Demand for Smartphones and Digital Divide

Scientific Conference "Understanding and Addressing Digital Inequalities"

20-21 November 2025

Toker Doganoglu, Lukasz Grzybowski, Ambre Nicolle

University of Würzburg, University of Warsaw, University of Munich

Motivation

- Mobile communication and internet access are critical drivers of market efficiency and economic growth (See Jensen (2007, QJE), Aker (2010, AEJ), Hjort and Poulsen (2019, AER)).
- Smartphones remain expensive relative to income levels in many developing countries.
- The digital divide in access to devices and the Internet persists: rich vs. poor, urban vs. rural populations.

Motivation

- What policies can improve smartphone adoption (access to mobile Internet) and reduce the digital divide?
- Understanding differences in adoption of smartphones between poor and well-off consumers.
- Trade-off between cost of devices and cost of data in the adoption of smartphones.
- Is the decision to purchase a smartphone a static or dynamic problem?

Our paper

- We use panel data of subscribers from **one** mobile operator in South Africa to study decisions to switch from feature phones to smartphones.
- If the decision problem to adopt a smartphone is dynamic then it is not easy to model for individual-level data when the state space is large (discrete choice dynamic programming models).
- We use the approach of **De Groote and Verboven (2019, AER)** to estimate a dynamic adoption model in application to aggregate data.
- They follow Hotz and Miller (1993) and Scott (2013) and simplify the dynamic decision model to a simple linear equation, which can be estimated using a standard GMM estimator accounting for the endogeneity of prices.

Data: a sample of mobile subscribers

- **Subscribers data:** 300,000 bi-annual observations of $\approx$ 85,000 subscribers between March 2016 and September 2018 from one operator with full country coverage.
- We observe the exact devices used by them and focus on mobile phone users $\approx$ 204,000 observations.
- Next, we focus on users of prepaid SIM cards $\approx$ 185,000 observations who must pay full market price for a mobile phone.

Data: coverage, income and smartphones characteristics

- **Coverage:** of 2G, 3G, and 4G networks at the 'main place' (2,434 geographic areas) with a large variation for 4G roll-out.
- **Average household income:** for 5090 sub-places.
- **Historical handset prices:** from IDC with gaps filled by model-level interpolation.
- **Handset characteristics:** from Imei.info ($\approx$146,000 observations left for 53,110 consumers).

Data: switching patterns in the matched dataset

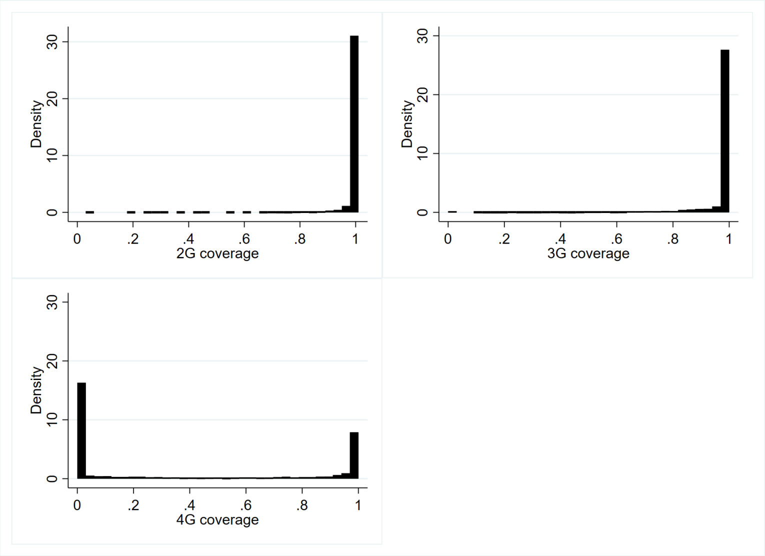
	Freq.	Percent
Stays with smartphone	23,612	48.2%
Stays with feature phone	19,666	37.0%
Upgrade to smartphone	4,370	8.2%
Downgrade to feature phone	1,909	3.6%
Other switching patterns	1.553	2.9%
Total	53,110	100%

- Observations correspond to unique individuals.
- We keep consumers who: (i) never adopted a smartphone, or (ii) switched from a feature phone to a smartphone.
- We use this data to compute market shares of devices for 5 income groups split by rural and urban areas over time.

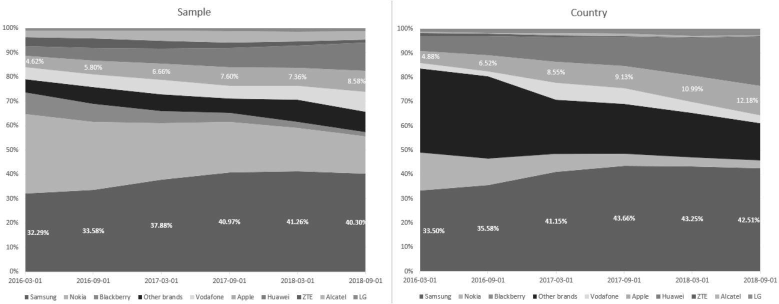
Data: smartphone use and coverage in the prepaid dataset

Variable(%)	March2016	Sep2016	March2017	Sep2017	March2018	Sep2018
Smartphone penetration	45.41	49.58	53.94	57.34	61.27	65.3
Living in urban area	84.35	85.06	85.20	85.39	86.16	86.07
2G coverage	99.41	99.45	99.50	99.48	99.45	99.41
3G coverage	97.07	97.21	98.34	98.33	98.53	98.47
4G coverage	56.68	57.85	76.67	76.67	80.17	80.28

Data: coverage by geographic area



Data: manufacturers' market shares



- In our data, 394 unique handsets belong to 20 brands. We group small brands and define a total of 10 main brands.
- On the figure, other brands include Motorola, Sony, Xiaomi and other small brands. The percentages indicated are for Samsung (bottom) and Apple (top). The source for country data is Statscounter.

Approximating income

- South African society is highly unequal $\Rightarrow$ large differences in income live in segregated neighborhoods with no mobility between them.
- Income levels of locations can be identified at a *sub-place* resolution $\Rightarrow$ there are 5090 *sub-places*.
- We can identify *sub-places* of subscribers live using the most frequently connected antenna information.
- We define five income groups for data aggregation:
  - Below the poverty line
  - Four income groups above the poverty line
- And in split by urban vs. rural locations.

Data aggregation

- **Product markets (choice sets) are defined based on income groups and location (urban vs. rural).**
- We define 8 markets $\Rightarrow$ 5 income groups in urban areas and 3 poorest income groups in rural areas (the two richest income groups are not represented in rural areas).
- We investigate whether income/geography groups have different price responsiveness by estimating separately two models:
    - Model (1) includes 2 poorest groups in urban and 2 poorest groups in rural areas.
    - Model (2) includes 3 richest groups in urban and the middle-income group in rural areas.

Choice model

- The adoption of a smartphone is terminating state $\Rightarrow$ If a consumer purchases a smartphone, she exits the market.

- The present value of using the smartphone (i consumer, j product, t time)

$$\underbrace{\gamma X_{jt} - \alpha p_{jt} + \theta c_t + \xi_{jt}}_{\delta_{jt}} + \epsilon_{ijt}$$

- X_{jt} are smartphone characteristics, p_{jt} is the price of the smartphone (and mobile data), c_t is LTE coverage, ξ_{jt} are unobserved characteristics, and ϵ_{ijt} is the type I extreme value error term.

- If we know δ_{jt}, the choice probabilities or predicted market shares for each alternative are given by logit formulas:

$$S_{jt} = s_{jt}(\delta_t) \equiv \frac{exp(\delta_{jt})}{\sum_{k=0}^{J} exp(\delta_{kt})}$$

Choice model: Value function

- The utility of not adopting a smartphone can be written as:

$$\underbrace{u_{i0t} + \beta E_t(\bar{V}_{t+1})}_{\delta_{0t}} + \epsilon_{i0t}$$

with β the discount factor, and $\bar{V}_{t+1}$ the value of considering to adopt a smartphone at the next date right before the type I extreme value taste shocks are revealed.

- The trick from De Groote and Verboven (2019) following Scott (2013) is to say:

$$\eta_t \equiv E(\bar{V}_{t+1}) - \bar{V}_{t+1}$$

where the expectation error is assumed to be on average zero, and thus:

$$\delta_{0t} = u_{0t} + \beta E(\bar{V}_{t+1}) = u_{0t} + \beta(\bar{V}_{t+1} + \eta_t)$$

Model: value function

- Given the logit structure, expected value of adopting the smartphone before tastes are realized is easy to compute:

$$\bar{V}_{t+1} = 0.577 + \ln \sum_{j=0}^{J} exp(\delta_{jt+1})$$

where 0.577 is again Euler's constant (the mean of the extreme value distribution).

- Based on Hotz and Miller (1993) we can write this as:

$$\bar{V}_{t+1} = 0.577 + \delta_{kt+1} - \ln S_{kt+1}$$

where k is some alternative available at $t+1$.

Model: estimating equation

- Thus, the utility of staying with the feature phone:

$$\delta_{0t} = u_{0t} + \beta E(\bar{V}_{t+1})$$

- can be written as:

$$\delta_{0t} = u_{0t} + \beta(0.577 + \delta_{kt+1} - \ln S_{kt+1} + \eta_t)$$

- and after normalizing $u_{0t} + \beta 0.577 = 0$, as in De Groote and Verboven (2019), takes the form:

$$\delta_{0t} = \beta(\delta_{kt+1} - \ln S_{kt+1} + \eta_t)$$

- We also model u_{0t} as a function of characteristics of feature phones (which vary over time).

Model: estimating equation

- Following Berry (1994) we have:

$$\ln(S_{jt}/S_{0t}) = \delta_{jt} - \delta_{0t}$$

- where:

$$\delta_{jt} = x_j\gamma - \alpha p_{jt} + \theta c_t + \xi_{jt}$$

$$\delta_{0t} = \beta(\delta_{k,t+1} - \ln S_{k,t+1} + \eta_t)$$

Model: estimating equation

- After substituting δ's the estimating equation is:

$$\ln\left(\frac{S_{jt}/S_{kt+1}^{\beta}}{S_{0t}}\right) = (x_{jt} - \beta x_{kt+1})\gamma - \alpha(p_{jt} - \beta p_{kt+1}) + \varepsilon_{jt}$$

where the econometric error term is defined as:

$$\varepsilon_{jt} \equiv \xi_{jt} - \beta(\xi_{kt+1} - \eta_t)$$

- For $\beta = 0$ we have a standard static logit model as in Berry (1994).
- Our instruments for prices: handset prices in UK, LTE coverage.
- We follow De Groote and Verboven (2019) and assume that $\beta = 0.99$ (can be also estimated).

Static demand model: Berry (1994)

	$p_H + p_D$				p_H only			
	OLS		GMM		OLS		GMM	
VARIABLES	Poor	Well-off	Poor	Well-off	Poor	Well-off	Poor	Well-off
price	-0.002	-0.013	-0.860***	-0.352***	-0.022	-0.019**	-0.363*	-0.265**
	(0.017)	(0.008)	(0.186)	(0.047)	(0.020)	(0.008)	(0.188)	(0.108)
mar2017	0.389***	0.308***	0.502***	0.410***	0.389***	0.304***	0.367***	0.287***
	(0.065)	(0.053)	(0.111)	(0.071)	(0.065)	(0.053)	(0.071)	(0.058)
sep2017	0.441***	0.419***	1.438***	0.930***	0.455***	0.420***	0.710***	0.659***
	(0.078)	(0.057)	(0.231)	(0.109)	(0.076)	(0.056)	(0.161)	(0.126)
mar2018	0.531***	0.481***	1.301***	0.834***	0.548***	0.490***	0.773***	0.724***
	(0.080)	(0.059)	(0.231)	(0.122)	(0.079)	(0.059)	(0.188)	(0.149)
handset age: 1 year	0.067	0.066	-0.398*	-0.064	0.053	0.061	-0.217	-0.026
	(0.113)	(0.076)	(0.233)	(0.123)	(0.113)	(0.076)	(0.160)	(0.103)
handset age: 2 year	-0.091	-0.052	-0.977***	-0.527***	-0.115	-0.062	-0.516**	-0.374**
	(0.135)	(0.083)	(0.316)	(0.164)	(0.135)	(0.083)	(0.239)	(0.185)
handset age: 3 year	-0.211	-0.065	-1.117***	-0.562***	-0.236	-0.076	-0.736***	-0.400*
	(0.158)	(0.086)	(0.365)	(0.194)	(0.158)	(0.086)	(0.257)	(0.209)
handset age: 4 year	-0.255	-0.200**	-1.518***	-1.052***	-0.291	-0.218**	-0.975***	-0.838***
	(0.186)	(0.095)	(0.448)	(0.235)	(0.186)	(0.094)	(0.341)	(0.311)
handset age: 5+ year	-0.269	-0.138	-1.829***	-1.229***	-0.311	-0.160	-1.077***	-0.924**
	(0.209)	(0.123)	(0.529)	(0.281)	(0.209)	(0.122)	(0.413)	(0.386)
Os Android	0.236	-0.262	0.100	-0.419*	0.213	-0.274	0.114	-0.497**
	(0.368)	(0.179)	(0.876)	(0.243)	(0.372)	(0.178)	(0.506)	(0.220)
OS Windows	0.362	-0.296*	-0.909	-0.555*	0.323	-0.304*	-0.154	-0.524**
	(0.258)	(0.176)	(0.633)	(0.297)	(0.258)	(0.176)	(0.455)	(0.255)
Constant	-6.300***	-5.501***	-4.156***	-4.884***	-6.242***	-6.478***	-5.943***	-5.619***
	(0.397)	(0.221)	(0.996)	(0.415)	(0.393)	(0.298)	(0.602)	(0.367)
Product dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Brand dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	918	862	918	862	918	862	918	862
R-squared	0.567	0.488			0.567	0.488		

Dynamic demand model: De Groote and Verboven (2019)

	$p_H + p_D$				p_H only			
	OLS		GMM		OLS		GMM	
VARIABLES	Poor	Well-off	Poor	Well-off	Poor	Well-off	Poor	Well-off
price	-0.013	-0.016	-1.327***	-0.435***	-0.030	-0.025**	-1.103***	-0.382**
	(0.019)	(0.010)	(0.452)	(0.135)	(0.020)	(0.010)	(0.405)	(0.150)
handset age: 1 year	-0.000	-0.084	-0.303	-0.110	-0.010	-0.089	-0.222	-0.074
	(0.148)	(0.098)	(0.376)	(0.158)	(0.147)	(0.097)	(0.334)	(0.151)
handset age: 2 year	-0.086	-0.231**	-1.355**	-0.608**	-0.101	-0.243**	-1.154**	-0.525**
	(0.180)	(0.108)	(0.573)	(0.242)	(0.180)	(0.108)	(0.506)	(0.247)
handset age: 3 year	-0.598***	-0.232**	-1.742***	-0.659**	-0.617***	-0.245**	-1.620***	-0.564**
	(0.221)	(0.108)	(0.645)	(0.281)	(0.220)	(0.108)	(0.578)	(0.282)
handset age: 4 year	-0.463*	-0.330***	-2.355***	-1.245***	-0.490*	-0.354***	-2.083***	-1.095**
	(0.256)	(0.119)	(0.897)	(0.419)	(0.255)	(0.119)	(0.796)	(0.436)
handset age: 5+ year	-0.650**	-0.243*	-2.833***	-1.401***	-0.682**	-0.271*	-2.492***	-1.218**
	(0.264)	(0.147)	(1.030)	(0.500)	(0.265)	(0.146)	(0.918)	(0.527)
Os Android	0.022	-0.193	0.359	-0.662**	0.002	-0.203	0.432	-0.648**
	(0.402)	(0.185)	(1.478)	(0.278)	(0.407)	(0.185)	(1.304)	(0.268)
OS Windows	0.217	-0.152	-1.911	-0.692*	0.184	-0.161	-1.492	-0.659**
	(0.228)	(0.185)	(1.197)	(0.359)	(0.229)	(0.186)	(1.056)	(0.335)
Constant	-0.941***	-1.087***	-0.651***	-1.121***	-0.936***	-1.088***	-0.657***	-1.148***
	(0.186)	(0.108)	(0.188)	(0.090)	(0.185)	(0.107)	(0.176)	(0.088)
Product dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Brand dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	918	862	918	862	918	862	918	862
R-squared	0.552	0.465			0.533	0.466		

Counterfactual simulations: Eliminate VAT for smartphones

Table 1: Static model

	Penetration					Simulations				
Rural	inc 1	inc 2	inc 3	inc 4	inc 5	inc 1	inc 2	inc 3	inc 4	inc 5
2016q3	37.7%	41.5%	51.5%			40.3%	43.4%	53.0%		
2017q1	44.6%	45.2%	56.7%			51.6%	51.1%	60.6%		
2017q3	51.6%	50.4%	61.9%			62.3%	61.2%	67.6%		
2018q1	57.7%	55.1%	66.1%			69.7%	68.4%	72.6%		
Urban	inc 1	inc 2	inc 3	inc 4	inc 5	inc 1	inc 2	inc 3	inc 4	inc 5
2016q3	46.4%	48.0%	57.6%	65.4%	73.0%	51.0%	51.3%	60.0%	68.0%	74.6%
2017q1	55.0%	53.2%	65.4%	70.4%	75.3%	64.9%	61.8%	70.8%	76.0%	79.1%
2017q3	63.7%	57.9%	71.9%	74.3%	78.2%	77.0%	70.9%	79.0%	81.5%	84.0%
2018q1	71.8%	62.7%	77.1%	78.6%	80.9%	85.1%	77.7%	84.4%	86.4%	87.6%

Counterfactual simulations: Eliminate VAT for smartphones

Table 2: Dynamic model

	Penetration					Simulations				
Rural	inc 1	inc 2	inc 3	inc 4	inc 5	inc 1	inc 2	inc 3	inc 4	inc 5
2016q3	36.0%	44.1%	51.4%			36.7%	44.0%	51.3%		
2017q1	41.4%	48.9%	55.9%			43.9%	49.5%	56.2%		
2017q3	48.4%	56.6%	62.0%			53.8%	60.8%	62.3%		
2018q1	58.0%	62.4%	65.8%			63.1%	67.3%	66.4%		
Urban	inc 1	inc 2	inc 3	inc 4	inc 5	inc 1	inc 2	inc 3	inc 4	inc 5
2016q3	50.9%	54.2%	56.1%	66.5%	73.7%	51.8%	54.6%	56.2%	66.6%	73.8%
2017q1	61.1%	64.5%	63.9%	71.7%	76.1%	63.7%	65.7%	64.3%	72.1%	76.5%
2017q3	70.0%	70.9%	68.6%	75.6%	80.1%	75.2%	74.5%	69.4%	76.9%	80.7%
2018q1	80.7%	78.3%	73.9%	80.1%	82.9%	85.3%	82.2%	74.9%	81.7%	83.9%

Conclusions (work in progress)

- Our preliminary estimates indicate taking dynamics into account is important $\Rightarrow$ the static model overestimates policy effects.
- We are working on other policy experiments (suggestions?).
- And trying to separate the effect of handset price and data usage price $\Rightarrow$ subsidize handset or usage?
- Our next paper will utilize individual-level data if we find a way to deal with the large state space.