Internet (Power) to the People: How to Bridge the Digital Divide

Julian Hidalgo and Michelle Sovinsky*

August 3, 2023

Abstract

Due to global lockdowns driven by the COVID-19 pandemic, the pervasiveness of inequalities in digital access is more salient. As a result, finding strategies to narrow the digital divide has moved to the forefront of public policy. We examine the impact of a pricing subsidy in Colombia to learn about what hinders low-income populations from adopting internet services. Our model allows take-up to depend on the types of plans offered as well as the rate of diffusion in the neighborhood. We find that increasing the diffusion rate (via internet literacy programs) is more beneficial to takeup among households in the lowest socioeconomic markets and in less technically savvy markets relative to a pricing subsidy. Pricing subsidies reduce the digital divide by 13% points but over half of that is due to the impact on the diffusion rate. Our results suggest that non-price policies are equally important to bridge the digital divide.

JEL Classification: L15, L51, L86, D12, D31
Keywords: digital divide, network effects, internet diffusion, limited choice sets, data restrictions, discrete choice models, price subsidies, internet access

---

*Hidalgo is at KU Leuven. Sovinsky is at University of Mannheim and CEPR. We wish to thank Mike Riordian, Matt Backus, Andrea Pozzi, Martin Peitz, Juan Oviedo, Humberto Pintado, Frank Verboven, Jan De Loecker, Rosa Ferrer, and conference participants at Banco de la República (Cali), CEPR Applied IO workshop (Cambridge), DG Competition, EARIE (Athens), FCC, Helsinki, Instituto Tecnológico Autónomo de México (ITAM), MaCCI Summer Institute, Pontificia Universidad Católica de Chile, Stockholm University, University of Bristol, Universidade NOVA de Lisboa, and the University of Vienna for helpful comments. Sovinsky acknowledges support from the European Research Council Grant #725081 FORENSICS and from the German Research Foundation (DFG) through CRC TR 224 (Project A02). This paper was previously circulated under the title "Internet (Power) to the People: The Impact of Demand-side Subsidies in Colombia ". Hidalgo: Department of Economics, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium. Phone: +32 163-224-92 . E-mail:julianfelipe.hidalgorodriguez@kuleuven.be. Sovinsky: Department of Economics, University of Mannheim, L7 3-5, 68161 Mannheim, Germany. Phone: +49 621-181-3503. E-mail:msovinsky@econ.uni-mannheim.de

1 Introduction

The Internet has been the cornerstone of economic development during the last decades. It has structurally changed the way individuals, companies and institutions make decisions and has driven major advances in all sectors of the economy. In spite of this, almost half of world households are not yet connected to the internet - with the majority of those offline in poorer countries (UNESCO, 2019). As a result, the digitalization of the economy brings new inequality concerns arising from this “digital divide” - the uneven distribution of the potential benefits from the internet economy.

Due to global lockdowns driven by the COVID-19 pandemic, the pervasiveness of this societal divide is more salient. Access to the internet has a direct impact on education, timely health information, and resource distribution. UNICEF estimates that 463 million children (one out of three school-aged children) do not have adequate resources to participate in virtual schooling (UNICEF, 2020). Regarding the ability to work from home, in developed nations one in three jobs can be done from home, whereas that number is only one in 26 in low-income countries (World Bank, 2020). Finally, there is evidence that lacking internet access affects vaccine uptake for older and vulnerable groups (World Economic Forum, 2020) and even impacts the ability to social distance (Chiou and Tucker, 2020). To that end, UNESCO and the International Telecommunication Union have a joint target of connecting 75% of the world’s population to broadband by 2025. Finding effective solutions to bridge the divide is a prominent challenge for policymakers. However, evidence on the (in)effectiveness of interventions in developing countries is scarce, and the impact is likely to be structurally distinct from developed nations given their relatively high levels of digital exclusion (PEW Research Center, 2015).

We examine a pricing subsidy program to lower Internet adoption barriers, which was aimed at increasing broadband connection among low-socioeconomic (SES) groups in Colombia. Specifically, we combine information on the program and data reported by Internet Service Providers (ISP) (at the plan and SES level) together with technology diffusion variables to estimate adoption of internet services. A novelty of our approach is that we allow take-up to depend on the (unobserved and endogenous) rate of internet diffusion in the locality. We use the estimated parameters to evaluate the effectiveness of (counterfactual) policies at increasing adoption rates via price subsidies, neighborhood diffusion, and access to more choice among plans. We provide evidence for the effectiveness of pricing subsidies on decreasing the divide. However, our results also show that non-price policy tools aimed at increasing internet diffusion are more vital, particularly for the less-internet savvy consumers who tend to be the poorest among the low-SES groups.

There are a few important components that our model incorporates. First, the value of adopting may be impacted by the adoption rate in the market. For example, a higher degree of diffusion in the neighborhood could positively impact a family if it provides more opportunities to gain experience with the internet. On the other hand, if the adoption rate is very fast the household may decide not to adopt as they have more opportunities to access the internet via their neighbors. However, we do not observe the speed of internet diffusion in the market, and, furthermore, the diffusion rate is likely to be endogenous to the adoption decision. To overcome the first hurdle we use data on the number of plan subscribers, together with models of innovation diffusion, to estimate the rate of internet diffusion in the geographic-SES market. To overcome the second hurdle, we take advantage of data that impacts the social network but not the unobserved characteristics of the plans to instrument for the diffusion rate.

We should note that not explicitly accounting for the impact of neighborhood diffusion on adoption could generate an inconsistent estimate of the pricing subsidy effect. That is, if not modeled, at least part of the impact of neighborhood diffusion (which may be influenced by the pricing subsidy) will be allocated to unobserved heterogeneity. Our approach has the advantage that it allows us to separate the impact of the diffusion rate on adoption from the impact of product or consumer characteristics, allowing us to correct for the potential endogeneity of diffusion to obtain a causal impact of the pricing subsidy.

Second, low-SES consumers have limited choice of available plans because ISPs can (and do) offer different plans to SES groups in the same region. We take this choice restriction into account when we estimate the demand model. This allows us to gauge the impact on adoption that could arise from an increase in the set of available plans. Third, ISPs may strategically change their prices in response to the subsidy so we estimate ISP pricing decisions to allow for strategic price responses to any policies (actual or counterfactual) that are put into place. Finally, barriers to adoption may exist due to lack of information about the functionality and potential benefits of Internet services. We divide the market according to the level of adoption prior to the subsidy, which enables us to examine the policy impact within SES groups.

We find that the pricing subsidy reduced the digital divide by 13% points. However, most of the increase in adoption was driven by households in markets that had high internet penetration rates prior to the subsidy. Furthermore, the benefit was concentrated in households from the upper income of the low-SES households. Our results show that over half of the increase in adoption from the pricing subsidy was due to the impact on the diffusion rate. In contrast to a pricing subsidy, we find a policy targeted only at increasing the diffusion of services (via internet literacy plans) is more effective at closing the digital divide both in markets less connected initially and among households from the lowest-SES groups. So, while our findings suggest that pricing subsidies are effective in closing the digital divide, we find that non-price policies are equally important and most effective at reaching households who are the most vulnerable.

There is a rich empirical literature that examines the influence of social network effects on individual choice.[1] [2] The majority of the literature incorporates demand interactions by allowing a measure of aggregate choices to influence individual behavior. As discussed in Bhattacharya et al. (2023), this method of incorporating social effects has the drawback that it is often not possible to disentangle the impact a policy has on an individual separate from its impact on the group. We deviate from this literature by estimating the diffusion rate, which captures a social effect of adoption, rather than controlling for the social effect via a group adoption rate. Our approach has the advantage that we can decompose the impact of policies into the impact on demand directly and the impact on demand arising from the social network (via the diffusion rate). This allows us to conduct counterfactual

[1] For some examples, see Guiteras et al. (2019) and Grigolon and Lasio (2023). The theoretical literature is equally rich, with Arieli et al. (2020) being the closest related to our context.

[2] There is related literature on indirect network effects where consumers care about how many users there are of a complementary good (e.g., Rysman, 2004; Fan, 2013; Lee, 2013; Tucker, 2008; Ryan and Tucker, 2006; Ackerberg and Gowrisankaran, 2006; Björkegren, 2019.)

policy experiments targetting the diffusion rate directly, which we find is critical to connect the most vulnerable of the low-SES group.

Our work is related to a large literature that evaluates residential internet adoption policies (e.g., Hausman et al., 2001; Goolsbee and Klenow, 2006; Goldfarb and Prince, 2008; Nevo et al., 2016).[3] More specifically, we focus on broadband adoption in a developing country and, so, are related to work in development including Galperin and Ruzzier (2013) (for Latin American and Caribbean countries), Belloc et al. (2012) (for OECD countries), Jordán et al. (2013)(for Latin American countries), and Hjort and Poulsen (2019) (for African countries). Regarding Colombia, Vélez-Velásquez (2019) examines the impact of mergers in telecommunication on broadband provision, Vélez-Velásquez (2020) studies the implications of price discrimination on telecommunication services, Sovinsky and Hidalgo (2022) examines consumer switching behavior between plan speeds, and Hidalgo and Oviedo (2014) provide descriptive analysis of the impact of standards on download speed on internet provision.

Our paper combines the evaluation of public policies with the analysis of a vulnerable group of consumers. In this sense, it is related to the literature that studies the digital exclusion of low-income populations (Savage and Waldman, 2009; Greenstein and Prince, 2006; Ackerberg et al., 2014). Finally, our work is related to a growing literature using tools from structural industrial organization to examine issues in developing nations (e.g., Chaudhuri et al., 2006; Björkegren, 2019; Walsh, 2020).

This paper differs from the above-mentioned literature in a number of ways. First, we present a structural demand model of consumers' preferences for Internet adoption with respect to price and the speed of technology diffusion. Second, we apply this model to a developing nation to assess the effectiveness of demand-side interventions in ameliorating the digital divide. Third, we show how to estimate such a model under data limitations. Finally, we take into account the limited nature of the choice set facing each consumer in that not all consumers in the same geographic market have access to the same plans.

In the next section we provide an overview of the subsidy policy, discuss the data, and present reduced form evidence of the impact of the subsidy. We outline the empirical model and the estimation approach in Sections 3 and 4. We present parameter estimates in Section

[3] See Briglauer et al. (2015) and references therein.

5. We discuss counterfactual results that inform the impact of the subsidy and alternative policies on adoption in Section 6. Then we conclude.

2 Data

2.1 Subsidy Policy

From 2012 until 2015, the Colombian Ministry of Information and Telecommunication Technologies (MinTIC) implemented a subsidy scheme aimed to increase internet adoption among low-SES households. The subsidy took the form of a reduction in monthly fees for fixed-line internet plans with a broadband connection.[4] The discount varied by geographic department, which is a region with similar legal status as a state in the US.[5] There are 32 departments in Colombia which are in turn divided into approximately one thousand municipalities. Figure 1 shows the distribution of the subsidies across Colombia. The average monthly discount across departments was about $4 US dollars (ranging from about $2 to $7 US dollars), which is about 21% of the average monthly tariff.

The subsidy was only available to low-SES households. Household SES classification is determined by a geographical targeting instrument, known as the Socio-economic Stratification. The method assesses characteristics of residential properties (dwelling and neighborhood) and classifies them into six SES strata within municipalities. The strata were designed to measure amenities surrounding the dwelling, and so are highly correlated with income as richer individuals tend to live in areas with more amenities.[6] Strata 1 represents the lowest-income/SES group whereas strata 6 denotes the highest-income/SES group. According to Gran Encuesta Integrada de Hogares (GEIH), a survey of selected households from 13 cities and metropolitan areas, monthly income averages around 262 (real US)$ for strata

[4] Section 2 of article 58 of the Act 1450 of 2011. The government paid the ISP a subsidy based on subscriptions and households paid the discounted monthly rate. Broadband is defined as a connection with download speeds ≥ 1 Mbps and upload speed ≥ 0.5 Mbps (CRC resolution 2352 of 2010).

[5] The MinTIC determined the amount of subsidy according to the density of the location and the distance with respect to the regional hub. That is, the higher the cost of the last-mile connections, the higher the monthly fee reduction. MinTIC resolution 0000926 of 5 May 2014.

[6] The method has been used to target water, electricity and basic sanitation subsidy programs.

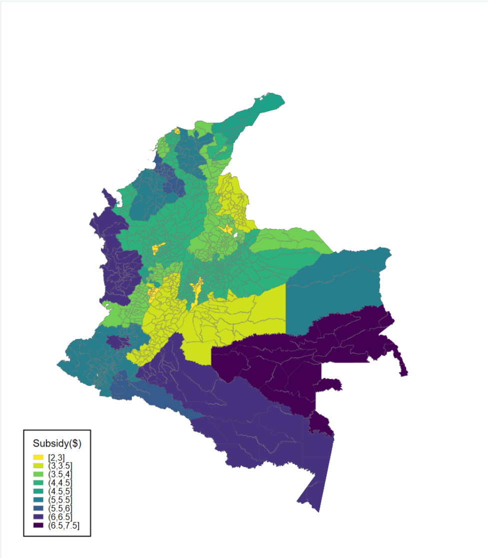

Figure 1: Subsidy Scheme

1 and around 316 (real US)$ for strata 2. Households in stratas 1 or 2 are considered the most vulnerable and, hence, are the focus of various subsidy policies including the subsidy we examine. Given the GEIH are only a subsample of the population, we use data on the number of households in the strata-municipality from the 2018 Census as the number of potential subscribers (i.e., the market size).[7]

2.2 Internet Plan Subscriptions

The Colombian Comisión de Regulación de Comunicaciones (CRC) provides information on services offered by all Internet Service Providers (ISP).[8] These data include (i) transmission speeds (i.e., download and upload speed); (ii) monthly service fee; (iii) type of Internet access technology, (iv) municipality and SES group to which the service was offered; (v) number of subscribers; and (vi) the ISP offering the service. We define a plan as a combination of ISP, upload speed, download speed, and technology.

[7] Census also reports projections until 2050 which we use to linearly interpolate the number of households for the quarters in our sample. We considered using the number of fixed telephone lines as the market size, but in 20% of the markets the number of adopters is greater than the number of fixed lines.

[8] The CRC is the Colombian analog of the US Federal Communications Commission.

ISPs offer subscription plans that vary across municipality and strata. That is, they treat each strata within a municipality as a separate market, and hence, our data also vary by strata within the municipality. Given that we are interested in a subsidy targeted at low-income households, we focus on residential clients in stratas 1 and 2 only.[9] We define a market as a municipality-strata combination.

The ISPs are required to report technical, commercial and financial data to the Telecommunications Information System. However, disaggregated information is not always required, and so the information provided by each ISP is limited - including at times missing information on plan monthly fees. One of the big components driving adoption decisions is the price of the plan (and it is what the subsidy targets) so we focus our analysis on the longest consecutive period for which we have data with plan prices: 2013:1-2014:4. However, we note that we also have prices on plans offered for part of 2011, which we use to provide reduced-form evidence of the impact of the subsidy and to test the fit of our model in section 6. After we drop some misclassified data and outliers our main sample for strata 1 and 2 residential clients consists of 44,515 observations from 2013:1-2014:4.[10] Table 1 provides an overview of residential plans offered to stratas 1 or 2. Notice, in some municipalities some ISPs do not offer a plan in strata 1, while in others some ISPs do not offer a plan in strata 2. We return to the issue of limited choice of plans in section 3.

In total, 35 ISPs offer plans to strata 1 or 2. Among those, 12 ISPs offered plans that were subsidized (i.e, broadband connections that were part of the subsidy program). As Table 2 shows, the major market players, in terms of market share, offered subsidized plans to both low-SES strata. The top 4 ISPs had a combined aggregate market share at the country-level of 89%. These top 4 providers have a national scope, except for ETB which primarily operates in the metropolitan area of the capital city. The remaining (other) ISPs serve specific regional markets. The remaining, smaller ISPs that offered subsidized plans served around 9% of the country, while those that did not grant subsidies have, on average,

[9] Residential clients constitute 89% of all subscribers across all strata. Our data cover about 90% of the population (which is spread across approximately 500 municipalities). About 400 of the municipalities have a very small population.

[10] We drop observations with download speed below 64kbps; below the 5th or above the 98th price percentiles. The bottom 5th price percentile prices are below $8 US dollars whereas the top percentile are prices above $200 US dollars (about two-thirds the average income).

				Total Markets	
Period	Plans	ISP	Munic.	Strata 1	Strata 2
2013:1	196	27	438	380	425
2013:2	198	26	452	389	439
2013:3	220	28	465	401	452
2013:4	202	25	457	400	444
2014:1	210	26	442	350	399
2014:2	218	28	481	363	438
2014:3	232	28	447	313	409
2014:4	227	29	527	423	506

Notes: Descriptive Statistics for services offered in Strata 1 or 2. A plan is defined as an ISP, technology, download speed and upload speed combination.

Table 1: Market Sample Description

	Subsidy Scheme	Number of Depart.	Country Share	Market Share
Colombia Telecom	1	28	28.19	79.67
UNE epm	1	19	19.95	39.6
ETB	1	5	11.81	27.77
Telmex	1	23	29	36.79
Other ISP Subsidy (8)	1	2	9.02	37.34
Other ISP No Subsidy (23)	0	4	2.04	25.09

Notes: Country share is the average share of subscribers over all markets. The market share is the average share of subscribers across markets (average within markets).

Table 2: Market Shares of ISPs

the lowest share of consumers.

Table 3 shows descriptive statistics for the internet services available to low-income strata broken down by year. The top panel presents plan-level statistics and the bottom panel presents market-level statistics. As panel A shows, there were 315 (305) unique plans offered in 2013 (2014) at an average (subsidized) price of around $20.

Most of the services were provided through xDSL (53%) and Cable (27%) technologies. Early deployment of wired internet infrastructure, using voice-grade copper telephone wires, is correlated with geographic coverage. This is consistent with the prevailing presence of the technology xDSL, which is provided in nearly 500 municipalities.[11]

Around 40% of the connections had download speeds of between 1-4 Megabits per second

[11] We aggregate the technologies into xDSL, Cable, and Other, which includes WiFi, WiMAX, microwave radio transmission, satellite and fiber-optic based technologies.

(Mbps) (Mid-speed) in 2013, and this increased to over 45% the following year. Likewise, the amount of high-speed connections (more than 4 Mbps) increased from 34% to 36% in the same period.[12] This is also reflected in the increase in the proportion of broadband connections from 2013 to 2014 (from 74% to 82%).

How established the firm is may impact consumers' perception of the quality of the product. For example, more established firms are more likely to have their own high-speed lines, which allows them to provide better services. In addition, more established firms tend to invest in advertising, which is broadcast on national television, and consumers may interpret this as a signal of good service. To allow for both potential impacts on demand we include variables for the amount of time the firm has been in operation, where we allow the impact to differ if the firm has been in operation for a period of at least 4 years (i.e., an established firm). The variable ISP seniority shows the number of quarters that the ISP has been operating in the municipality since 2010, which was under 4 years on average in 2013.

Panel B of Table 3 reports descriptive statistics at the market-level.[13] For each market, there were on average four plans offered by two ISPs. The Herfindahl measure of concentration of subscribers among ISPs (HHI ISP-subscribers) indicates that the level of concentration of subscribers is relatively high over time and hypothetically one ISP hoards most of the demand in each market. The index HHI ISP-plans measures the distribution of Internet plans among ISPs and shows that the variety of plans is highly concentrated.

Two issues merit discussion. First, our analysis focuses on subsidies for fixed-line internet connections. In practice, many individuals access the internet via cellular connections. However, during the period we study the adoption of mobile internet services in Colombia was in its infancy. Mobile internet penetration in Colombia was low (6%) during 2013 and it rose only a bit (to 9%) by the end of 2014.[14] According to the OECD(2014), these low penetration indicators, jointly with the high prices of mobile devices (smartphones, tablets, and laptops), are a sign that mobile internet services in Colombia were still under-developed

[12] Recall that the subsidies applied to broadband connections, which were associated with download speeds greater or equal to approximately 1 Mbps and upload speeds greater or equal to 0.5 Mbps.

[13] In some municipalities Internet services for low-income households are offered only to one strata.

[14] Source: MinTIC - SIUST. http://colombiatic.mintic.gov.co/estadisticas/stats.php?s=1. Retrieved April 30, 2018.

	2013		2014		Change
	Mean	Std. Dev.	Mean	Std. Dev.	(%)
(A) *Plan level*					
Price w/ subsidy	21.16	10.18	20.03	9.70	-5.3
Market price	21.74	10.41	20.87	9.84	-4.0
Mid-speed	0.40	0.49	0.46	0.50	14.6
High-speed	0.34	0.47	0.36	0.48	7.0
Tech: Cable	0.27	0.45	0.23	0.42	-17.1
Tech: xDSL	0.53	0.50	0.48	0.50	-10.3
ISP seniority	11.98	4.39	15.45	5.28	29.0
Broadband (1/0)	0.74	0.44	0.82	0.39	11.3
# Unique plans	315		305		
(B) *Market level*					
# Plans	6.45	6.58	7.17	7.46	11.1
# ISP	1.53	0.98	1.68	1.13	9.7
Plans per ISP	3.91	2.32	3.90	2.31	0.0
ISP Seniority	19.10	13.38	24.33	18.86	27.4
HHI ISP-subsc.	0.89	0.20	0.87	0.20	-2.2
HHI ISP-plans	0.84	0.25	0.81	0.27	-4.4

Notes: Prices are real 2008 US dollars. Mid-speed indicates download speed is 1-4 Mbps; high-speed is more than 4Mpbs. ISP seniority is number of quarters the ISP has been operating in the municipality since 2010:1.

Table 3: Internet Services Descriptive Statistics for Lower Strata

and were targeted mainly to well-off segments of the population. As a result, cellular services were not (yet) included in the choice set of low-income consumers.

Second, a major barrier to connection could be inadequate infrastructure to deliver the service. We have information on plans offered in all six residential stratas as well as to businesses, which we use to examine in which municipalities and time periods internet connections were available. Specifically, if an internet plan was provided in a specific municipality in a certain period, then we assume that the technology employed to offer that plan was available to residents of stratas 1 and 2. The data show that there is sufficient deployment of infrastructure nationwide. The exception being municipalities located in the Amazon region (southern Colombia), which are typified by low population density and challenging geography.

2.3 Data on Innovation Diffusion

The rate at which the internet diffuses in the population may impact take-up, but measures of diffusion do not exist in the data. Fortunately, there are many papers that examine how technology diffuses across populations. In order to estimate a model of innovation diffusion, we require information on the number of potential adopters, variables that impact the magnitude of innovation, and those that impact the growth of innovation.

Timing variables are those that impact the magnitude of innovation, but not the growth rate. The literature notes timing is impacted by items such as access to fiber optic network facilities, the number of fixed phone line connections, and the economic importance of the area. Whereas, variables that impact how rapidly innovation grows are related to internet literacy and ease of access to digital devices. We obtain data on these variables from MinTIC (with the exception of GDP, which we obtain from the Departamento Administrativo Nacional de Estadistica).

	Mean	Std. Dev.	Min	Max
Internet Subscribers	1.68	14.66	0.001	615.3
Landline Subscribers	1.82	18.20	0	642.7
GDP per-capita (10^6)	0.008	0.012	0.001	0.3
Has Fiber Network	0.94	0.24	0	1.0
Lit & Equip: Kiosks (10^2)	0.04	0.08	0	1.5
Lit & Equip: Centers (10^2)	0.04	0.07	0	0.6
Equip: PC Schools	0.13	0.77	0	46.6
Equip: Tablet Schools	0.19	0.70	0	17.7
Lit: Schools	0.17	0.53	0	11.2
Lit: Citizen	0.13	0.84	0	43.1
Subsidy in Place	0.59	0.49	0	1.0
# time periods per Market	22.12	9.33	2	34.0

Notes: The unit of observation is the market-quarter combination. Statistics are in thousands unless otherwise indicated. The total number of observations is 44,258 from 2011:2-2019:3. There are 2001 (1050) unique markets (municipalities).

Table 4: Innovation Diffusion Variables Descriptive Statistics for Lower Strata

Table 4 provides descriptive statistics. We have data on the number of subscribers from 2011:2-2019:3 so we use this longer time series to estimate innovation diffusion. We aggregate the data across plans, so an observation is a market-quarter combination, which yields 44,258 observations (for comparison, 6,531 of these are from 2013-2014). An average market consists of 1,680 subscribers.

Regarding timing variables, 94% of the markets have fiberoptic backbone facilities and there are about 1,820 landline subscribers per market.[15] The average municipality has a GDP per-capita of $8,000.[16] Regarding growth variables, in 2010 Colombia initiated the Plan Vive Digital with the objective to increase both access to the internet and digital literacy. The program does so in a number of ways, including providing access to computers and literacy education through large centers (Lit & Equip: Centers) and through smaller centers (Liter & Equip: Kiosks); providing IT devices to public schools (Equip: PC Schools; Equip: Tablet Schools); providing training on digital skills to teachers and parents (Lit: Schools); and providing training in basic computing skills (Lit: Citizen).

As Table 4 shows, an average municipality has 4 centers and 4 kiosks, but the number of centers varies with some having no centers and one as many as 60, with as many as 150 kiosks in larger municipalities. Regarding the impact on public schools, an average school receives 130 PCs and 190 Tablets annually, where 170 teachers and parents receive training to develop digital skills. Finally, about 130 citizens of an average municipality receive training in basic digital skills regarding online security, digital rights and e-commerce.

2.4 Descriptive Analysis

Figure 2 shows the extent of the digital divide in Colombia. Adoption by the lower-SES stratas 1 and 2 (that are the focus of the subsidy) is given by the lower (green) line, while adoption by the remaining population (i.e, stratas 3-6) is given by the upper (blue) line. The period covered by the vertical dashed lines corresponds to the policy intervention for the lower stratas 1 and 2 (2012:3-2014:4). The figure illustrates three notable trends. First, the difference in penetration rates across SES strata groups shows the extent of the digital divide in Colombia - where penetration rates are approximately three-fold higher in the mid-

[15] Data from 2011:2-2011:4 is imputed using linear interpolation.

[16] Proxy for GDP computed as the product of the (1) weight assigned by the Colombian statistical office of economic importance of the municipality and (2) GDP by department.

upper SES strata relative to the lower two strata (stratas 1 and 2). Second, the penetration rate among the lower SES strata stabilized after the targeted subsidy policy. Third, prior to and during the subsidy period, penetration among strata 1 and 2 households was increasing at a higher rate than that of the other stratas. These trends suggest the subsidy helped to decrease the digital divide.

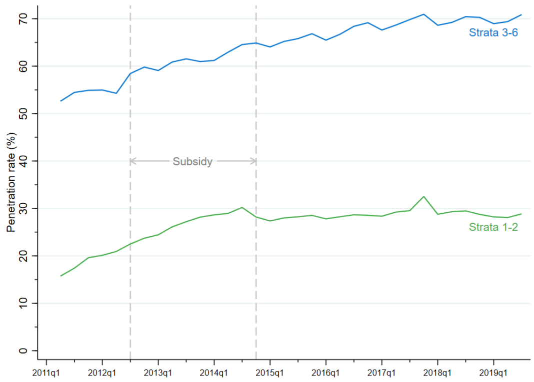

Figure 2: Penetration rates

To substantiate this suggestive finding, we quantify the impact of the policy using a Difference-in-Differences design. We focus on the time period prior to and during the subsidy (2011:2-2014:4), and aggregate the data across plans such that an observation is at the strata-municipality-period level. Recall that ISPs offered different plans to households depending on their strata and municipality. In some markets ISPs did not offer broadband plans or grant subsidies, and, hence, none of the available plans were subsidized. We use this fact to define the treatment group. Specifically, households in strata-municipality-periods that were offered (subsidized) broadband plans are our treatment group (10,478 observations). The remaining (1,604 observations) are part of the control group. We leverage the fact that the treatment (i.e., the price subsidy) occurs at the same point in time (post 2012:3) for all households in stratas 1 and 2 and estimate the following equation:

$$r_{lkt} = \psi V_{lkt} + \rho_{DD}(Post \times Treatment)_{lkt} + \kappa_t + \kappa_k + \kappa_l + \varepsilon_{lkt}, \tag{1}$$

where r_{lkt} denotes the penetration rate for strata l in municipality k in period t. The vector

V_{lkt} contains control variables that may influence the take-up of Internet services and vary across strata or municipalities. These include municipality GDP per-capita, the presence of a fiber network, the number of ISPs, and number of technologies in each municipality-strata. The coefficient of interest is ρ_{DD}, which measures the effect of the subsidy scheme on the penetration rate in treatment markets relative to control markets. *Post* is an indicator equal to one starting at the beginning of the subsidy policy (in 2012:3). We control for time, market, and strata fixed effects (denoted by the κ terms). We cluster the standard errors (ε_{lkt}) at the municipality level. This allows for arbitrary serial correlation in errors within a municipality over time.

Columns (1)-(3) of Table 5 present the results. The first two columns show that the policy intervention led to an increase of the penetration rate of two percentage points. Column (3) allows the effect of the policy to vary by strata. We find that the effect on the penetration rate is quite moderate, but significant, for the lowest SES group (1 percentage point). The impact for strata 2 is three times larger than that of strata 1. However, we note that our control group is small relative to our treatment group, therefore, we do not treat our findings as causal but rather descriptive and suggestive.

	Strata 1 & 2			Strata 3 & 4	
	(1)	(2)	(3)	(4)	(5)
Post×Treatment	0.021***	0.023***	0.011**	0.019	0.02
	-0.01	-0.01	0	-0.01	-0.01
Post×Treatment×Strata 2			0.023**		
			-0.01		
Strata/Time FE	No	Yes	Yes	No	Yes
Adjusted R-squared	0.954	0.961	0.961	0.933	0.934
Observations	12082	12082	12082	8116	8116

Notes: All regressions include time, muncipality/strata fixed effects. and control variables, which are municipality GDP per-capita, fiber network, number of ISPs, and number of technologies in each market. Standard errors are robust to heteroskedasticity and allow intragroup correlation within municipalities. $***p < 0.001, **p < 0.05, *p < 0.1$.

Table 5: Reduced form: Penetration rates

The latter columns present results of a validity test. We estimate the same model on a market which should not have been impacted by the subsidy - the middle class (stratas 3 and 4).[17] The results show no impact of the subsidy for the middle-class, which lends

[17] In this exercise, we use strata 3 and 4 in treated municipalities as treatment groups (7,560 observations),

credibility to the positive impact in the low-SES strata.[18]

In summary, preliminary evidence is consistent with the subsidy intervention having a positive effect on the take-up of residential Internet services among the low SES strata. As noted previously the subsidy may impact the penetration rate via the price directly, via its impact on the diffusion rate, or due to supply-side reactions to the policy. We now present a model of supply and demand for Internet services to explore both the impact of the subsidy as well as the mechanisms through which it may impact penetration.

3 Model

In this section we develop a model of internet adoption among low-income populations. The framework we use is an "aggregate" demand model, along the lines of Berry et al. (1995) (hereafter BLP) in that we don't observe individual purchase data. However, we do observe more about the consumers than what is typical for aggregate demand models in that, for each municipality, we observe the number of individuals in a particular SES group (strata) that subscribe to each plan. We begin this section by laying out the primitives of the demand model. Then we incorporate the fact that ISPs do not offer all plans in all markets. Finally, we show how we account for the impact of the (unobserved) diffusion rate on demand.

3.1 Demand

Following in the spirit of BLP, an individual chooses from a set of plans offered in her municipality and SES (income) strata at time t. We refer to a municipality-strata combination as a market. The indirect utility consumer i obtains from plan j offered in her market m is

$$u_{ijmt} = \delta_{jmt} + \mu_{ijmt} + \epsilon_{ijmt}. \tag{2}$$

and the same strata in untreated municipalities as the control group (556 observations). The treatment status of the municipalities is given by the presence of subsidies in the two lowest strata.

[18] We also conduct an event study in which we compare treatment and control groups on pre-treatment dynamics. The results from the event study (presented in the Online Appendix) support the findings that the subsidies boosted the adoption of Internet services.

The δ_{jmt} captures the mean utility every consumer in m derives from plan j at time t. The composite random shock, $\mu_{ijmt}+\epsilon_{ijmt}$, captures heterogeneity in tastes for product attributes, and ϵ_{ijmt} is a mean zero stochastic term distributed i.i.d. type I extreme value across products and consumers.[19] Henceforth, we suppress the time index for ease of exposition.

Each plan is characterized by a monthly subscription fee (i.e., the price) (p_{jm}), a discount on the monthly price (d_{jm}) (which may be equal to zero), and non-price observed attributes (x_{jm}) that include transmission speeds, internet access technology, and whether the plan is offered by an established ISP.

The value of adopting an internet connection may be impacted by the rate at which it is adopted in the market. For example, a higher degree of internet diffusion could positively impact a family if it provides more opportunities to gain experience with the internet. On the other hand, if the adoption rate is very fast the household may decide not to adopt as they have more opportunities to access the internet via their neighbors. To account for these potential effects, we allow the speed of diffusion in the market, captured by b_m, to impact the utility from adopting. In the next section we discuss how we obtain a measure of the diffusion rate.

One point is useful to mention. The impact of the diffusion rate on demand, if not controlled for, is captured in the unobserved market heterogeneity term (i.e., the ξ_{jm}). To the extent that the price subsidy increases the diffusion rate, the subsidy would be correlated with the error term. Accounting for the impact of the diffusion rate on demand allows us to control for this potential endogeneity, and so to obtain a causal impact of the pricing subsidy on adoption. We discuss identification in section 4.

Market-specific variables that may impact residential internet services are represented by G_m which includes a set of municipality-fixed effects, firm (ISP)-fixed effects and SES-strata-fixed effects. Finally, plan attributes that are unobserved to the researcher but known to consumers and ISP providers are captured by ξ_{jm}. Mean utility is given by

$$\delta_{jm} = \alpha_l(p_{jm} - d_{jm}) + \lambda_l b_m + \beta x_{jm} + \gamma G_m + \xi_{jm}. \tag{3}$$

[19] Choices of an individual are invariant to multiplication of utility by a person-specific constant, so we fix the standard deviation of the ϵ_{ijmt}. Estimating an unrestricted variance-covariance matrix is not feasible because there are over $25,000$ plans.

The parameter α_l captures price sensitivity (that may vary by strata l), λ_l captures the effect of the speed of broadband adoption growth in the market (i.e., the diffusion rate) that may vary by strata, and β and γ capture the importance of non-price plan attributes.

The μ_{ijm} term includes interactions between unobserved (to the econometrician) consumer tastes (ν_i) and service fees ($p_{jm} - d_{jm}$). Specifically,

$$\mu_{ijm} = (p_{jm} - d_{jm})\sigma_v \nu_i \qquad \nu_i \sim N(0,1). \tag{4}$$

where σ_v is a scaler. Finally, consumers may decide not to purchase an internet plan. Normalizing service fees to zero, the indirect utility from the outside option is

$$u_{i0m} = \xi_{0m} + \epsilon_{i0m}.$$

We also normalize ξ_{0m} to zero, because we cannot identify relative utility levels.

3.2 Limited Choice Sets

When deciding which plan to purchase, consumers face a restriction on their available choices as not all plans are offered in all municipalities nor to all strata within a municipality. At the municipality level, households are only able to subscribe to services offered within their local network. Within each municipality, households are only able to subscribe to services offered to their SES strata. Therefore, the choice set of individual i depends on her municipality and strata. We model this restriction following previous literature on limited choice sets (e.g., Sovinsky (2008)). However unlike Sovinsky (2008), we observe the set of plans that are available to an individual. Furthermore, as we discussed in section 2, there are not a large number of plans available in each market. Therefore, we assume that consumers are aware of the plans offered in their market. This yields the following (conditional) probability that consumer i purchases j

$$s_{ijm} = \frac{\exp\{\delta_j + \mu_{ij}\}}{1 + \sum_{r \in \mathbf{J}_m} \exp\{\delta_r + \mu_{ir}\}} \mid j \in \mathbf{J}_m \tag{5}$$

where the summand is over plans offered in consumer i's market.

We assume that a consumer purchases at most one plan per period, that which provides the highest utility, U, from the plans available. Let $R_{jm} \equiv \{v_i : U(v, p_{jm}, d_{jm}, x_{jm}, G_m, \xi_{jm}, \epsilon_{ijm}) \geq U(v, p_{rm}, d_{jm}, x_{rm}, G_m, \xi_{rm}, \epsilon_{irm}) \quad \forall r, j \in \mathbf{J}_m, r \neq j\}$ define the set of variables that results in the purchase of j given the parameters. The market share of plan j in market m is

$$s_{jm} = \int_{R_{jm} j \in \mathbf{J}_m} dF(\nu, \epsilon) = \int_{R_{jm} j \in \mathbf{J}_m} s_{ijm} dF_\nu(\nu) \tag{6}$$

where $F(\cdot)$ denotes the respective distribution functions. The second equality follows from independence assumptions. The conditional probability that i purchases j is given in (5).

Demand for plan j in market m at time t is

$$\mathcal{M}_{mt} s_{jmt}$$

where $\mathcal{M}_{mt}$ is the number of households by strata and municipality as reported by the Census. Total demand for plan j is then given by

$$D_{jt} = \sum_{m=1}^{M} \mathcal{M}_{mt} s_{jmt}. \tag{7}$$

3.3 Innovation Diffusion Rate

As we discussed earlier, the rate at which the internet diffuses in the population may impact take-up, but measures of diffusion do not exist in the data. Fortunately, there are many papers that examine how technology diffuses across populations. We use innovation diffusion models, together with data on the number of subscribers in each market across time (which we have from 2011:2-2019:3), to obtain an estimate of the speed of internet diffusion (b_{mt}).[20]

Following the literature (see Geroski (2000) and Gruber and Verboven (2001) for example), we estimate an innovation model given by

$$y_{mt} = \frac{\mathcal{M}_{mt}}{1 + \exp(-a_{mt} - b_{mt}t)} \tag{8}$$

where y_{mt} denotes the number of consumers that have adopted internet services, a_{mt} the timing of diffusion, and b_{mt} the speed of diffusion.

The timing variable a_{mt} shifts the magnitude of the function, but does not impact the shape. We specify the timing variable as

$$a_{mt} = \sum_{l} \pi_l z_{mt}$$

where the vector z_{mt} contains the GDP of the municipality, access to a fiber network, and the number of landline subscribers per-capita. The parameter vector π_l allows for potentially differential impacts across SES strata.

The speed of diffusion could depend on a variety of factors related to internet literacy (such as those implemented by the Vive Digital program) as well as whether the subsidy policy is in place. These are captured in the vector v_{mt}. The diffusion rate is given by

$$b_{mt} = \sum_{l} \varphi_l v_{mt}, \tag{9}$$

where φ_l is a parameter vector that varies by strata. We jointly estimate the model of innovation diffusion together with demand to obtain an estimate of the diffusion rate.

Our control variables may not capture all features of the diffusion process for some markets, especially for those at an early stage of diffusion. To address this we include municipality fixed effects. To avoid compounding the computational burden and estimation issues related to non-linear models (e.g., incidental parameters problems), we follow Gruber

[20] Ideally, we would estimate a market-specific diffusion rate. However, this proved problematic primarily because there were an infeasible number of parameters (estimating market-specific innovation diffusion models jointly with demand yielded more than 3000 parameters given the large number of markets). In addition, some of the markets are at more rudimentary stages of innovation so the market-specific model did not converge for these smaller markets.

and Verboven (2001) and transform the estimation equation. In particular, we estimate

$$\log\left(\frac{y_{mt}}{\mathcal{M}_{mt} - y_{mt}}\right) \;=\; a_{mt} + b_{mt}t. \tag{10}$$

This allows us to control for municipality fixed effects through the timing variable a_{mt}, and differences in the autonomous diffusion rate, through the speed of diffusion b_{mt}.

3.4 Firm Behavior

We assume that there are F non-cooperative ISPs which are Bertrand-Nash competitors. Each ISP offers a subset of the J plans in a market, $\mathcal{J}_{fm}$. Suppressing time notation, profits of firm f are

$$\sum_{j,m \in \mathcal{J}_f} (p_{jm} - mc_{jm})\mathcal{M}_{mt} s_{jm}(p_{jm} - d_{jm}) - \mathcal{C}_{fm} \tag{11}$$

where s_{jm} is the market share given in (6); mc_{jm} is marginal cost of production; and $\mathcal{C}_{fm}$ are fixed costs of production.

Following BLP and others, we assume mc_{jm} are log-linear and composed of unobserved (ϑ_{jm}) and observed cost characteristics (h_{jm}) which include transmission speeds, the internet access technology, the seniority of the ISP offering the plan, population density and a time trend. The vector of parameters to be estimated is η. The (log) marginal cost function is

$$\ln(mc_{jm}) = h'_{jm}\eta + \vartheta_{jm}. \tag{12}$$

Given their plans, the plans offered by competitors, and any potential reduction in the monthly fee from the subsidy policy, firms choose prices simultaneously to maximize profits. Product attributes that affect demand and those that affect marginal costs are treated as exogenous to the firm's pricing decisions. Constant marginal costs imply pricing decisions are independent across sectors. Optimal prices satisfy

$$s_{jm}(p_{jm} - d_{jm}) + \sum_{r,m \in \mathcal{J}_f} (p_{rm} - mc_{rm}) \frac{\partial s_{rm}(p_{rm} - d_{rm})}{\partial p_{jm}} = 0, \tag{13}$$

where the subsidy d_{jm} is non-zero if plan j is a broadband plan offered by an ISP in the subsidy program.[21]

4 Estimation and Identification

The parameters of the model are $\theta = \{\alpha, \beta, \gamma, \lambda, \sigma\}$, $\Gamma = \{\pi, \varphi\}$, and η. We are ultimately interested in the estimates of the demand model (θ), but in order to obtain these estimates we require a measure of the diffusion rate (b_{mt}), which is estimated from the innovation diffusion model. We estimate the parameters $\Theta = \{\theta, \Gamma, \eta\}$ simultaneously by Generalized Method of Moments. We begin by discussing the residuals/unobservables and the associated sample moments which we use to form the joint GMM objective function. These moments are estimated over the period for which we have adequate pricing data (2013 and 2014). Then we provide details on how the innovation and demand models are linked, followed by a discussion of identification, and details of the estimation algorithm.

Innovation Moment Residual Identification of the parameters of the innovation diffusion model (Γ) requires information on a time series of subscribers and control variables. As it does not require pricing data we can use data over our whole sample (2011:2 until 2019:3) to estimate Γ. We complete the econometric model by adding an error term (ε_{mt}) to equation 10.

$$\log\left(\frac{y_{mt}}{\mathcal{M}_{mt} - y_{mt}}\right) - a_{mt} - b_{mt}t = \varepsilon_{mt} \tag{14}$$

The model can be expressed to define the moment residual

[21] To participate in the subsidy program, the ISP had to set a subscriber growth goal. We provide more detail on how we link subsidized prices to plans in Appendix A.

$$g(z_{mt}, \upsilon_{mt}; \Gamma_0) = \varepsilon_{mt} \tag{15}$$

where the population moment restriction is given by

$$E[\varepsilon_{mt}] = E[g(z_{mt}, \upsilon_{mt}; \Gamma_0)] = 0. \tag{16}$$

Demand Unobservables We compute the observed market share for a plan as

$$S^{obs}_{jmt} = \frac{hh_{jmt}}{\mathcal{M}_{mt}} \tag{17}$$

where hh_{jmt} represents the number of subscribers. First, following BLP, we restrict the model predictions for j's market share to match the observed shares as computed in equation 17. To do so, we solve for $\delta(S, \theta)$ that is the implicit solution to

$$S^{obs}_t - s_t(\delta, \theta) = 0 \tag{18}$$

where S^{obs}_t represents the vector of observed shares and s_t is the vector of predicted shares. We substitute the $\delta(S, \theta)$ for δ in equation 3 when calculating the moments and we measure the diffusion rate by $b_{mt} = \sum_l \varphi_l \upsilon_{mt}$. The moment unobservable is

$$\xi_{jmt} = \delta_{jmt}(S, \theta) - \alpha_l(p_{jmt} - d_{jmt}) - \lambda_l \sum_l \varphi_l \upsilon_{mt} - \beta x_{jmt} - \gamma G_{mt}. \tag{19}$$

Notice that the subsidy policy impacts take-up via the price effect $(p_{jmt} - d_{jmt})$ and via the diffusion rate $(\sum_l \varphi_l \upsilon_{mt})$ as the latter is a function of whether the subsidy is in place. We return to this in our analysis of the impact of the subsidy policy in section 6.

Cost Unobservables We use the demand system estimates to compute marginal costs. In vector form, the J FOCs from (13) imply

$$mc = p - \Delta(\theta,\delta)^{-1}s(\theta,\delta) \tag{20}$$

where $\Delta_{j,r} = -\frac{\partial s_{rm}}{\partial p_{jm}} I_{j,r}$ with $I_{j,r}$ an indicator equal to one when j and r are plans from same ISP. Combining (20) and (12) yields the second moment unobservable:

$$\vartheta = \ln(p - \Delta(\theta,\delta)^{-1}s(\theta,\delta)) - h'\eta. \tag{21}$$

Linking the Innovation Model to the Demand Model As noted above, the demand parameters are chosen such that observed and predicted market shares are equated ($S_t^{obs} - s_t(\delta,\theta) = 0$) and this system of equations is solved market by market. One of the plans ($j = 0$) is the plan not to purchase any internet connection. This outside good is defined as the difference between one (since the market shares of all products must sum to one) and the sum of the market shares for all "inside" goods. That is, for each market $s_0 = 1 - \sum_j S_j^{obs}$ which can be rewritten in the notation of our innovation model as $s_0 = 1 - \frac{y_{mt}}{\mathcal{M}_{mt}}$.

The parameters of demand are chosen such that the "observed" market share for the outside good (which depends on the number of observed plan users) matches the model's prediction for that good. Concurrently, the innovation parameters are chosen such that they match the number of observed plan users to those implied by the innovation model

$$y_{mt} - \frac{\mathcal{M}_{mt}}{1 + \exp(-\sum\limits_l \pi_l z_{mt} - \sum\limits_l \varphi_l v_{mt} t)} = \varepsilon_{mt}.$$

The moment restrictions implied by the demand and innovation models are given in equations (15) and (19). We jointly estimate the parameters of innovation diffusion, demand, and supply, which takes into account the cross-equation restrictions of the moments. This yields more efficient estimates, as well as ensures the link between the innovation model and the demand model is explicitly incorporated via the cross-equation moment restrictions. We arrive at parameters that are generated from models that are internally consistent but that utilize different variation in the data to aid in more precise identification of the estimates.

Identification Regarding the identification of the parameters of demand and supply, as in previous studies, we assume that the demand and pricing unobservables (evaluated at the true value of the parameters $\Theta_0 = (\theta_0 \ , \eta_0))$ are mean independent of a vector of observable product characteristics and cost shifters (x, h) :

$$E\left[\xi_j(\Theta_0) \mid (x,h)\right] = E\left[\vartheta_j(\Theta_0) \mid (x,h)\right] = 0 \tag{22}$$

The primitives of the structural model include unobserved characteristics (ξ_{jm} and ϑ_{jm}) that are taken into account by market participants when they decide in which plan to enroll but are unobserved by the econometrician. This leads to an endogeneity problem caused mainly by the relationship between prices and unobserved attributes of the services.

To account for the potential endogeneity of prices we use instruments that are related to the cost of providing internet services. These include the monthly cost to an ISP of a network internet connection and its interaction with the connection speed.[22] The variation of the cost shifter comes from the highly fragmented telecom network and the respective interconnection charges. The telecommunications fixed network in Colombia consists of multiple geographical segments owned by different (private) operators. To offer Internet services in a particular market, a provider must pay an interconnection charge to the incumbent operator.[23] The interconnect pricing rule is based on the network capacity (in Mbps) and the location of the local market. Accordingly, our main cost shifter exhibits substantial variation across municipalities (location of the market) and ISPs (variation in network capacity).

The inclusion of the diffusion rate in the indirect utility also poses identification issues. The demand model (Section 3.1) describes the take-up choice, ultimately, as a function of the number of consumers that have already taken up the Internet service in the market at a specific time period. In our setting, the functional form capturing this relationship is given

[22] We obtain these data from the telecommunications competition authority (Form 7)

[23] When the ISP is the operator of the local fixed network, the interconnection fee is zero.

by the timing of the diffusion process (diffusion rate). In light of the reflection problem of Manski (1993), local spillovers between consumers hinders the identification of the diffusion rate in demand (Rysman, 2019).

The take-up of Internet services can be affected by the diffusion rate in two ways. First, consumers may obtain higher utility the greater is the number of peers that they can interact with (e.g., on social media). Since the number of consumers is positively correlated with how fast the service is spread across the population, the diffusion rate may positively influence take-up. Second, a higher diffusion rate implies a higher probability that a non-subscriber meets a subscriber and they exchange information about the service (e.g., quality features, functionalities of the service, etc.). In the context of low-income and non-tech savvy consumers, the exchange-of-information channel may be potentially as important as the direct network effect. The information flow is, to some extent, correlated with the attributes of the service itself. For example, a high-quality plan (e.g., no latency issues and high connection speed) may prompt a subscriber to convey information to a non-subscriber about the experience of using Internet services. Neglecting this relationship between (unobserved) service characteristics and the diffusion rate may yield biased estimates.

Based on recent literature on structural models of the diffusion of innovation (Arieli et al., 2020; Board and Meyer-ter-Vehn, 2021), the speed of innovation diffusion depends on the topology of the social network. Specifically, the learning dynamics depend both on the size (number of links) and the strength of the network. To identify the diffusion rate parameter, we can use a variable that is related to the structure of the social network, and hence to the diffusion rate, if it is also orthogonal to the unobserved characteristics of the service. We argue disruptions in the social network brought on by the Colombian internal armed conflict facilitate such instruments. The five-decade war has drastically disrupted the social order, across municipalities and over time, through different violent actions such as kidnapping, child recruitment, ambushes, armed assaults, military clashes, etc.. These violent incidents can alter the size of the social network (e.g., through forced displacement) and weaken the degree of connectivity of the social network. Further, the Internet product portfolio, and the characteristics thereof, is a decision made by providers at a regional (or even national) level and it is unlikely to be affected by the armed conflict.

We obtained a number of variables related to the armed conflict.[24] Overidentification tests indicate that we cannot use all the conflict variables at our disposal. Therefore, we chose the variables that had the most variation and were equally disruptive in cities as in more remote areas. Specifically, we use (lagged) variables of the number of ambushes of illegal armed groups in the municipality, the average number of illegal ambushes in neighboring municipalities, and the number of assaults from armed groups in the municipality.[25] We discuss the validity of the instruments in Section 5.2.

Identification of the other parameters of the demand function is in the same spirit as previous studies. Namely, associated with each plan is a mean utility, which is chosen to match observed and predicted market shares. If consumers were identical, then all variation in choices would be driven by variation in plan attributes. Variation in plan market shares corresponding to variation in the observable attributes of those plans (such as connection speed) is used to identify the parameters of mean utility. While a plan may have attributes that are preferred by many consumers (high β's), it may also have attributes that appeal to certain types of consumers. Identification of the taste distribution parameters (σ) relies on information on how consumers substitute. ISPs offer different plans over time and variation of this sort is helpful for identification of σ. The distribution of unobserved tastes, ν_i, is fixed over time, but the choice set of available plans is changing over time. Variation in sales patterns over time as the choice sets change allows for identification of σ.

The parameters that determine innovation diffusion are identified by covariation in measures of access to a fiber connection and a landline, the GDP of the municipality, the implementation of the price subsidy policy, and literacy programs with the number of subscribers to a plan in a particular strata-municipality over the period 2011:2-2019:3.

Estimation Algorithm and Properties of the Estimator We use Generalized Method of Moments (GMM) to find the parameter values that minimize the objective function, $\Lambda' Z A^{-1} Z' \Lambda$, where A is a weighting matrix which is a consistent estimate of $E[Z'\Lambda\Lambda' Z]$ and Z are instruments orthogonal to the composite error term Λ. Let Z_ε, Z_ξ and Z_ϑ be the

[24] We thank the Centro de Estudios sobre Desarrollo Económico (CEDE) for sharing the data.

[25] An assault is defined as a surprise attack with the aim to create nuisance and military stress without producing great damage or a large number of casualties.

respective instruments for each residual/unobservable, the sample moments are

$$Z'\Lambda = \begin{bmatrix} \frac{1}{MT_{innov}} \sum_n Z_{\varepsilon,n} g_n(\Gamma) \\ \frac{1}{J} \sum_j Z_{\xi,j} \xi_j(\delta, \alpha, \beta, \gamma, \varphi_l) \\ \frac{1}{J} \sum_j Z_{\vartheta,j} \vartheta_j(\delta, \theta, \eta) \end{bmatrix}.$$

where MT_{innov} denotes the longer time period (2011:2 until 2019:3) used in the innovation moments, $Z_{\varepsilon,n}$ is column n of Z_ε, $Z_{\xi,j}$ is column j of Z_ξ and $Z_{\vartheta,j}$ is column j of Z_ϑ. Joint estimation takes into account the cross-equation restrictions on the parameters that affect both estimation of innovation diffusion, demand, and supply which yields more efficient estimates. This comes at the cost of increased computation time since joint estimation requires a non-linear search over most of the parameters of the model.[26] If the parameters don't minimize the moments (according to some criteria) we make a new guess of the parameters. We repeat the estimation steps until the moments are close to zero.

In summary, we employ the following estimation algorithm: Calculate the instruments and keep them fixed for the duration of the estimation. Given a value of the parameters, Θ,

(i) Calculate the diffusion rate (b) given in equation 9.

(ii) Compute the simulated market shares and solve for the vector δ that equates simulated and observed shares.

(iii) Calculate α, β, γ, and λ and compute the demand unobservables, ξ (see 19).

(iv) Calculate η and compute the cost side unobservables, ϑ (see 21).

(v) Calculate Γ and compute the diffusion sample moments.

(vi) Search for the parameter values that minimize the objective function: $\widehat{\Lambda}' Z A^{-1} Z' \widehat{\Lambda}$, where $\widehat{\Lambda}$ is the composite error term resulting from the simulated moments. If the parameters don't minimize the moments (according to some criteria) make a new guess of the parameters. Repeat until moments are close to zero.

There is no closed-form solution for the market shares (equation 6) so they must be simulated. We follow standard simulation techniques by sampling a set of "individuals"

[26] As in Nevo (2000), we restrict the non-linear search to a subset of the parameters.

where each consists of taste parameters drawn from a normal distribution.[27] The parameters are simultaneously estimated using two-step feasible GMM. We restrict the non-linear search to a subset of the parameters, namely: the standard deviation of the random coefficients of the demand model, and the parameter associated with the subsidy variable in the model of innovation diffusion. The estimates are obtained using the pattern search optimization routine.[28]

Using the results of Pakes and Pollard (1989), the resulting estimator is consistent and asymptotically normal. As the number of pseudo-random draws used in simulation $R \to \infty$ the method of simulated moments covariance matrix approaches the method of moments covariance matrix. The reported (asymptotic) standard errors are derived from the inverse of the simulated information matrix which allows for possible heteroskedasticity.

5 Results

5.1 Innovation Diffusion Estimates

We begin by investigating what elements of the innovation model indicate about the rate of diffusion of residential internet services among the poorest households. Table 6 presents the estimates from different specifications of the innovation diffusion model (see equation 14). The most simplistic specification (in Column (1)) shows that both the timing and the speed of diffusion is faster in the higher of the two low-income strata. In addition, the presence of the subsidy is associated with a faster speed of diffusion.

Not surprisingly, we also find that timing of internet diffusion is faster when the municipality strata has access to a fiber network and when there are more landlines per-capita.

[27] To reduce simulation error, we employ 500 latin hypercube sampling draws. The market share simulator is then the average over individuals of the choice probabilities. To find the vector δ that equates simulated and observed market shares, we use SQUAREM (Varadhan and Roland, 2008). The contraction mapping tolerance is 1e-12.

[28] To speed up the optimization routine, we restrict the parameters search such that the random coefficient is non-negative and the subsidy parameter of the innovation diffusion lies in the interval [-0.1,0.1]. The former constraint is not necessary as the value of the objective function should be symmetric about zero (Conlon and Gortmaker, 2020). The latter constraint only prevents the routine from evaluating unreasonable values.

	(1)	(2)	(3)	(4)	(5)	(6)
Timing of Diffusion						
Strata 2	1.7057***	1.7097***	1.6941***	1.6941***	1.6942***	1.6881***
	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)
Fiber Network		0.4861***	0.4853***	0.4854***	0.4650***	0.4649***
		(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Landline per-capita			0.0762***	0.0762***	0.0760***	0.0753***
			(0.018)	(0.018)	(0.018)	(0.018)
GDP per-capita				-0.1314	0.1021	0.1006
				(1.740)	(1.738)	(1.738)
Speed of Diffusion						
Strata 2	0.0053***	0.0052***	0.0054***	0.0054***	0.0054***	0.0054***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Equip: PC Schools					0.0003	0.0009***
					(0.000)	(0.000)
...× Strata 2						-0.0012***
						(0.000)
Equip: Tablet Schools					0.0024***	0.0020***
					(0.000)	(0.000)
...× Strata 2						0.0009*
						(0.001)
Lit: Schools					0.0026***	0.0013***
					(0.000)	(0.000)
...× Strata 2						0.0026***
						(0.001)
Lit: Citizen					0.0012***	0.0017***
					(0.000)	(0.000)
...× Strata 2						-0.0010**
						(0.000)
Lit & Equip: Kiosks					0.0324**	0.0325**
					(0.014)	(0.014)
Lit & Equip: Centers					0.0162**	0.0162**
					(0.008)	(0.008)
Subsidy (1/0)	0.0105***	0.0061***	0.0060***	0.0060***	0.0082***	0.0082***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Implied Average Diffusion Rate	0.059	0.054	0.054	0.054	0.054	0.054
	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)
Adjusted R-squared	0.653	0.655	0.656	0.656	0.657	0.657

Notes: 44,258 observations. All specifications include constants and municipality fixed effects (timing and speed of diffusion). Robust standard errors are reported in parentheses. $***p < 0.001, **p < 0.05, *p < 0.1$.

Table 6: Preliminary Innovation Diffusion Estimates

However, conditional on these variables, the GDP per-capita does not significantly impact the diffusion timing. As specifications (5) and (6) show, policies implemented to increase internet literacy and access to equipment are significantly associated with faster diffusion speed. Specification (6) allows the coefficients of non-subsidy policies to vary by income strata. The provision of computers to public schools is associated with faster diffusion speed only among households in strata 1. The relationship between tablets provision and diffusion speed, on the other hand, is positive for both strata and higher for strata 2. These findings could be related to the technology adoption by income strata. Computers might be at an earlier stage of adoption for strata 1 than for households in strata 2, while tablets are more recent technology and can be perceived as innovative for both strata. Regarding digital training, both policies are associated with a faster diffusion speed. The initiative targeting citizens is related to a faster rate for strata 1 compared to strata 2 which could reflect the engagement in this program by strata 1 households. Based on goodness-of-fit, we use the specification in column (6) to estimate the full model.

5.2 Demand Estimates

Table 7 provides preliminary estimates of what elements of the demand inform consumers choices among the poorest households using the diffusion rates implied by specification (6) of Table 6.

We begin with motivation for our price and diffusion rate instruments. Recall that, for price we use cost instruments and for the diffusion rate we use Colombian conflict variables. Column (1) presents the Logit OLS estimates, where neither price nor diffusion rate endogeneity concerns are addressed. In column (2), we control for price endogeneity and find that the impact of price becomes more negative. This result is expected as (endogenous) prices are likely to be positively correlated with unobserved quality. In addition, column (3) addresses diffusion rate endogeneity issues. High unobserved quality may indicate a market that has reached a later stage of the diffusion process, typified by a higher proportion of tech-savvy households and a slower diffusion rate. This negative correlation between diffusion rate and unobserved quality produces downward biased OLS estimates. In column (3), we address this bias and find that the diffusion rate coefficient becomes higher than the OLS estimator.

Our findings indicate that the diffusion rate significantly and positively impacts the probability an individual enrolls in an Internet plan. This effect needs not to be uniform across individuals. Column (4) incorporates an interaction between diffusion rate and the strata 2 indicator, and column (5) adds a similar interaction with price. The estimates indicate that strata 2 individuals are less price sensitive and less affected by the diffusion rate than strata 1 individuals. While we present the results only for our main specification (due to space considerations), we estimated many alternative specifications, where we observed the same patterns as those presented in Table 7. In particular, alternative specifications also

	Logit OLS	Price IV	Price & Diffusion IV		
	(1)	(2)	(3)	(4)	(5)
Price - subsidy	-0.103***	-0.276***	-0.284***	-0.284***	-0.364***
	(0.002)	(0.026)	(0.027)	(0.027)	(0.043)
Price - subsidy × Strata 2					0.149***
					(0.041)
Diffusion Rate	0.106***	0.137***	0.865***	1.197***	1.199***
	(0.028)	(0.031)	(0.238)	(0.254)	(0.273)
Diffusion Rate × Strata 2				-0.581***	-0.667***
				(0.107)	(0.127)
Mid-speed	2.226***	2.156***	2.140***	2.139***	2.104***
	(0.031)	(0.041)	(0.043)	(0.044)	(0.047)
High-speed	1.547***	3.004***	3.061***	3.001***	2.930***
	(0.038)	(0.223)	(0.232)	(0.240)	(0.248)
Tech: Cable	1.350***	1.184***	1.189***	1.201***	1.224***
	(0.058)	(0.070)	(0.072)	(0.075)	(0.077)
Tech: xDSL	1.233***	1.141***	1.140***	1.149***	1.185***
	(0.047)	(0.061)	(0.062)	(0.064)	(0.066)
Seniority (No Est.)	0.075***	0.079***	0.079***	0.086***	0.085***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)
Seniority (Est.)	0.130***	0.171***	0.191***	0.203***	0.197***
	(0.014)	(0.017)	(0.019)	(0.020)	(0.021)
Trend	-0.070***	-0.156***	-0.190***	-0.191***	-0.181***
	(0.008)	(0.016)	(0.021)	(0.022)	(0.022)
Constant	-8.167***	-4.644***	-7.210***	-8.989***	-7.426***
	(0.260)	(0.598)	(0.912)	(1.005)	(1.002)
Strata 2	1.156***	1.259***	0.975***	3.734***	1.518**
	(0.022)	(0.029)	(0.092)	(0.521)	(0.560)
Weak IV		62.435	24.181	17.272	16.474
J-stat (pval)		0.172	0.080	0.211	0.210

Notes: Total number of observations is 44,515. The time period is 2013:1-2014:4. All specifications include municipality fixed effects and firm fixed effects. The weak IV is for price instruments and diffusion instruments (specifications 3-5) and corresponds to the Kleibergen-Paap statistic. J-stat denotes the p-value for the Hansen J test. Robust standard errors are reported in parentheses. $***p < 0.001, **p < 0.05, *p < 0.1$.

Table 7: Preliminary Demand Estimates

show that the (instrumented) rate of diffusion impacts take-up significantly and positively. The Weak IV statistic on the excluded instruments indicates that the pricing and diffusion rate instruments are not weak. We present the first-stage estimates for the instruments in the Online Appendix.

5.3 Full Model Estimates

Table 8 provides the joint estimates of demand (panel A), cost (panel B), and innovation model (panels C and D). Not surprisingly, we find that consumers prefer lower prices, and the random coefficient price estimate shows that there is heterogeneity across households in how they value the price (subsidy). As with the individual demand estimates, we find that the consumers in markets with higher rates of diffusion have a significantly higher valuation for an internet connection, where individuals in strata 2 have a lower valuation on diffusion than those in Strata 1. Higher connection speeds are more costly, but more valuable to consumers.

Providing internet via cable technology is associated with lower marginal costs of provision than xDSL connections and is valued more by consumers. Finally, plans provided by more established ISPs provide more utility. Overall, individuals in Strata 2 obtain more utility from an internet connection than those in Strata 1.

The impact of covariates on innovation diffusion is consistent with those from our preliminary estimates in section 5.1. Policies implemented to increase internet literacy and access to equipment are significantly associated with faster diffusion speed. Taken together our estimates imply that internet adoption diffuses through the population at a rate of around 5 percent on average across markets. This is consistent with estimated diffusion rates in other developing countries. For example, in a cross-country study, Andrés et al. (2010) found internet diffusion rates in low-middle income countries of around 7%. Our estimates also indicate high variation in the diffusion rates (standard deviation around 7%).

In summary, our estimates reveal a number of targets for policymakers interested in impacting internet adoption. First, prices do matter significantly for adoption decisions, suggesting that pricing subsidies should have an impact on adoption decisions. Second, consumers value faster connection speeds and the internet technology through which the service is provided, suggesting the types of plans offered may influence takeup. Third, consumers value plans offered through more established firms perhaps due to quality concerns, Finally, the diffusion of the services plays a prominent role in adoption decisions suggesting that policies aimed at improving internet literacy or providing equipment may be effective in encouraging takeup. We examine the implications of these findings via counterfactual policy experiments.[29]

[29] Our results are robust to alternative specifications (estimates available upon request).

	(A) Demand		(B) log(mc)	
	Parameter	SE	Parameter	SE
Price - subsidy	-0.355***	0.041		
Std. dev. Price - subsidy	0.03***	0.011		
Price - subsidy × Strata 2	0.118***	0.035		
Diffusion Rate	1.065***	0.251		
Diffusion Rate × Strata 2	-0.585***	0.111		
Mid-speed	2.098***	0.047	0.202***	0.011
High-speed	2.888***	0.239	0.678***	0.01
Tech: Cable	1.222***	0.075	-0.104***	0.013
Tech: xDSL	1.179***	0.064	-0.085***	0.011
Seniority (Not Est.)	0.084***	0.007	0.005***	0.001
Seniority (Est.)	0.197***	0.02	0.016***	0.003
Trend	-0.195***	0.025	-0.043***	0.002
Strata 2	1.935***	0.559		
Constant	-7.775***	1.102	11.302***	1.056
Density			-0.001***	0
Fiber Network			0.088***	0.021

	(C) Timing of Diffusion		(D) Speed of Diffusion	
	Parameter	SE	Parameter	SE
Strata 2	1.687***	0.017	0.005***	0.001
Fiber Network	0.432***	0.029		
Landline per-capita	0.075***	0.018		
GDP per-capita	0.166	1.7		
Equip: PC Schools			0.001***	0.000
Equip: PC Schools × Strata 2			-0.001***	0.000
Equip: Tablet Schools			0.002***	0.000
Equip: Tablet Schools × Strata 2			0.001*	0.001
Lit: Schools			0.001***	0.000
Lit: Schools × Strata 2			0.003***	0.001
Lit: Citizen			0.002***	0.000
Lit: Citizen × Strata 2			-0.001**	0.000
Lit & Equip: Kiosks			0.06***	0.014
Lit & Equip: Centers			0.037***	0.008
Subsidy (1/0)			0.016***	0.001

Notes: Demand and cost regressions include municipality fixed effects and firm fixed effects and use 44,515 obsertavions from 2013:1-2014:4. In demand, prices are instrumented with costs IVs, and the diffusion rate is instrumented with conflict IVs. The diffusion rate is scaled by 100. Innovation regressions include municipality fixed effects and use 44,258 observations from 2011:2-2019:3. $***p < 0.001, **p < 0.05, *p < 0.1$.

Table 8: Joint Estimates of Demand, Supply and Innovation Diffusion

6 Subsidy Policy Evaluation

We conduct a number of counterfactual experiments to determine what drives internet adoption. Specifically, we use the estimates of the joint model to calculate take-up under a variety of policies such as changes in: the types of plans the subsidy applies to, the seniority of the provider, and changes in policies that impact the diffusion rate. We report the impact of enacting these policies in the last quarter of our sample (2014:4).

One issue merits discussion first. The rebates that the ISPs receive from the subsidy program are conditional on reaching a target (that they set) for the number of subscribers. This may impact the prices the ISPs charge separate from the reduction granted from the subsidy. If there is a supply-side reaction to the subsidy in that post-subsidy prices are different than those the firms would have chosen had the policy not been in place then our counterfactual results would not reflect this and could lead to incorrect policy conclusions.

To address this issue, we allow for strategic price responses to any policies (actual or counterfactual) that are put into place. Specifically, we compute the equilibrium price in the (counterfactual) world without a pricing subsidy (denoted p^{pre}) which is associated with a (no subsidy) rate of internet penetration. When we implement a new policy we compute the new equilibrium price that firms choose which implies a new probability of adoption - i.e., the post-policy penetration rate. We compare the post-policy penetration rate to the no-subsidy penetration rate when evaluating the impact of hypothetical policies. Our model implies that the predicted pre-subsidy equilibrium price p^{pre} is the implicit solution to

$$p^{pre} = \widehat{mc} + \Delta(p^{pre})^{-1} s(p^{pre})$$

where the matrix $\Delta_{j,r} = \frac{\partial s_{rm}(p,d)}{\partial p_{jm}} I_{j,r}$ with $I_{j,r}$ an indicator function equal to one when plans j and r are offered by the same ISP, $\widehat{mc}$ are predicted marginal costs, and p^{pre} is the vector of pre-subsidy prices.[30]

[30] We use a market fixed-point iteration (see Conlon and Gortmaker (2020)) to compute the new equilibrium. If it did not converge, we solved the system of equations by the Newton method. In the counterfactuals, we evaluate 929 markets from 2014:4.

6.1 Impact of Implemented Price Subsidy

We begin by exploring the impact of the implemented price subsidy on adoption decisions. Therefore, we use our estimated parameters to compute the pre-subsidy penetration rate. As we discussed in section 2, price subsidies are district specific (and are represented by d in equation 3). The number of subscribers (reported by the ISPs) corresponds to those that subscribed *after* the price subsidy was in place. Therefore, to evaluate the impact of the subsidy we have to simulate how many households *would have* subscribed in the absence of the subsidy. We implement this in our model by (i) removing the discount from the subsidy (i.e, setting $d = 0$) while (ii) setting the subsidy variable equal to zero in the diffusion equation and computing the implied diffusion rate, (iii) computing a new equilibrium price response as detailed above, and (iv) calculating the corresponding probability that households adopt.

	Penetration Rate		Change in Penetration Rate (pp)		Decomposition	
	Observed	Predicted	Supply Response?		% of (3) Due to Impact of	
	Post subsidy	Pre-subsidy*	Yes	No	Price	Diffusion
	(1)	(2)	(3)	(4)	(5)	(6)
Stratas 1 & 2	29.0	16.0	13.0	15.2	39.0	61.0
Strata 1	11.8	2.3	9.5	10.7	45.8	54.2
Strata 2	44.2	28.1	16.2	19.1	35.5	64.5
Panel A: Strata 1						
Low Coverage	1.2	0.1	1.2		73.4	26.6
Mid Coverage	3.7	0.3	3.3		65.9	34.1
High Coverage	18.0	3.7	14.4		56.6	43.4
Panel B: Strata 2						
Low Coverage	6.5	2.2	4.2		66.3	33.7
Mid Coverage	18.3	7.3	11.0		65.8	34.2
High Coverage	52.2	33.9	18.3		45.3	54.7

Notes: *Includes the strategic price response. pp denotes percentage points.

Table 9: Impact of Pricing Subsidy on Adoption

The counterfactual pre-price subsidy probability of adopting is presented in column (2) of Table 9. The first column contains the probability of adoption after the subsidy was in place and is determined by dividing the observed number of subscribers by the market size. The table shows that the price subsidy had a moderate impact on adoption on average. Prior to the subsidy, our estimates imply that 16% of lower strata households had internet access.[31] Penetration increased to 29% after the subsidy.

[31] Our pre-subsidy penetration rate estimates are not significantly different from the observed pre-subsidy penetration rates (which was 15.7%).

The Colombian government spent approximately 78 Million US$ on pricing subsidies over the two year period. Our estimates on the impact on penetration suggest the result was almost 1 million new subscribers - the bulk of which were from strata 2. This implies the cost of the pricing subsidy was approximately 10.50$ per quarter per new customer. However, columns (3) and (4) show that the passthrough of the subsidy to consumers was not complete. That is, if firms would not have strategically reacted to the subsidy by changing their prices, the penetration rate would have increased by 15.2% points (Column 4) instead of only 13% points (Column 3). Comparing columns (3) and (4) shows that the strategic prices changes were split among strata, but were even greater for strata 2. In addition, the impact of the pricing subsidy varies greatly across municipalities (estimates not shown), where in many municipalities the subsidy had no effect on take-up in the lowest-SES strata 1.

To further explore this issue, we examine the impact within economic strata. To do so, we divide each strata into three groups with differing levels of internet knowledge according to the pre-subsidy levels of adoption. These three groups are determined according to terciles of low, medium, and high coverage in the pre-subsidy predicted scenario. This allows us to evaluate the evolution of service adoption bearing in mind pre-existing differences. The lower panel of Table 9 shows that the largest percentage point increase was in markets that had higher initial coverage in both strata.

Overall, the evidence for the price subsidy in closing the digital divide is mixed. On the one hand, the policy increased the probability that low income households adopted, but the level effect of the subsidy was smaller among those households in markets with the lowest initial coverage. Hence, the subsidy does not help to close the digital gap in the sense that it shows the largest percentage point increase in the terciles that have larger average initial penetration rates, but it is closing the gap in the sense that the probability of adoption increases among low SES households across all markets.

It is also worthwhile to examine how much of the total effect of the subsidy on penetration rates can be attributed to (i) the price reduction itself, and how much to (ii) the impact of the subsidy on the diffusion rate. To address this question, we conduct a series of counterfactual analyses to decompose the effect of the price subsidy into these two components. Starting from the predicted pre-subsidy scenario (Column 2 of Table 9), we compute the decomposition in two steps. First, we compute the penetration rate if the subsidy only impacted the diffusion rate and did not otherwise reduce the price paid by the consumers. The comparison of the penetration rate between this auxiliary scenario and the pre-subsidy case amounts to the effect of the subsidy on adoption due to a change in the diffusion rate. In the second step, starting from the auxiliary scenario, we hold the diffusion rate at the pre-subsidy level while allowing consumers to benefit from a lower monthly fee from the subsidy. The comparison of the post-policy penetration rate to the auxiliary penetration rate is the impact of the subsidy policy on adoption through the price reduction.

Columns (5) and (6) show that the effect of the price subsidy on the diffusion rate is more important than the impact of reducing the monthly fee (i.e., the impact on utility directly). The last column shows that 61% of the increase in adoption is due to the impact of the subsidy on the diffusion rate. The analysis across groups of markets suggests that the price reduction is a sizeable part of the total effect for households in markets with the lowest initial coverage, but that the impact on the diffusion rate plays a prominent role as well. Taken together, this evidence suggests that policies aiming at increasing the diffusion rate can be effective at closing the digital divide.

6.2 Impact of Alternative Policies

Our parameter estimates indicated that, in addition to the price, plan characteristics, ISP seniority, and the diffusion rate are important for take-up decisions. To understand the magnitude of these effects we conduct policy counterfactuals. Table 10 presents the results.

The second column gives the probability of adoption under status quo (prior to the price subsidy) (i.e., column 2 from Table 9), and Column 3 gives the penetration rate change from the subsidy (i.e., column 3 in Table 9) to ease comparison with the counterfactual policies presented in the remaining columns.

Our first counterfactual examines the impact of subsidizing more plans. As discussed earlier, ISPs do not offer the same plans to all markets. This alternative policy requires ISPs to offer plans that meet the lowest criteria for broadband (speed of 1-2 Mbps) to consumers in all municipalities that it currently serves. Expanding the choice set results in higher

Tercile	Penetration rate (%) No Subsidy	Change in Penetration Rate (pp)			
		Current Subsidy	Subsidy to more plans	All firms are Established ISPs	Diffusion 80%
Strata 1 & 2	16	13.03	13.57	0.91	18.96
Panel A: Strata 1					
Low Coverage	0.1	1.2	1.2	0	16.5
Mid Coverage	0.3	3.3	3.5	0.1	13.4
High Coverage	3.7	14.4	15.2	0.4	24.6
Total	2.3	9.5	10	0.3	20.7
Panel B: Strata 2					
Low Coverage	2.2	4.2	4.6	1.2	18.7
Mid Coverage	7.3	11	11.8	2.5	25.1
High Coverage	33.9	18.3	18.8	1.4	16.3
Total	28.1	16.2	16.7	1.5	17.4

Table 10: Impact of Alternative Policies on Adoption

take-up but by less than 1 percentage point. The lower panels show that the effect is evenly distributed across strata.

Our demand estimates also indicated that whether the ISP is established is important to consumers, perhaps proxying for aspects related to the functionality or reliability of the plan. The fourth column gives the probability of adoption if all ISPs were treated as established instead of a price subsidy. This counterfactual mimics the positive features of an established ISP while not changing the other ISP characteristics nor the number of competitors in the market. The results show, that while this feature is important to consumers, it does not drive take-up decisions. Adoption would increase only marginally (by less than 1% point). However, this effect is more beneficial (relative to a price subsidy) for households with lower to mid-levels of adoption pre-subsidy so, in this sense, is beneficial.

The final column shows the impact on the penetration rate if, instead of a pricing subsidy, the diffusion rate was increased by 80%. The lower panels of Table 10 indicate this has the biggest impact on the lower income strata households, which is particularly notable when compared to the opposite effect of the price subsidy. Also, in contrast to the pricing subsidy, diffusion rate increases are more evenly distributed within each strata. In particular, it benefits those households who had low initial coverage much more than other households within the strata.

An increase of diffusion by 80% is rather arbitrary, so we present the impact on penetration from gradual increases in the diffusion rate in Figure 3. As the figure shows, in

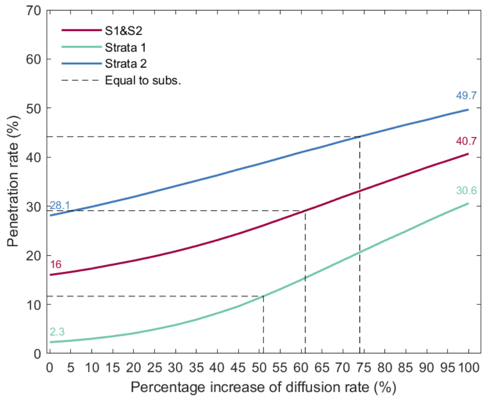

Figure 3: Impact of Alternative Diffusion Rates on Adoption

order to impact penetration as much as the price subsidy, the diffusion rate would need to be increased by around 60% across both strata or by around 50% for households in strata 1. We next examine how this can be accomplished.

6.3 Increasing Diffusion

Our parameter estimates indicate that there are many ways to increase the diffusion rate, with varying degrees of impact on the penetration rate. The model estimates from Table 8 show that providing training and equipment via internet community kiosks (Lit & Eqiup: Kiosks) has the largest impact on the diffusion rate.

Figure 4 presents the impact on the penetration rate from a counterfactual where we increase the number of Internet kiosks. First, we note that, unlike pricing subsidies, there is a fixed cost to providing internet kiosks, with benefits spread over time. The x-axis shows the number of quarters since the internet kiosks were installed. The implied change in the penetration rate is on the vertical axis. The curves represent the impact of kiosks on adoption where we allow for the quality of the kiosks to depreciate at rate τ. The long dashed line shows the number of quarters it would take after the kiosks were installed to generate a change in the penetration rate of the same magnitude as the price subsidy (13 points) for different levels of depreciation. For example, at zero depreciation, it takes 5 quarters for the kiosks to increase penetration as much as the price subsidy.

However, given the fixed costs of setting up an internet kiosk, the price per new customer is high. The most cost effective point occurs after 11 quarters at a cost of 238 US$ per new customer. Although we also note that the benefit of the kiosks may spill over into other strata leading to higher cost effectiveness of the kiosks. Furthermore, even when we don't consider the spillovers, the counterfactual indicates UNESCOs target of 75% penetration (which corresponds to an increase in penetration of 59 points) would have taken less than 7 years to reach for the low-SES population (at zero depreciation).

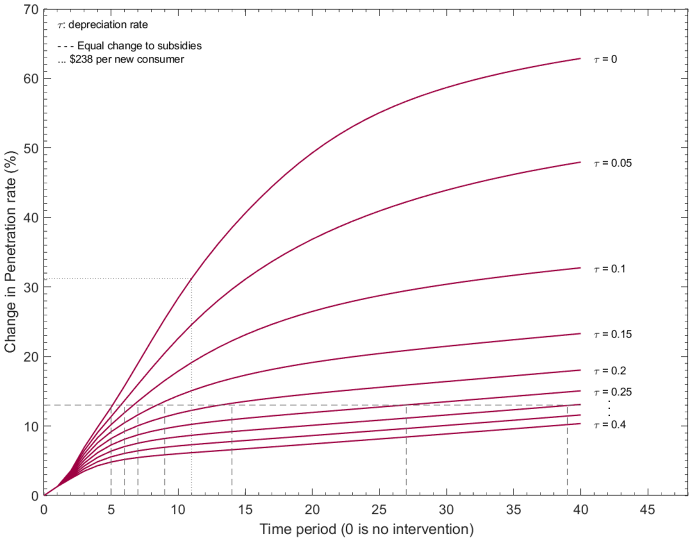

Figure 4: Impact of Kiosks and Centers on Adoption

In summary, we find that increasing the diffusion rate directly (via literacy tools) is more effective than the price subsidy at reaching the most vulnerable among low-SES households. And it is more effective at reaching households that are less technically knowledgeable. However, this comes at an increased cost relative to a price subsidy, albeit one that has positive spillovers to other strata.

7 Conclusion

The provision of affordable access to Information and Telecommunication Technologies is a fundamental priority mandated by the United Nations (United Nations, 2015), and studies show this leads to increased productivity (Thompson and Garbacz, 2011), higher educational outcomes (Hepp et al., 2004) and stimulates economic activity (Qiang and Rossotto, 2009). Therefore, it is vital to understand the impact of interventions on take-up as well as what prevents adoption, especially in developing nations.

We examine a subsidy policy aimed at increasing the adoption of internet services for low-SES households in Colombia. To gauge the impact on market coverage we conduct counterfactual experiments using estimates obtained from a structural model of demand for internet access among low-SES individuals. The model incorporates the influence that your neighbors' connection may have on your take-up decisions (while accounting for the potential endogeneity). Our method allows us to decompose the policy impact into the impact due to lower prices directly and that arising from the impact from the social network.

The evidence for the price subsidy in closing the digital divide is mixed. On the one hand, the policy increased the probability that low-SES households adopt by 13% points, but the impact was less pronounced for the poorest of the poor. Our results indicate a policy targeted at providing more plans as well as increasing the diffusion of services via internet literacy plans is even more effective at closing the digital divide in markets with the poorest populations, and in markets with the lowest connection rates pre-subsidy, although these policies are more costly in the short-run. Our findings are particularly salient for developing countries, where the benefits from the internet economy are more unevenly distributed among SES classes. Our results suggest that increasing the diffusion of services via access and literacy campaigns is the most effective way to reach UNESCOs connection goal - effectively closing the digital divide in Colombia in under a decade.

References

Ackerberg, D. A., DeRemer, D. R., Riordan, M. H., Rosston, G. L., and Wimmer, B. S. (2014). Estimating the impact of low-income universal service programs. *International Journal of Industrial Organization*, 37:84–98.

Ackerberg, D. A. and Gowrisankaran, G. (2006). Quantifying equilibrium network externalities in the ach banking industry. *The RAND Journal of Economics*, 37(3):738–761.

Andrés, L., Cuberes, D., Diouf, M., and Serebrisky, T. (2010). The diffusion of the internet: A cross-country analysis. *Telecommunications Policy*, 34(5):323 – 340.

Arieli, I., Babichenko, Y., Peretz, R., and Young, H. P. (2020). The speed of innovation diffusion in social networks. *Econometrica*, 88(2):569–594.

Belloc, F., Nicita, A., and Rossi, M. A. (2012). Whither policy design for broadband penetration? evidence from 30 oecd countries. *Telecommunications Policy*, 36(5):382–398.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–90.

Bhattacharya, D., Dupas, P., and Kanaya, S. (2023). Demand and Welfare Analysis in Discrete Choice Models with Social Interactions*. *The Review of Economic Studies*. rdad053.

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

Board, S. and Meyer-ter-Vehn, M. (2021). Learning Dynamics in Social Networks. *Econometrica*, 89(6):2601–2635.

Briglauer, W., Frübing, S., and Vogelsang, I. (2015). The Impact of Alternative Public Policies on the Deployment of New Communications Infrastructure. A Survey. *Review of Network Economics*, 13(3):227–270.

Chaudhuri, S., Goldberg, P. K., and Jia, P. (2006). Estimating the effects of global patent protection in pharmaceuticals: A case study of quinolones in india. *American Economic Review*, 96(5):1477–1514.

Chiou, L. and Tucker, C. (2020). Social distancing, internet access and inequality. Technical report, National Bureau of Economic Research.

Conlon, C. and Gortmaker, J. (2020). Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics*, 51(4):1108–1161.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628.

Galperin, H. and Ruzzier, C. A. (2013). Price elasticity of demand for broadband: Evidence from latin america and the caribbean. *Telecommunications Policy*, 37(6-7):429 – 438.

Geroski, P. (2000). Models of technology diffusion. *Research Policy*, 29(4-5):603–625.

Goldfarb, A. and Prince, J. (2008). Internet adoption and usage patterns are different: Implications for the digital divide. *Information Economics and Policy*, 20(1):2–15.

Goolsbee, A. and Klenow, P. J. (2006). Valuing consumer products by the time spent using them: An application to the internet. *American Economic Review*, 96(2):108–113.

Greenstein, S. and Prince, J. (2006). The Diffusion of the Internet and the Geography of the Digital Divide in the United States. Nber working paper 12182.

Grigolon, L. and Lasio, L. (2023). Biased beliefs and stigma as barriers to treatment and innovation adoption. CEPR Discussion Papers 17938, C.E.P.R. Discussion Papers.

Gruber, H. and Verboven, F. (2001). The evolution of markets under entry and standards regulation. *International Journal of Industrial Organization*, 19(7):1189–1212.

Guiteras, R., Levinsohn, J., and Mobarak, A. M. (2019). Demand estimation with strategic complementarities: Sanitation in bangladesh. *Available at SSRN 3328509*.

Hausman, J. A., Sidak, J. G., and Singer, H. (2001). Cable modems and dsl: Broadband internet access for residential customers. *American Economic Review*, 91(2):302–307.

Hepp, K., Hinostroza, E., Laval, E., and Rehbein, L. (2004). Technology in schools : Education, ict and the knowledge society. Technical report, World Bank.

Hidalgo, J. and Oviedo, J. D. (2014). The impact of broadband quality standards on internet services market structure in colombia. International Telecom Society Conference.

Hjort, J. and Poulsen, J. (2019). The arrival of fast internet and employment in africa. *American Economic Review*, 109(3):1032–79.

Jordán, V., Galperin, H., and Peres, W. (2013). *Broadband in Latin America : beyond connectivity*. Santiago de Chile.

Lee, R. S. (2013). Vertical integration and exclusivity in platform and two-sided markets. *American Economic Review*, 103(7):2960–3000.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.

Nevo, A., Turner, J. L., and Williams, J. W. (2016). Usage-based pricing and demand for residential broadband. *Econometrica*, 84(2):411–443.

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–1057.

PEW Research Center (2015). Spring 2015 global attitudes survey. Technical report.

Qiang, C. and Rossotto, C. (2009). Economic impacts of broadband. Technical report, World bank.

Ryan, S. and Tucker, C. (2006). Heterogeneity and the dynamics of technology adoption. *Quantitative Marketing and Economics*, 10.

Rysman, M. (2004). Competition between networks: A study of the market for yellow pages. *Review of Economic Studies*, 71:483–512.

Rysman, M. (2019). The reflection problem in network effect estimation. *Journal of Economics & Management Strategy*, 28(1):153–158.

Savage, S. and Waldman, D. (2009). Ability, location and household demand for internet bandwidth. *International Journal of Industrial Organization*, 27(2):166–174.

Sovinsky, M. (2008). Limited information and advertising in the u.s. personal computer industry. *Econometrica*, 76(5):1017–1074.

Sovinsky, M. and Hidalgo, J. (2022). Subsidies, speed and switching? impacts of an internet subsidy in colombia. Technical report, CEPR Press Discussion Paper.

Thompson, Jr., H. G. and Garbacz, C. (2011). Economic impacts of mobile versus fixed broadband. *Telecommunications Policy*, 35(11):999–1009.

Tucker, C. (2008). Identifying formal and informal influence in technology adoption with network externalities. *Management Science*, 54(12):2024–2038.

UNESCO (2019). State of broadband report 2019. Geneva: International Telecommunication Union and United Nations Educational, Scientific and Cultural Organization.

UNICEF (2020). Covid-19: Are children able to continue learning during school closures?

United Nations (2015). Transforming our world: The 2030 agenda for sustainable development.

Varadhan, R. and Roland, C. (2008). Simple and globally convergent methods for accelerating the convergence of any em algorithm. *Scandinavian Journal of Stat*, 35(2):335–353.

Vélez-Velásquez, J. S. (2019). Merger effects with product complementarity: Evidence from colombia's telecommunications. *Information Economics and Policy*, 49:100831.

Vélez-Velásquez, J. S. (2020). Banning price discrimination under imperfect competition: Evidence from colombia's broadband. *Borradores de Economía*, 1148.

Walsh, C. (2020). Social impacts of new radio markets in ghana: A dynamic structural analysis. working paper.

World Bank (2020). Who on earth can work from home? Technical report. Policy Research Working Paper;No. 9347.

World Economic Forum (2020). Covid-19 could widen the digital gap. https://www.weforum.org/agenda/2020/07/covid-19-could-widen-the-digital-gap-here-is-what-is-needed-now/.

A Linking Subsidized Prices to Plans

We do not have plan-strata level data on how many consumers have a subsidized plan. Instead we have municipality ISP-level data. We link prices to plan-stratas by comparing the number of subsidies granted (from MinTIC) with the number of observed subscribers. For example, Telmex sold subsidized plans to 6,000 subscribers in Medellin. If Telmex had 6,000 subscribers in Medellin, we assume that all Telmex broadband plans in Medellin were subsidized. However, Telmex has about 28,000 subscribers in Medellin. We then compare at the plan-level. If the number of consumers subscribed to, say, plans with 1Mbps is close to the number of granted subsidies, we assume that the provider subsidized only this plan.[32] We repeat the same method sequentially increasing the speed of the plans.[33] In a couple of cases the procedure is not informative, so we assume that the subsidies were granted to plans with speed 1-2Mbps. These Internet plans just fulfill the requirement of the intervention and are likely to be the first products subsidized by the Internet providers. Our approach fits the data well. A projection of the (log of) the number of granted subsidies on the (log of) assigned subsidies, yields a coefficient of 0.95 (not statistically different from 1). We also examined the fit at the municipality level, for all but one region (the NE) the approach fits well. In the northeast, the share is slightly higher using the heuristic approach (so our method to allocate the subsidies is an upper bound for the NE).

[32] We define closeness based on the ratio of the number of observed subscribers to the number of consumers with subsidy. In particular, the operator Telmex grants the subsidies to all consumers subscribed to Internet plan 1Mbps if this ratio is less or equal to 1.1.

[33] In the order 1Mbps; 1-2 Mbps; 1-5Mbps; 1-10Mbps; and all broadband.

B Online Appendix

This is the Online Appendix referred to in the manuscript: "Internet (Power) to the People: How to Bridge the Digital Divide" by Julian Hidalgo (KU Leuven) and Michelle Sovinsky (Univ of Mannheim).

B.1 Price Data Cleaning Details

We had a few data cleaning issues. First, 219 prices show omitted/extra zeroes in the local currency. For example, the price of a plan with 1 Mbps speed from ISP Axesat in Yopal is over 4 million COP (about 2,000 USD). Prices in other jurisdictions for the same plan-ISP are around 40,000 COP. We assumed the 4 million was an error and corrected it to reflect the average price of the same ISP-plan in other areas. Using this method allowed us to account for all 219 discrepancies. Second, we drop observations with prices less than 15 thousand COP (7 USD) as they were offered mostly by the ISP Azteca. This provider was in charge of the deployment of the fiber backbone and provided services at a reduced, even zero, price. Third, we observed a large increase in price for some of the plans offered by Colombia Telecomunicaciones (Telecom) post 2013:2. Using the Internet Archive, we contrasted these prices to the ones posted on the operator's website in those years. The issues was that the operator reported fares related to telecom bundles instead of those of residential Internet services starting in 2013:2. To address this issue, we replaced the prices for the period 2013:2-2014:4 with those reported on the operator's website when we can determine it (at the plan level), when we cannot we use the price of the plan from 2013:1 adjusted for average changes in time over prices. The adjustment we obtain from the Internet Archive which shows a 16% decrease for the Internet plans with a connection speed between 2-3 Mbps, and a 7% price reduction for all plans with a connection speed greater than 3Mbps. Finally, we discard the observations in the top 5% of the price distribution whose monthly fee is greater than 57 USD, which would constitute more than 15% of the average monthly income of households categorized in strata 1 and 2.

B.2 Reduced-Form Estimations

To estimate the event study, we include a series of treatment leads and lags in our baseline regression (strata 1 and 2). Figure 5 displays the results of the event study.

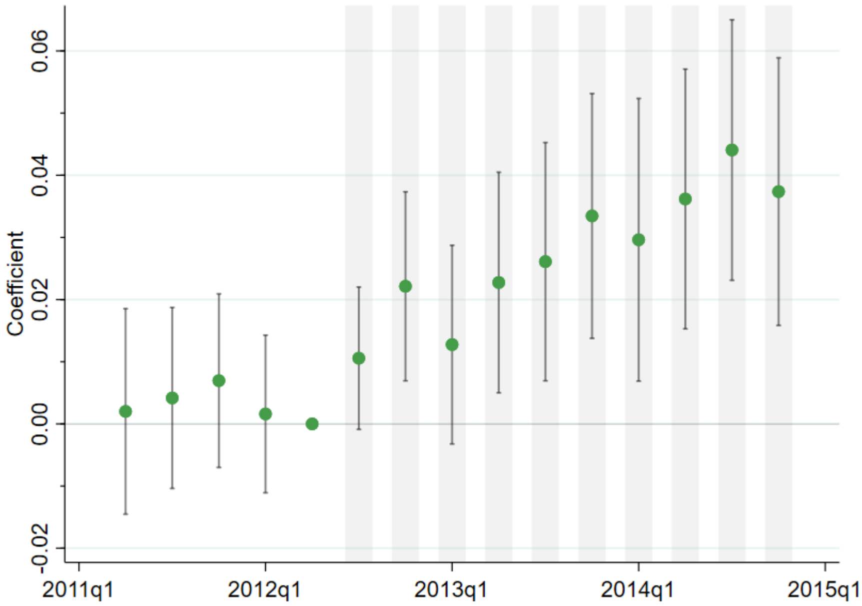

Figure 5: Event study: Strata 1 and 2

The omitted category is the time period right before the start of the subsidy policy. Hence, each time-specific coefficient denotes the changes in the penetration rates in treated markets relative to untreated ones in each time period, as measured from the time period immediately prior to the policy. We can see that during the pre-policy period, the coefficients are small and not statistically significant. After the subsidies were implemented, the penetration rate of treated markets relative to untreated ones increased by an average of 1 percentage point. However, the effect is only significant at the 90% confidence interval.[34]

[34] The associated p-value indicates that it is significant at the 10% level.

B.3 First-Stage Instrumental Variables Regressions

Recall that, for price we use cost instruments and for the diffusion rate we use Colombian conflict variables. Both sets of instruments were described in the main paper in Section 4. Table 11 shows the first-stage regressions and statistics used to assess the validity of the instrumental variables. In all specifications, the weak-IV statistic is greater than 10, thereby indicating that the instruments are not weak. The P-value associated with the Hansen J-Test indicates that we cannot reject the hypothesis that the overidentifying restrictions are valid, except in the case of specification (6).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Price - subsidy	-0.29***	-0.28***	-0.28***	-0.28***	-0.28***	-0.33***	-0.36***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Price - subsidy × Strata 2						0.07**	0.15***
						(0.03)	(0.04)
Diffusion Rate	0.99***	0.86***	0.86***	1.20***	1.20***	0.98***	1.20***
	(0.26)	(0.24)	(0.24)	(0.26)	(0.25)	(0.24)	(0.27)
Diffusion Rate × Strata 2				-0.58***	-0.58***	-0.37***	-0.67***
				(0.11)	(0.11)	(0.08)	(0.13)
Mid-speed	2.13***	2.14***	2.14***	2.14***	2.14***	2.12***	2.10***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
High-speed	3.12***	3.06***	3.06***	3.00***	3.00***	3.11***	2.93***
	(0.24)	(0.23)	(0.23)	(0.24)	(0.24)	(0.24)	(0.25)
Tech: Cable	1.19***	1.19***	1.19***	1.20***	1.20***	1.19***	1.22***
	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)
Tech: xDSL	1.14***	1.14***	1.14***	1.15***	1.15***	1.15***	1.19***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
Seniority (No Established)	0.08***	0.08***	0.08***	0.09***	0.09***	0.08***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Seniority (Established)	0.20***	0.19***	0.19***	0.20***	0.20***	0.20***	0.20***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Trend	-0.20***	-0.19***	-0.19***	-0.19***	-0.19***	-0.19***	-0.18***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-7.52***	-7.20***	-7.21***	-8.99***	-8.99***	-6.93***	-7.43***
	(0.95)	(0.91)	(0.91)	(1.01)	(1.00)	(0.91)	(1.00)
Strata 2	0.93***	0.98***	0.97***	3.73***	3.73***	1.53***	1.52***
	(0.10)	(0.09)	(0.09)	(0.52)	(0.52)	(0.52)	(0.56)
First-Stage Diffusion rate							
Lag Avg Neighbour Ambush	-0.83***	-0.82***	-0.82***	-0.85***	-0.85***	-0.85***	-0.85***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Lag Assault		-0.09***	-0.10***	-0.11***	-0.11***	-0.11***	-0.11***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag Ambush			-0.04***		-0.04***	-0.04***	-0.04***
			(0.01)		(0.01)	(0.01)	(0.01)
Lag Avg Neighbour Ambush × Strata 2				0.05*	0.05*	0.04	0.04
				(0.03)	(0.03)	(0.03)	(0.03)
Lag Assault × Strata 2				0.03***	0.03***	0.03***	0.03***
				(0.01)	(0.01)	(0.01)	(0.01)
First-Stage Price							
Network Capacity Cost	-0.08***	-0.08***	-0.08***	-0.08***	-0.08***	-0.04***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Network Capacity Cost × Speed	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Network Capacity Cost × Strata 2						-0.05***	
						(0.01)	
Network Capacity Cost × Speed × Strata 2						0.01**	-0.00
						(0.00)	(0.00)
Observations	44515	44515	44515	44515	44515	44515	44515
Weak IV (F-test)	36.59	30.08	24.18	20.05	17.27	12.51	16.47
J-stat (pval)	0.029	0.034	0.080	0.119	0.211	0.004	0.210

Notes: All estimates are based on regressions with fixed effects for Internet provider and municipality. The diffusion rate variable is scaled by 100. The weak IV statistic corresponds to the Kleibergen-Paap F statistic, and J-stat denotes the p-value for the Hansen J test. Robust standard errors are reported in parentheses. $***p < 0.001, **p < 0.05, *p < 0.1$.

Table 11: First-Stage Estimates for Price and Diffusion Rate Instruments