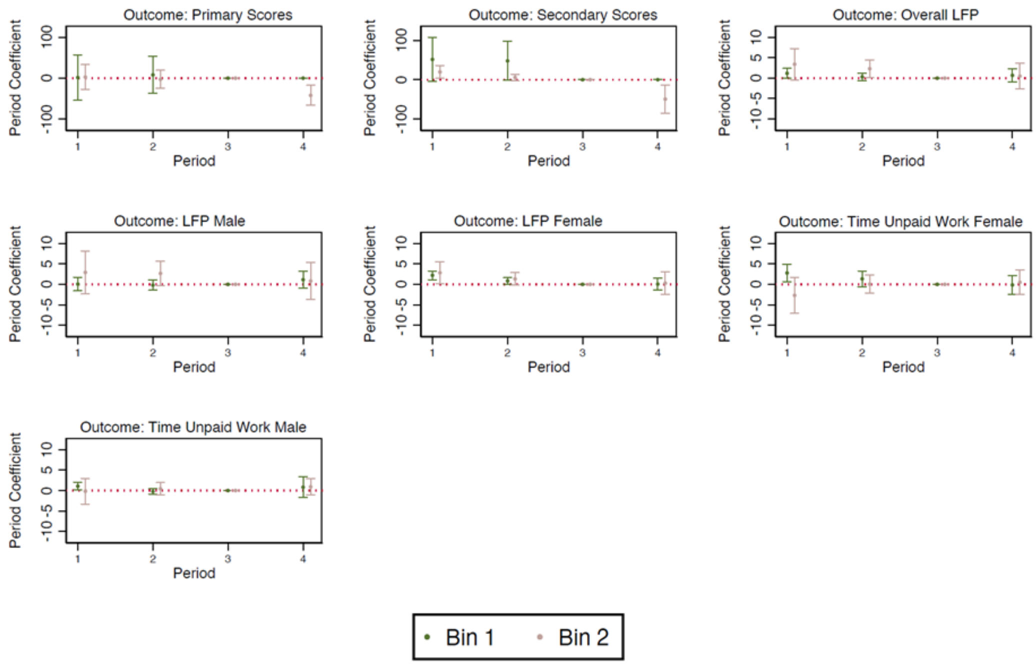
We use the program *rlasso* available in the STATA package 'LASSOPACK' for estimation (Chernozhukov et al. [2021], Ahrens et al. [2020]). The LASSO estimator $\hat{\beta}$ solves the following problem:

$$\min_{\beta} \frac{1}{N} RSS + \frac{\lambda}{N} \|\psi \times \beta\|_1 \tag{9}$$

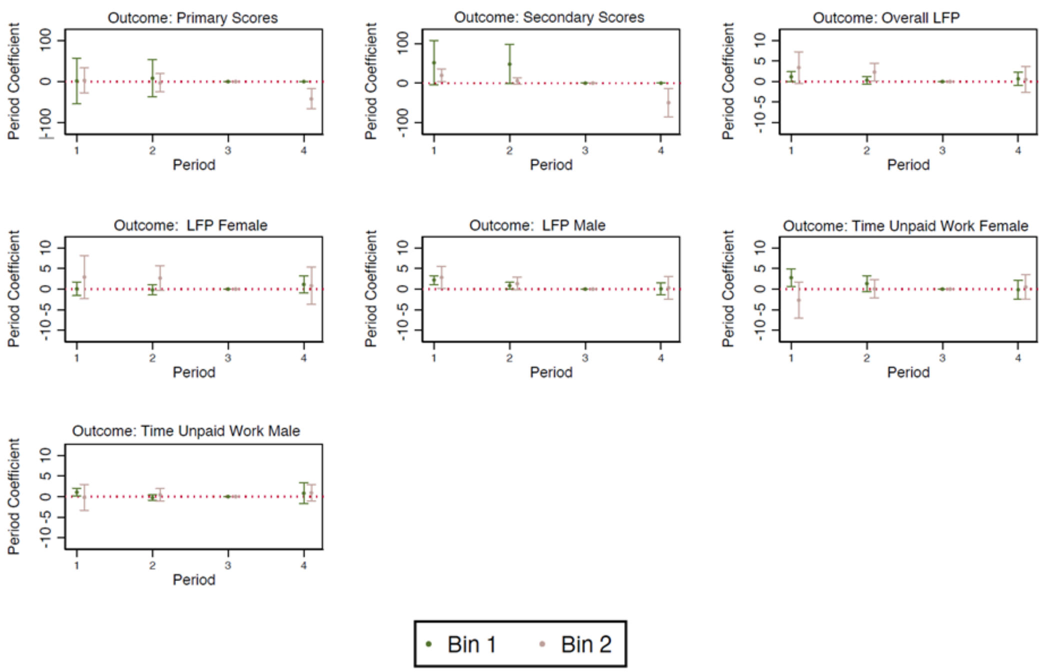
Where $RSS = \sum (y_i - x_i'\beta)^2$ denotes the residual sum of squares, β is the p-dimensional parameter vector, λ is the overall penalty level, $\|.\|_1$denotes the L1-norm, i.e., $\sum_i |a_i|$, ψ is a p by p diagonal matrix of predictor-specific penalty loadings (rLASSO treats ψ as a row vector), *N* is the number of observations. We partial out GDP per capita, country fixed effects, and all the dummy variables for imputes values corresponding to the included predicting variables. The default Bartlett kernel with bandwidth 11 (order$T^{\frac{1}{4}}$) has been used.

Annex V. Validating Parallel Trends Assumption

Parallel Trends—Simple Difference-in-Differences



Source: Authors' calculations



Source: Authors' calculations

References

Achim Ahrens, Christian B Hansen, and Mark Schaffer. "Lassopack: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation," 2020.

Ana C Dammert, Jose C Galdo, and Virgilio Galdo. "Preventing dengue through mobile phones: Evidence from a field experiment in Peru," *Journal of Health Economics*, 35:147– 161, 2014.

Ana C Dammert, Jose Galdo, and Virgilio Galdo. "Integrating mobile phone technologies into labor-market intermediation: A multi-treatment experimental design. *IZA Journal of Labor & Development*, 4(1):1–27, 2015.

Anders Akerman, Edwin Leuven, and Magne Mogstad. "Information frictions, internet, and the relationship between distance and trade," *American Economic Journal: Applied Economics*," 14(1):133–63, 2022.

Daniel Björkegren, Joshua Blumenstock, Omowunmi Folajimi-Senjobi, Jacqueline Mauro, and Suraj R Nair. "Instant loans can lift subjective well-being: A randomized evaluation of digital credit in Nigeria," *arXiv preprint arXiv:2202.13540*, 2022.

Daniel Björkegren. "The adoption of network goods: Evidence from the spread of mobile phones in Rwanda," *Review of Economic Studies*, 86(3):1033–1060, 2019.

Daron Acemoglu, Suresh Naidu, Pascual Restrepo, and James A Robinson. "Democracy does cause growth," *Journal of Political Economy*, 127(1):47–100, 2019.

Dean Karlan, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. "Getting to the top of mind: How reminders increase saving," *Management Science*, 62(12):3393–3411, 2016.

Dean Karlan, Melanie Morten, and Jonathan Zinman. "A personal touch: Text messaging for loan repayment. Technical report," National Bureau of Economic Research, 2012.

Eric Bettinger, Nina Cunha, Guilherme Lichand, and Ricardo Madeira. "What is it about communicating with parents?" *AEA RCT Registry*, 2020.

Felipe Kast, Stephan Meier, and Dina Pomeranz. "Under-savers anonymous: Evidence on self-help groups and peer pressure as a savings commitment device," National Bureau of Economic Research, 2012.

Gaurav Chiplunkar and Pinelopi Goldberg. "Employment effects of mobile internet in developing countries," National Bureau of Economic Research, 2022.

Giorgia Barboni, Erica Field, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore. "A tough call: Understanding barriers to and impacts of women's mobile phone adoption in India," *Boston, Massachusetts: Harvard Kennedy School*, 2018.

Hal R. Varian. "Price discrimination," *Handbook of Industrial Organization*, 1, pp.597–654, 1998.

Hunt Allcott, Matthew Gentzkow, and Lena Song. "Digital addiction," *American Economic Review*, 112(7):2424–63, 2022.

Jenny C Aker and Christopher Ksoll. "Can mobile phones improve agricultural outcomes? Evidence from a randomized experiment in Niger," *Food Policy*, 60:44–51, 2016.

Jenny C Aker and Isaac M Mbiti. "Mobile phones and economic development in Africa," *Journal of Economic Perspectives*, 24(3):207–32, 2010.

Jenny C Aker, Christopher Ksoll, and Travis J Lybbert. "Can mobile phones improve learning? Evidence from a field experiment in Niger," *American Economic Journal: Applied Economics*, 4(4):94–120, 2012.

Johannes Haushofer and Jeremy Shapiro. "The short-term impact of unconditional cash transfers to the poor: experimental evidence from Kenya," *Quarterly Journal of Economics*, 131(4):1973–2042, 2016.

Joshua E Blumenstock, Nathan Eagle, and Marcel Fafchamps. "Airtime transfers and mobile communications: Evidence in the aftermath of natural disasters," *Journal of Development Economics*, 120:157–181, 2016.

Julian Hidalgo and Michelle Sovinsky. Internet (power) to the people: How to bridge the digital divide. Unpublished, 2022.

Karthik Muralidharan and Abhijeet Singh. "India's new national education policy: Evidence and challenges. *Science*," 372(6537):36–38, 2021.

Karthik Muralidharan, Abhijeet Singh, and Alejandro J Ganimian. "Disrupting education? Experimental evidence on technology-aided instruction in India," *American Economic Review*, 109(4):1426–60, 2019.

Karthik Muralidharan, Paul Niehaus, Sandip Sukhtankar, and Jeffrey Weaver. "Improving last-mile service delivery using phone-based monitoring," *American Economic Journal: Applied Economics*, 13(2):52–82, 2021.

Piet Buys, Susmita Dasgupta, Timothy S Thomas, and David Wheeler. "Determinants of a digital divide in Sub-Saharan Africa: A spatial econometric analysis of cell phone coverage," *World Development*, 37(9):1494–1505, 2009.

Robert Jensen and Nolan H Miller. "Market integration, demand, and the growth of firms: Evidence from a natural experiment in India," *American Economic Review*, 108(12):3583– 3625, 2018.

Robert Jensen. "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector," *Quarterly Journal of Economics*, 122(3):879–924, 2007.

Todd Schoellman. "Education quality and development accounting," *Review of Economic Studies*, 79(1):388–417, 2012.

Valerie L Flax, Mekebeb Negerie, Alawiyatu Usman Ibrahim, Sheila Leatherman, Eric J Daza, and Margaret E Bentley. "Integrating group counseling, cell phone messaging, and participant-generated songs and dramas into a microcredit program increases Nigerian women's adherence to international breastfeeding recommendations," *Journal of Nutrition*, 144(7):1120–1124, 2014.

Victor Chernozhukov, Wolfgang Karl Härdle, Chen Huang, and Weining Wang. "Lasso-driven inference in time and space," *Annals of Statistics*, 49(3):1702–1735, 2021.

William Jack and Tavneet Suri. "Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution," *American Economic Review*, 104(1):183–223, 2014.

World Bank Group, "World development report 2016: Digital dividends," *World Bank Publications*, 2016.

PUBLICATIONS

Determinants and Social Dividends of Digital Adoption

Working Paper No. WP/2023/065