Demand for Smartphones and Digital Divide*

Toker Doganoglu† Lukasz Grzybowski‡ Ambre Nicolle§

March 2023

Work in Progress! Please do not cite!

Abstract

In this paper we estimate a static and dynamic model of smartphone adoption. We construct a unique database of adopters of smartphones with different levels of income in South Africa, which is a developing economy with large income inequality. We use our model to assess the impact of policies which aim at stimulating the adoption of smartphones by people living below the poverty line. We find that the main driver of adoption is coverage by LTE networks, while the price of smartphones has only marginal impact. We conclude that to reduce digital divide it is critical to implement regulations which aims at developing LTE infrastructure in poorer areas. The static and dynamic models yield comparable results suggesting that consumers do not take future price and quality into account when adopting smartphones, while this might be the case when consumers replace their smartphones.

Key Words: *Smartphones; Dynamic demand model; Infrastructure regulation; Income inequality*

JEL Classification: L13, L50, L96

*We are indebted to PriceCheck.za for providing historical data on handset prices. Ambre Nicolle acknowledges funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Sklodowska-Curie Grant Agreement No. 75388 and from LMU Munich's Institutional Strategy LMUexcellent within the framework of the German Excellence Initiative (No. ZUK22). All errors are our own.

†Department of Economics, University of Würzburg, Chair of Economics, especially Industrial Economics, Sanderring 2, 97070 Würzburg Germany. E-mail: toker.doganoglu@uni-wuerzburg.de

‡Telecom Paris, Institut Polytechnique de Paris, 19 place Marguerite Perey, 91120 Palaiseau, France & University of Cape Town, School of Economics, Rondebosch, 7701, Cape Town, South Africa. E-mail: lukasz.grzybowski@telecom-paris.fr

§LMU Munich, ISTO, Kaulbachstr. 45, 80539, Munich, Germany. E-mail: ambre.nicolle@ensai.fr

1 Introduction

Mobile communications offer a major opportunity to advance economic growth in developing countries, where fixed-line infrastructure is non-existent or of limited coverage. Even when fixed networks exist, they are typically available in urban areas and to better-off households who are the minority. Therefore, the majority of population has to rely on mobile networks to access internet. As the main communications technology, mobile phones can stimulate inclusive economic growth, and reduce poverty and inequality through different mechanisms, for example by: saving wasted trips; providing information about prices; improving management of supplies; and increasing productive efficiency of firms (see Aker and Mbiti, (2010)). They can also serve as a channel for provision of services which are in general not available to poor people, such as mobile-based financial, health, educational and agricultural services.[1]

Yet, mobile phones and especially smartphones are still expensive and not affordable to the majority of population in low-income countries. Many poor people in these countries do not have stable jobs or work in the informal sector. Thus, they are not eligible for tariff plans which would enable them to pay off the cost of a smartphone over one or two years. Also, because of limited budgets, they cannot afford purchasing bigger bundles of data and minutes at one time and instead make frequent purchases of small bundles. This increases the average price paid for mobile services. Overall, poor households spend a relatively high share of their income on telecommunications services. Another issue is that 4G mobile networks, which enable high-speed internet access, are first rolled out in densely populated urban areas which tend to be wealthier on average. This further contributes to widening the digital divide between rich and poor, and between urban and rural areas.

In this paper, we use biannual panel data of subscribers of mobile telecommunications services in South Africa to analyze how the price of smartphones impacts the adoption decision in different income groups. This is of key importance because as mentioned above, a smartphone is the only mean of accessing internet by poor people, which has economic and societal consequences. The governments and international organizations such as World Bank

[1]For evidence on the impact of mobile phones on economic development see papers by Jensen (2007), Muto and Yamano (2009) and Aker (2010).

have been trying to design policies to increase smartphone penetration and internet access among poor people in Africa and in other low-income countries. The potential policies are lower taxes on the imports and sale of smartphones, or smartphone subsidy schemes, which could lower the upfront prices paid by poor consumers. However, the effectiveness of these policies cannot be evaluated without knowing how price responsiveness varies across different income groups, which is what we estimate in this paper. It is important to note that the price of smartphones is not the only factor which impacts adoption. There is little value in having a smartphone in areas without coverage with 3G or 4G/LTE technologies, i.e., without mobile internet access. Also, the cost of mobile data may determine the adoption decision. We incorporate these two additional factors in our analyses.

We estimate a number of discrete choice models, where we take into consideration the fact that smartphones are durable goods and consumers are forward-looking. In doing this, we follow De Groote and Verboven (2019) who develop a simple dynamic model of new technology adoption. Similarly to other technological products, the quality of smartphones has been increasing over time (better processor, memory size, quality of camera, etc.), while the quality-adjusted price has been declining. Therefore, modeling demand for smartphones may require a dynamic framework, where in each period consumers decide whether to adopt a smartphone or to postpone this decision until the price goes down enough and quality increases. A static model may underestimate the price elasticities when consumers indeed postpone the adoption decision (see Gowrisankaran and Rysman (2012)). We investigate whether this is the case for the demand for smartphones by comparing dynamic and static adoption models.

We do not have information on the exact income of individuals in our sample. However, in South Africa the suburb in which an individual lives provides a very good approximation of the income level of households because of racial segregation policies implemented during the Apartheid. This spatial segregation has not changed much since then. Even though underprivileged and wealthy suburbs often share a border with one another, there is virtually no mobility between them. The differences in property prices and rents are extreme and people remain segregated due to structural poverty and inequality. Our dataset includes information on the location of mobile antennas to which each individual in our sample was connected most of the time in the time period observed. We use this information to identify the suburbs in which individuals live. Next, we are able to complement our data with a very detailed census data on 5,090 so called 'sub-places'. This allows us to retrieve information about the average household income as well as other statistics, such as the share of population having a fixed-line connection or a computer. We use the income information obtained from this data to construct five income groups. The first group corresponds to individuals living in areas where the income is below the national poverty line. The remaining individuals are divided into four income groups based on the four quartiles of the distribution of income for people living above the poverty line.

We analyze the characteristics of smartphones used by individuals belonging to these five income groups. There are substantial differences in the type, quality and price of smartphones across these groups, suggesting that the alternatives considered by individuals vary widely based on their wealth. Therefore, in our models, we assume that consumers have a different choice set of handsets depending on their income group. In general, consumers with stable jobs and regular streams of income have access to long-term contracts with handset subsidies and can afford more expensive handsets. In the prior literature, the choice sets of consumers are typically specified at the level of a geographic market, or based on different product segments. In this paper, we define separate markets and choice sets for different income groups. We proceed by aggregating the individual data at the income group-level which we use in the estimation. This approach allows us to estimate the income-group-specific price coefficients, which we then use in the counterfactual simulations.

Our results confirm our intuition that there is a digital divide in the population studied because poor people live in areas without network coverage. Surprisingly, the cost of smartphones does not have a significant impact on smartphone adoption. In particular, if poor and richer areas were fully covered by LTE networks in the period we study, the adoption of smartphones would increase from 57.1% to 76.0% among people below the poverty line, while it would only change from 81.5% to 82.2% in the richest group of consumers as of 1st quarter of 2018. This is precisely because richer areas have almost full LTE coverage, while poorer regions are only partially covered. At the same time, removing 15% VAT on smartphones would increase adoption only to 57.7% among people below the poverty line, and to 81.7% among richest consumers as of 1st quarter of 2018. The handset prices seem to have only a marginal impact on adoption in all income groups. We conclude from these findings that it is critical to develop LTE infrastructure in poorer areas to reduce the existing digital divide as people will likely respond by adopting smartphones, irrespective of their income. Our static and dynamic models yield comparable results suggesting that consumers do not take future prices and quality into account when purchasing smartphones.

The remainder of this paper is organized as follows. Section 2 provides a discussion of related literature. Section 3 presents the data used in the estimation. Section 4 introduces the econometric framework. Section 5 presents the estimation results, and finally Section 6 concludes.

2 Literature review

Our paper is methodologically related to the stream of literature focusing on discrete choice dynamic programming models (DCDP). Methods to estimate such models have been reviewed by Aguirregabiria and Mira (2010) and Arcidiacono and Ellickson (2011), and can be divided into two types. The first one includes estimations using full solution methods, i.e., GMM approaches relying on the Nested Fixed Point (NFP) algorithm suggested by Rust (1987), and the BLP contraction mapping proposed by Berry, Levinshon and Pakes (1995). The second type are estimations relying on a simplification of Rust's algorithm, with the seminal paper by Hotz and Miller (1993) and more recent papers by Aguirregabiria and Mira (2002) and Arcidiacono and Miller (2011). This seceond stream of literature relies on the Conditional Choice Probability (CCP) estimators, which is a relatively easy way to estimate dynamic discrete choice models. The estimation methods from both streams of literature can be implemented using individual-level or aggregate market-level data.

In the seminal paper, Rust (1987) formulated a simple regenerative optimal stopping model of bus engine replacement. The optimal stopping rule is the solution to a stochastic dynamic programming problem that formalizes the trade-off between the conflicting objectives of minimizing maintenance costs versus minimizing unexpected engine failures. The paper proposes a "nested fixed point" algorithm for estimating discrete choice dynamic programming models. For each combination of parameters, a fixed point algorithm is used to compute value functions which characterize the expected future utility associated with the choices.

In another seminal paper, Hotz and Miller (1993) develop a method for estimating the structural parameters of DCDP models. Under some conditions, this method avoids the computation of the value functions and is computationally less burdensome. Their method relies on the realization that the value functions can be represented as an easily computed function of the state variables, structural parameters, and the probabilities of choosing alternative actions for states that are feasible in the future. They estimate a dynamic model of parental contraceptive choice and fertility to illustrate application of this method. Subsequent work and extensions of these two seminal papers are surveyed in Aguirregabiria and Mira (2010) and Arcidiacono and Ellickson (2011).

Dynamic discrete choice models are estimated for various choices made by agents such as occupational choices (Miller, 1984), patent renewal (Pakes, 1986) or fertility choices (Wolpin (1984), Hotz and Miller (1993)). One category of decision that is of particular interest for us is the adoption of durable goods, for which prices tend to decrease over time and quality increases.[2] In an earlier paper, Nair (2007) estimates the demand for video-games when consumers forward-looking and firms use inter-temporal price discrimination. In other papers, Schiraldi (2011) estimates a model for new and used cars where consumers decide to replace their automobile when a secondary market exists, and Conlon (2012) estimates a dynamic demand model for LCD TVs. Gowrisankaran and Rysman (2012) estimate dynamic demand for cameras for video recording (camcorders) which experienced rapid price declines and quality improvements at the same time. In this paper, they consider a dynamic decision that involves endogenous repeat purchases over time (similar to Rust (1987)) and when a large number of differentiated products is available on the market (as in BLP (1995)). They highlight

[2]For a discussion about the challenges associated with estimating the demand for durable goods, see Dubé et al (2014).

significant differences in the elasticities of price and characteristics depending on whether the model is static or dynamic, with more intuitive results obtained with the dynamic one. This observation suggests that accounting for dynamic decisions is important and should be implemented when one needs to compute the welfare gains associated with the introduction of these products. The market of smartphones which we study is similar to camcorders in that consumers choose from many differentiated products and they can rationally expect that over time prices will decline and the quality will increase. Thus, consumers face the trade-off between adopting now or waiting for the next period. However, in contrast to their work that focuses on replacement of product, we study the adoption of products.

Our paper follows closely the adoption decision framework derived in De Groote and Verboven (2019). In this paper, they formulate the expected next period ex ante value function as the realized value function plus a prediction error, which is uncorrelated with any variables known by the household at the time of the adoption decision. Next, they show how to invert the demand model to solve for the unobserved error term, which yields a linear regression equation. The current adoption rate then depends on current and next period prices, as well as on the next period adoption rate. The model can be estimated using a standard nonlinear GMM estimator to account for the endogeneity of several variables. They apply the model to a program which promotes adoption of solar photovoltaic systems in Belgium and find that households significantly discounted the future benefits from the new technology. This implies that an upfront investment subsidy program would have promoted the technology at a much lower budgetary cost. The model is estimated using aggregate market data, which is similar to Scott (2017) who uses the CCP estimator developed by Hotz and Miller (1993) to study a change in land use in US where landowners optimize dynamically. His approach avoids the burden of explicitly modeling the evolution of market-level state variables like input and output prices. He then uses the model to estimate long-run cropland price elasticity. He concludes that the estimation of a static, myopic model yields to understating the long-run acreage elasticities, and therefore of the long-run land use effects and environmental costs. In this paper, we use aggregate data and follow the approach by Scott (2017) and De Groote and Verboven (2019), which allows us to address our research question in a relatively simple way. Estimating a dynamic demand model with micro-data is more challenging to implement although it would allow us to account for individual heterogeneity. We leave the estimation of a dynamic model of demand for smartphones with individual-level data and possibile individual heterogeneity to future research.

Finally, our paper contributes to the literature on consumer's choices in the telecommunications markets, which in the last years have been undergoing dramatic technological innovation. This includes deployment of mobile broadband networks (3G, 4G and soon 5G) and launch of smartphones. The usage and adoption of mobile services has been extensively studied as well as the adoption dynamics of mobile services and handsets due to switching costs (e.g. Cullen and Scherbakov (2010), Grzybowski and Liang (2014), Grzybowski and Nicolle (2020)). A novel aspect in our setting is that most consumers already use a mobile phone and decide to upgrade to a new product, which enables them to use mobile broadband services. Only a few papers estimated demand for smartphones in a structural framework with different objectives. For example, Sun (2012) explores the impact of the application stores on the brand value of operating systems. Sinkinson (2014) estimates price elasticities for smartphones and carriers to study the implications of exclusive contracts for smartphones. Hiller, Savage and Waldman (2018) simulate the impact of different hypothetical smartphone patent infringements on equilibrium outcomes. Fan and Yang (2018) explore the relationship between competition on the smartphone market and the number of products offered. They assess the welfare impact of various mergers between manufacturers, accounting for price and quality of products released. Finally, Luo (2018) explores the existence of OS-specific network effects in the smartphone industry. She also assesses the impact of long-term contracts offered by operators in this framework.

The adoption of mobile services in developing countries attracted much interest because of the potential which they have for the economic development (see for example in Aker and Mbiti (2010)). In a more recent paper, Bjorkegren (2019) develops a method to estimate and simulate the adoption of a network good. The demand for mobile phones is modeled as a function of individuals' social networks, coverage and prices. He uses transaction data over 4.5 years for nearly the entire mobile subscribers base in Rwanda. The model is then used to simulate the effects of two policies. First, a requirement to serve rural areas by the operators resulted in lower profits but increased net social welfare. Second, he finds that a shift from handset to usage taxes would have increased the surplus of poorer users by at least 26%.

In another recent paper, Shreeti (2019) uses aggregate data from 2007-2018 to explore the determinants of smartphone adoption in India. She exploits two shocks that occurred in the Indian market over the period of her study. First, the entry of a Chinese manufacturer resulted in a sharp decline of device prices. Second, the entry of a new mobile network operator (MNO) decreased mobile broadband prices. She concludes that the availability of cheap devices was not sufficient to allow for the adoption of smartphones to take-off. Her results suggest that what is needed to stimulate adoption are both cheap smartphones and cheap data services.

Our paper makes the following contributions to the literature discussed above. First, we apply the dynamic technology adoption model proposed by De Groote and Verboven (2019) to the case of smartphones in a developing country with large income inequality. We account for heterogeneity in price responsiveness in different income groups. We then use our estimates to conduct counterfactual simulations and comment on what should be the most effective policies to stimulate adoption of smartphones and access to internet among poor consumers. This is the first paper which sheds light on how consumers from different income groups choose smartphones.

3 Data

Dataset construction We combine different datasets for the purpose of this analysis. There are four mobile network operators in South Africa. We use data from one of them with full geographic coverage. The original dataset consists in 300,000 observations of 85,052 individuals who used mobile services between March 2016 and September 2018. Information about consumers is recorded twice per year in March and September. We observe the model of handset used by these individuals because the SIM card used by a consumer automatically detects and registers the model of handset based on a unique international code called the IMEI (International Mobile Equipment Identity). From this data, we drop all observations on devices which are not phones. This leaves us with 203,720 observations of 65,270 individuals. We also drop consumers who subscribe to postpaid contracts because they are often bundled with a handset that is subsidized by the operator. In South Africa, the majority of consumers use mobile communications services with prepaid SIM cards. The remaining dataset has 185,304 observations.

Next, we merge this data with detailed information on network coverage for 2G, 3G and 4G networks at the main place level (2,434 unique geographic areas). There is substantial variation in coverage across geographic regions, as presented in Figure A.1. While coverage for 2G and 3G has reached very high levels over the whole territory with a few exceptions, there are greater geographic differences in the roll-out of 4G network. We use this variation to identify how the development of high-speed networks, in particular 4G, impacts consumers' utility and smartphone adoption. Finally, we use the codes for sub-places to add information on average household income from the country's National Census of 2011, which covered about 10% of the country population. We loose a few observations at this stage, which leaves us with 181,915 observations.

This data is then merged with historical prices of handsets obtained from a price comparison website and complemented by data purchased from the International Data Company (IDC). For some handsets and periods information on price is missing. We fill the data gaps using linear interpolation at the model level, or if not possible, at the brand-level.[3] After dropping missing observations, we are left with 145,977 observations. We further merge this data with handset characteristics from Imei.info, which does not result in further loss of data.

After this process of merging and cleaning our dataset consists of 145,977 observations and 53,110 unique individuals. We focus our analysis on consumers who in the time period of our data: (i) never adopted a smartphone; or (ii) switched from a feature phone to a smartphone. We drop observations related to consumers who upgraded from a feature phone to a smartphone before the beginning of our sample. Table A.2 shows the number of individuals with different patterns of switching between feature phones and smartphones. For the small share of consumers who upgrade and then downgrade, we drop the observations related to

[3]We provide further details on our approach in the Appendix B.

downgrade, which reduces the sample to 65,556 observations on 26,988 unique consumers observed in the period between March 2016 and September 2018.

For each individual we also retrieve information on the average cost of data in Rands per MB. This is computed from available information on the bills of individuals by dividing data revenues by data usage in each period. The cost of one MB of data varies significantly depending on the selected tariff. We then compute an average price of data over time for consumers belonging to different income groups.

We form income groups using information on the average income at the sub-place level. The first group corresponds to individuals living in areas where the average income is below the national poverty line. In 2019 in South Africa, the poverty line was defined as per individual income of less than 47 US dollars per month. Individuals who live in the remaining areas with average income above the poverty line are divided into four income groups based on the four quartiles of the income distribution above the poverty line. We use these five income groups to define choice sets, as discussed below.

The society in South Africa is multi-racial, multi-lingual and highly segmented with respect to income, which results in differences in the affordability of mobile telecommunications services. The operator market shares vary by income group. Our dataset is for a single operator and might not be representative of consumers in the country. We compare the consumers in our sample with those surveyed in the representative national census, both in terms of socio-demographics, characteristics of the handsets used and coverage.[4] We present descriptive statistics in Appendix C. While our final sample is overall representative in terms of coverage and ethnic composition, the individuals we consider are on average wealthier (with an average income of about 140,000 RAND compared to the 83,931 country-wide)[5]. Also, the manufacturer market shares suggest that smaller brands grouped together under the name "other brands" represent about 50% of handsets country-wide, while they represent only between 8 and 13% in our sample over the period studied. Even if our final sample is not representative for the population in the whole country, it allows us to study the role of

[4]We gathered information on the manufacturer and OS market shares at the country level on the website of Statcounter Global Stats.

[5]This statistics is for so called small places and unweighted.

income in smartphone adoption decisions.

Aggregation of the data: We estimate the model for two levels of data aggregation. First, we aggregate the data at the *market-level*, which yields 135-166 observations of different smartphone models per period and 614 observations in total. Second, we divide all individuals into five income groups and aggregate the data at the *income-group level*, which yields 1,519 observations in total, with the number of handset models varying between 49 and 103, depending on income group and time period. In this case we consider that people living in locations with different average income level, and hence belonging to different income group, choose from a different set of handsets. The range of smartphone models which are available to consumers varies over time in both cases.

South Africa is one of the most unequal in terms of income in the world. People with extreme differences in income live in segregated neighbourhoods with no mobility between them. The rationale for grouping consumers by income based on the location of the individuals is to capture the differences in budget, tastes and usage of mobile services. The income of individuals constrains the choice of handsets. Thus, the aggregation at the *income-group level* allows us to identify how income affects the price sensitivity and preferences for product characteristics.

Descriptive statistics: Table A.1 presents statistics on individuals (upper panel) and observations (lower panel). We observe 26,988 unique consumers, among whom about 83% live in an urban area. They switch their handsets between 0 and 5 times over the two-years period which we study (0.48 on average). They consume per month between 0 and 31 GB of data (with an average of 115MB), make between 0 and 4805 minutes of calls (with an average of 203 minutes) and send between 0 and 1431 text messages (with an average of 13 texts).

In our sample, there are 413 unique handsets which belong to 20 brands. We group 11 small brands together under the name 'small brands', so that we use only 10 brands in the subsequent analysis. Table A.3 presents the average characteristics of handsets which were observed at least once in our data over the period. On average, the cost of a handset is about

200€, with a minimum price of 14€ and a maximum price of 1110€.[6] About 69% of handsets are smartphones and 35% are compatible with LTE networks. On average, the phones selected by the consumers in the sample were released 4.2 years prior to purchase. Figures A.4 and A.5 show the differences in the distribution of prices and ages for two categories of handsets: feature phones and smartphones.

The LTE coverage and price of data may be two key determinants of the adoption of smartphones. Figure A.6 shows how they evolved on average in our sample over time. The cost of data steadily decreased over the period, while the LTE coverage witnessed a jump at the beginning of 2017.

Table A.4 shows how the characteristics of selected phones differ across income groups. We observe that individuals who belong to higher income groups choose smartphones with more desirable characteristics such as LTE compatibility, bigger memory, higher quality camera, greater number of cores of CPU, greater display size and others. Moreover, data usage increases with income, while usage of text messages decreases. This trend can be explained by the substitution between text messages and data, which depends on the income level, and hence on the handset type and purchased usage bundle or tariff plan.

Figure A.2 presents how the share of smartphone users evolves over time in the full sample. The level of smartphone adoption differs by income groups, but there is a similar upward trend with comparable slope for all groups. Figure A.3 presents the share of smartphone adopters among feature phone users at each period for the five income groups. Here, we consider the sample which is used for our estimation, i.e. the individuals who did not have a smartphone in March 2016. The rate of adoption increases sharply between 2016 and 2017 and then levels off before going up again in 2018. This figure shows that the pace of adoption is comparable across income groups, even though the adoption curves differ marginally both in terms of level and slope after 2017, where the adoption is faster for higher income groups compared to lower ones.

The sample which we use for estimation at the *market-level* consists of 614 observations, with 135 observations for September 2016, 163 for March 2017, 159 observations for September

[6]We report these numbers in Euros but use prices in local currency in the estimation.

2017, 158 for March 2018 and 166 for September 2018. The sample we use for estimation at the *income-group* level consists of 1,519 observations in total. At each period of time, we observe between 49 and 103 models of smartphones, with the greatest variety in the first income group (see Table A.5).

4 Model

We follow De Groote and Verboven (2019) and specify a dynamic adoption model which can be estimated with aggregate market-level data. As discussed above, we create five income groups in the population based on the average income information linked to geographic location of individuals in our sample. Next, we define the choice set of smartphones which is specific to each income group. An individual i_m from income group m may decide to adopt a smartphone from the set of devices available for this income group at period t, denoted with j, with $j_m = 1, ..., J_m$. We consider that a smartphone is available if we observe that it is chosen at least once by one individual from the income group.[7] Alternatively, an individual may choose to continue using a feature phone denoted as $j_m = 0$. In this notation, we ignore the fact that the choice set differs between periods. Also, to simplify the notation, we ignore the subscript for the income segment m in the derivation which follows.

Given that we have individual-level data, we know the feature phone which is used by each individual. The utility derived from using a feature phone depends on its quality and price, but also on individual's taste. Since we estimate the model using aggregate *income-group level* data, we need to approximate this utility using the same value for all consumers belonging to group m. Furthermore, we make different assumptions regarding the value of continued use of a feature phone and test how they impact our estimates. First, we normalize the utility of using a feature phone to zero for all individuals. Second, we assume that the utility of choosing a feature phone is determined by the characteristics and price of the most-used feature phone in each income group at each period of time. Third, we compute an arithmetic average based on the characteristics and price of the 50 top-used feature phones in

[7] Models which are not selected by anyone from the income group have zero market shares and cannot be be considered in our model, which de facto does not mean that they are not available.

each group at each period. In this way we create a time-varying and group-specific 'average feature phone', which is used by consumers as an alternative to adopting a smartphone. In De Groote and Verboven (2019), non-adoption is associated with zero utility, as in our first approach.

There is no individual heterogeneity in the model apart from taste shock, which is assumed to be i.i.d. and drawn from a type I extreme value distribution. This assumption implies that the preferences individuals for similar smartphones are not correlated. We only estimate group-specific price coefficients to account for differences in price responsiveness of individuals with different levels of income.

We assume that at each period t, individuals choose the alternative j which maximizes their random utility denoted by $\delta_{ijt} + \epsilon_{ijt}$, where δ_{ijt} depends on smartphone characteristics and ϵ_{ijt} is the taste shock specific to individual i. The adoption of a smartphone is a terminating state, and we ignore the fact that a small group of consumers switches back to a feature phone, as shown in Table A.2. For the sake of simplicity, we drop these observation from the data. The indirect utility derived by an individual from owning a smartphone, which consists of a stream of utilities for a certain number of time periods can be specified as:

$$(1) \qquad \delta_{ijt} = x_j\gamma - \alpha p_{jt} + \theta c_t + \xi_{jt}$$

where x_{jt} is a vector of characteristics of smartphone j, which do not change over time, p_{jt} is the smartphone price variable, c_t is 3G network coverage which varies over time and by income segment, and ξ_{jt} is the unobserved quality of alternative j in period t. We do not use fixed effects for smartphone models which cannot be estimated together with other smartphone characteristics. But we use a set of dummy variables for the main brands and operating systems.

The conditional value of not adopting a smartphone can be written as:

$$(2) \qquad \delta_{i0t} = u_{0t} + \beta E_t(\bar{V}_{t+1})$$

where u_{0t} denotes the utility in period t and $\bar{V}_{t+1}$ is the ex ante value function, i.e., the continuation value from behaving optimally from period $t+1$ onward, before the random taste shocks are revealed. This value function captures the expected utility of making a decision to switch or not in the next period. With a type I extreme value distribution for the random taste shocks ϵ_{ijt}, the ex ante value function $\bar{V}_{t+1}$ has the well-known closed-form logsum expression:

$$(3) \quad \bar{V}_{t+1} = 0.577 + \ln \sum_{j=0}^{J} exp(\delta_{jt+1})$$

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

The random utility maximization yields the following choice probabilities or predicted market shares for each alternative $j = 0, ..., J$ in period t:

$$(4) \quad S_{jt} = s_{jt}(\delta_t) \equiv \frac{exp(\delta_{jt})}{\sum_{k=0}^{J} exp(\delta_{kt})}$$

where following Berry (1994) the predicted market share $s_{jt}(\delta_t)$ corresponds to the observed market shares S_{jt} because of the inclusion of unobserved quality ξ_{jt} for every product and period. The observed market share of smartphone j in period t is defined as $S_{jt} = q_{jt}/N_t$, where q_{jt} is the observed number of users of this smartphone and N_t is the total number of individuals in our sample in period t. Since the adoption of a smartphone is terminal action, the potential number of adopters in period t is the total number of individuals in our sample, N, less the number of individuals that adopted a smartphone in the past, i.e., $N_t = N - \sum_{\tau=1}^{t-1} \sum_{j=1}^{J} q_{j\tau}$. The aggregate market share of not adopting is $S_{0t} = 1 - \sum_{j=1}^{J} S_{jt}$.

De Groote and Verboven (2019) show how to obtain an analytic expression for the expected future value term $E_t(\bar{V}_{t+1})$, which enters the conditional value for not adopting given by (2) and is recursively defined by (3). The usual approach to compute $E_t(\bar{V}_{t+1})$ is by specifying an explicit stochastic process of the state transitions. Instead, they follow Scott (2013) and decompose $E_t(\bar{V}_{t+1})$ into the realized ex ante value function $\bar{V}_{t+1}$ and a short run prediction error $\eta_t \equiv \bar{V}_{t+1} - E_t(\bar{V}_{t+1})$. Assuming that individuals' expectations are on average correct, so that η_t is mean zero, equation (2) can be written as:

$$(5) \quad \delta_{0t} = u_{0t} + \beta(\bar{V}_{t+1} - \eta_t)$$

The next step follows Hotz and Miller (1993) who show how to write $\bar{V}_{t+1}$ in terms of the conditional choice probabilities. When the decision problem has a terminal action as in our setup, we can take the next period CCP for an arbitrary terminating choice, e.g. $j = 1$, as given by $s_{1t+1}(\delta_{t+1}) \equiv exp(\delta_{1t+1})/\sum_{j=0}^{J} exp(\delta_{jt+1})$. This expression after taking logs becomes:

$$(6) \quad \ln \sum_{j=0}^{J} exp(\delta_{jt+1}) = \delta_{1t+1} - \ln s_{1t+1}(\delta_{t+1})$$

After substituting (6) into (3), we obtain the following expression for the ex ante value function at $t + 1$:

$$(7) \quad \bar{V}_{t+1} = 0.577 + \delta_{1t+1} - \ln s_{1t+1}(\delta_{t+1})$$

When substituting (7) into the mean utility from not adopting (5) we get:

$$(8) \quad \delta_{0t} = u_{0t} + \beta(0.577 + \delta_{1t+1} - \ln s_{1t+1}(\delta_{t+1}) - \eta_t)$$

After normalizing $u_{0t} + \beta 0.577 = 0$ and given that the CCP at the realized mean utilities is equal to the observed market share ($S_{1t+1} = s_{1t+1}(\delta_{t+1})$), we get:

$$(9) \quad \delta_{0t} = \beta(\delta_{1t+1} - \ln S_{1t+1} - \eta_t)$$

Next, the market share equation can be inverted following Berry (1994). The choice probabilities S_{jt} for each $j = 1, ..., J$ given by the market share expressions (4) are divided by S_{0t} and after taking logs become:

$$(10) \quad \ln(S_{jt}/S_{0t}) = \delta_{jt} - \delta_{0t}$$

We get the following main estimating equation after substituting the expressions for the mean utilities (9) and (1) into (10):

$$(11) \quad ln(S_{jt}/S_{0t}) = (x_{jt} - \beta x_{1t+1})\gamma + \theta(c_t - \beta c_{t+1}) - \alpha(p_{jt} - \beta p_{1t+1}) + \beta ln(S_{1t+1}) + \epsilon_{jt}$$

where the econometric error term is defined as:

$$(12) \quad \epsilon_{jt} \equiv \xi_{jt} - \beta(\xi_{1t+1} - \eta_t)$$

When the discount factor $\beta = 0$, the model and error term simplify to standard static logit model as in Berry (1994). We follow De Groote and Verboven (2019) and assume that $\beta = 0.99$, for which the equation (11) is a regression of the change in the number of new adopters on the change in price and possibly other characteristics. In the model above, the outside option is normalized to zero. We also estimate alternative model specifications, where the utility of outside option, $u_{0t} + \beta 0.577 \neq 0$, is determined by the most-used feature phone in each income group and period:

$$ln\left(\frac{S_{jt}/S_{1t+1}^{\beta}}{S_{0t}}\right) = (x_{jt} - \beta x_{1t+1} - x_{Ft})\gamma + \theta(c_t - \beta c_{t+1}) - \alpha(p_{jt} - \beta p_{1t+1} - p_{Ft}) - \beta 0.577 + \epsilon_{jt}$$

where x_{Ft} and p_{Ft} denote a vector of characteristics and price of top-used feature phone in period t. In another specification instead of top-used feature phone we use the average characteristics and price of 50 top-used feature phones. The arbitrary terminating product choice, $j = 1$, differs between income groups and over time. For example, for the income group below poverty line we use Vodafone Smart Kicka as the arbitrary product in the first three periods, Vodafone Kicka 3 in March 2018 and Vodafone Kicka 4 in September 2018. These correspond to the top-used models in the lowest income group.

A vector of characteristics of smartphone j at time t is denoted by x_{jt} and includes: (i) brand; (ii) operating system; (iii) age of the model, (iv) a dummy for not having CPU, (v) dimensions in terms of height, width and thickness, (vi) weight.[8] Furthermore, p_{1t+1} denotes

[8]A smartphone CPU core is an individual processing unit found in the central processing unit (CPU) of

the price and x_{1t+1} denotes a vector of characteristics of the arbitrary terminating choice, $j = 1$, at time $t + 1$.

5 Results

We estimate the model using the Ordinary Least Squares (OLS) and the Generalized Method of Moments (GMM), where we instrument the handset price with two handset characteristics: a dummy variable for Graphics Processing Unit (GPU) and Random Access Memory (RAM).[9] These characteristics are responsible for the performance of smartphone and determine the cost of production. We do not include them as determinants of the utility function because consumers may be unaware of these technical features. These characteristics are therefore correlated with the price of smartphone as cost drivers, but should not be correlated with the error term which corresponds to the unobserved quality in our estimated equation.

Each observation represents a smartphone purchased in a given income group and period. In case the smartphone model was purchased in more than one income group and period, there are multiple observations on this model with different market shares. In our base specification, the outside option which is the utility of feature phone, is normalized to zero in each period. In two alternative specifications, which we report in the Appendix, the outside option is determined by the characteristics of: (i) top-used feature phone in each group at each period; (ii) the average characteristics of 50 top-used feature phones in each group at each period.

First, we estimate a static demand model, where we eliminate the terms with the discount factor in equation (11) by setting $\beta = 0$. The estimation results are shown in Table A.6. We pool together the data for five income groups (column All), and then estimate the model separately for two groups with the lowest income (column Poor) and three groups with the highest income (column Well-off). The estimation results for OLS are shown in columns (1)-(3) and for GMM in columns (4)-(6). Next, we estimate a dynamic demand model where the

a mobile phone. It is responsible for receiving and executing instructions that are sent from the user to the phone.

[9]The main purpose of GPU is to perform graphics processing operations or do floating point calculations. In simple terms, it is a specialized circuit whose main job is to generate images for the device to display. Every smartphone has a GPU in some form to generate pictures.

discount factor is set to $\beta = 0.99$ in equation (11). The estimation results for OLS and GMM are shown in Table (A.7).

The estimation results of both static and dynamic models are comparable in terms of signs and significance of variables. The price coefficient is significant and negative in both specifications. It is greater in absolute terms for poor consumers, as compared to the group of well-off consumers. The price coefficient increases when GMM is estimated for all three regressions. This suggests that in the OLS regression the price coefficient is biased downwards due to endogeneity. The coverage with LTE networks increases the adoption of smartphones with comparable impact for both poor and well-off consumers in the static framework. In the dynamic framework, there seem to be bigger impact on well-off consumers. Thus, investment in LTE coverage stimulates adoption of smartphones.

The age of smartphones reflects the quality and, as expected, newer smartphones are more valued by consumers. We include four dummy variables for 1, 2, 3 and 4 years passed since the release date. The estimates of these dummy variables are interpreted relative to smartphones released 5 or more years before. Smartphones without CPU are less valued relative to those with CPU. The dimensions of smartphones impact the utility with preference for smaller ones, where those with height of 120mm and less are valued more, width of 63mm and less are valued more and weight of 120 gram and less are valued more. The smartphones produced by Apple and Blackberry are valued more than other brands. These two manufacturers rely on own operating systems, as compared to other brands which use Android, Windows and other OS. There is no difference in the valuation of these operating systems because dummy variables for Android and Windows are insignificant. We also include in the estimation a set of dummy variables for top-used brands and models, which are not shown in the table due to space constraints.

5.1 Counterfactual scenarios

We use the model for the following counterfactual simulations. First, we consider that to stimulate adoption among poorest consumers the government eliminates VAT on smartphones, which is now 15%. This corresponds to price reduction of smartphones by 1/1.15=13%, while the price of feature phones remains unchanged, and thus the utility of outside option. Second, we consider that there was full LTE coverage from the first period in our data. The investment in deployment of LTE networks is a relatively slow process starting from urban and densely populated areas. Thus, individuals living in rural areas who are generally poorer are disadvantaged with respect to access to mobile Internet, which contributes to digital divide between geographic regions.

In the static framework, we first compute the share of smartphone adopters in each period in the sample used in the estimation, which includes only individuals without a smartphone in the first period. Next, we use this information to calculate smartphone adoption path for different income groups using full data set, which in addition includes people who used feature phones in the first period. The simulated adoption paths are show in Table (A.8) for price reduction due to eliminating VAT on smartphones, and in Table (A.9) for full LTE coverage.

We follow the same procedure in the dynamic framework, but the formula for computing the share of smartphone adopters differs, as shown in the Appendix. We then also calculate smartphone adoption path for different income groups using full data set. The adoption path for VAT-related price reduction is shown in Table (A.10). The results are comparable with the static model, which suggests that consumers do not take future price and quality into account when purchasing smartphones. We are not able to estimate the counterfactual change in smartphone adoption due to the assumption of full LTE coverage over the entire period. This is because in equation (11) LTE coverage is a difference between coverage in periods t and $t+1$.

We find that there is a 'digital divide' in access to mobile Internet caused by the fact that poor people live in areas without network coverage. The cost of smartphones does not have significant impact on adoption decisions. In particular in the static framework, if both poor and richer areas were fully covered by LTE networks in the entire period of our data, the adoption of smartphones would increase from 57.1% to 76.0% among people below poverty line, while it would only change from 81.5% to 82.2% in the richest group of consumers as of 1st quarter of 2018. This is precisely because richer areas have almost full LTE coverage, while poorer regions are only partially covered. At the same time, removing 15% VAT on smartphones would increase adoption only to 57.7% among people below poverty line, and to 81.7% among richest consumers as of 1st quarter of 2018. The price effect has only marginal impact on adoption in all income groups. We conclude that to reduce smartphones-related 'digital divide' it is critical to develop LTE infrastructure in poorer areas. People will respond by adopting smartphones irrespective of their income.

6 Conclusions

In this paper, we construct a unique database of adopters of smartphones with different levels of income in South Africa, which is a developing economy with large income inequality. We estimate a static and dynamic model of smartphone adoption for different income groups. We use the model to assess policies which can stimulate adoption of smartphones among people living below poverty line. We find that the main driver of adoption is coverage by LTE networks. The price of smartphones has only marginal impact on smartphone adoption decision. We conclude that to reduce smartphones-related 'digital divide' it is critical to develop LTE infrastructure in poorer areas and people will respond by adopting smartphones irrespective of their income. The static and dynamic models yield comparable results suggesting that consumers do not take future price and quality into account when purchasing smartphones.

In our further research, we will estimate a dynamic adoption model using individual-level data accounting for smartphone prices, differences in network coverage and the cost of mobile data.

References

Aker, J. C., and I.M. Mbiti (2010). Mobile phones and economic development in Africa. Journal of economic Perspectives, 24(3), pp.207-232.

Arcidiacono, P., and P.B. Ellickson (2011). Practical methods for estimation of dynamic discrete choice models. Annu. Rev. Econ., 3(1), pp.363-394.

Aguirregabiria, V. and P. Mira. 2002. "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models". Econometrica, 70 (July): 1519-1543.

Aguirregabiria, V. and P. Mira. 2010. "Dynamic discrete choice structural models: A survey", Journal of Econometrics, 156(1): 38-67.

Arcidiacono, P., and R.A. Miller. 2011. Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity. Econometrica 79 (6): 1823–67.

Berry, S.T. 1994. Estimating discrete-choice models of product differentiation. RAND Journal of Economics, 25(2): 242-262.

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

Bjorkegren, D. (2019). The adoption of network goods: Evidence from the spread of mobile phones in Rwanda. The Review of Economic Studies, 86(3), pp.1033-1060.

Conlon, Ch.T. 2012. A Dynamic Model of Prices and Margins in the LCD TV Industry. Unpublished.

De Groote, O., and F. Verboven, (2019). Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems. American Economic Review, 109(6), pp.2137-2172.

Fan, Y., and C. Yang, (2020). Competition, Product Proliferation, and Welfare: A Study of the US Smartphone Market. American Economic Journal: Microeconomics, 12(2), pp.99-134.

Gowrisankaran, G. and M. Rysman, (2012). Dynamics of consumer demand for new durable goods. Journal of political Economy, 120(6), pp.1173-1219.

Grigolon, L., Reynaert, M., and F. Verboven, (2018). Consumer valuation of fuel costs and tax policy: Evidence from the European car market. American Economic Journal: Economic Policy, 10(3), pp.193-225.

Grzybowski,. L. and A. Nicolle. 2021. Estimating consumer inertia in repeated choices of smartphones. Journal of Industrial Economics, 63(1): 33-82.

Grzybowski,. L. and J. Liang. 2015. Estimating demand for fixed-mobile bundles and switching costs between tariffs. Information Economics and Policy, 33: 1–10.

Hiller, R.S., Savage, S.J. and Waldman, D.M. 2018. Using Aggregate Market Data to Estimate Patent Value: An Application to United States Smartphones 2010 to 2015. International Journal of Industrial Organization, 60: 1-31.

Hotz, V. J., and R.A. Miller, (1993). Conditional choice probabilities and the estimation of dynamic models. The Review of Economic Studies, 60(3), pp.497-529.

Luo, R., 2018. Network Effect and Multi-Network Sellers' Dynamic Pricing: Evidence from the US Smartphone Market. Unpublished.

Miller, R., 1984. Job matching and occupational choice. Journal of Political Economy 92, 1086-1120.

Nair, H. 2007. Intertemporal Price Discrimination with Forward-looking Consumers: Application to the US Market for Console Video-Games. Quantitative Marketing and Economics, 5(3): 239–292.

Pakes, A., 1986. Patents as options: Some estimates of the value of holding European patent stocks. Econometrica 54, 755-784.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. Econometrica, pp.999-1033.

Schiraldi, P. 2011. Automobile Replacement: A Dynamic Structural Approach. RAND Journal of Economics, 42(2): 266-291.

Scott, P.T. 2013. "Dynamic Discrete Choice Estimation of Agricultural Land Use." Unpublished.

Shreeti, V. (2019). Evolution of the Mobile Handset Market in India. Unpublished.

Sinkinson, M., 2014. Pricing and Entry Incentives with Exclusive Contracts: Evidence from Smartphones. Unpublished Sun, Y., 2012. The Value of Branding in Two-Sided Platforms. Unpublished

Wolpin, K., 1984. An estimable dynamic stochastic model of fertility and child mortality. Journal of Political Economy 92, 852-874.

Appendix A: Main descriptive and Results

A.1 Descriptive statistics

Table A.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
Individuals				
Living in urban area (0/1)	0.83	0.38	0	1
Black	0.83	0.28	0	1
Colored	0.07	0.18	0	0.98
White	0.1	0.2	0	0.99
Number of handset switches	0.48	0.71	0	5
Average usage of sms (units)	13.27	60.47	0	1431
Average usage of voice (minutes)	202.73	293.82	0	4805.27
Average usage of data (mb)	115.18	603.54	0	31283.11
N	26,988			
Observations				
Text usage in units	13.76	70.95	0	2648
Voice usage in minutes	220.02	323.99	0	5250.7
Data usage in MB	116.17	663.1	0	44243.42
Smartphone 0/1	0.09	0.29	0	1
Handset = Samsung	0.22	0.42	0	1
Handset = Nokia	0.5	0.5	0	1
Handset = Apple	0.0011	0.03	0	1
Living in urban area (0/1)	0.83	0.38	0	1
Average household income	94308.69	103648.39	0	1311958
2G coverage	0.99	0.03	0.16	1.01
3G coverage	0.97	0.08	0	1
4G coverage	0.63	0.44	0	1
N	65,556			

Figure A.1: Coverage by geographic area

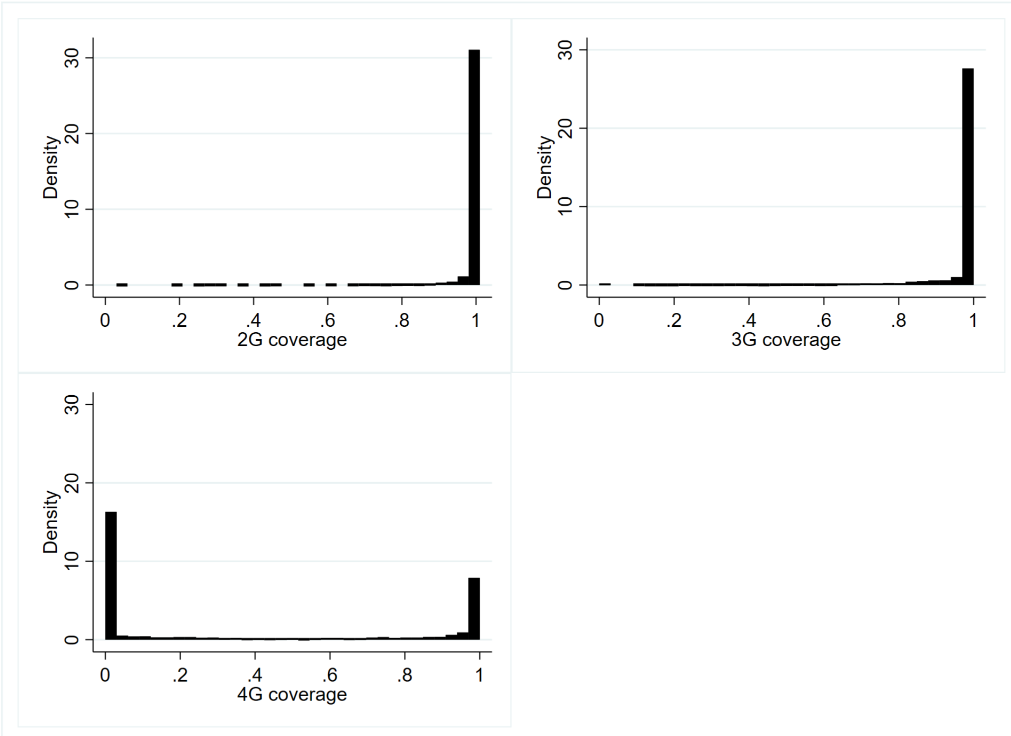
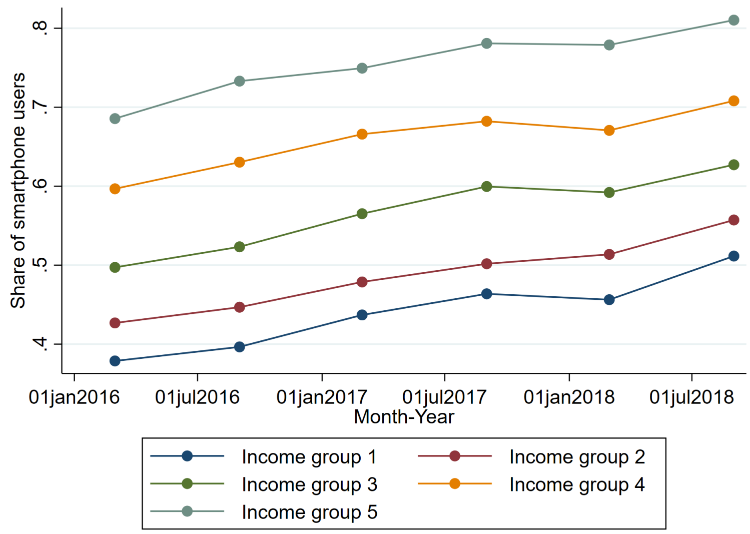


Table A.2: Switching patterns in the full dataset

	Freq.	Percent
Stays with smartphone	23,453	44.16
Stays with feature phone	20,917	39.38
Upgrade to smartphone	4,349	8.19
Downgrade to feature phone	2,669	5.03
Other switching patterns	1,722	3.24
Total	53,110	100

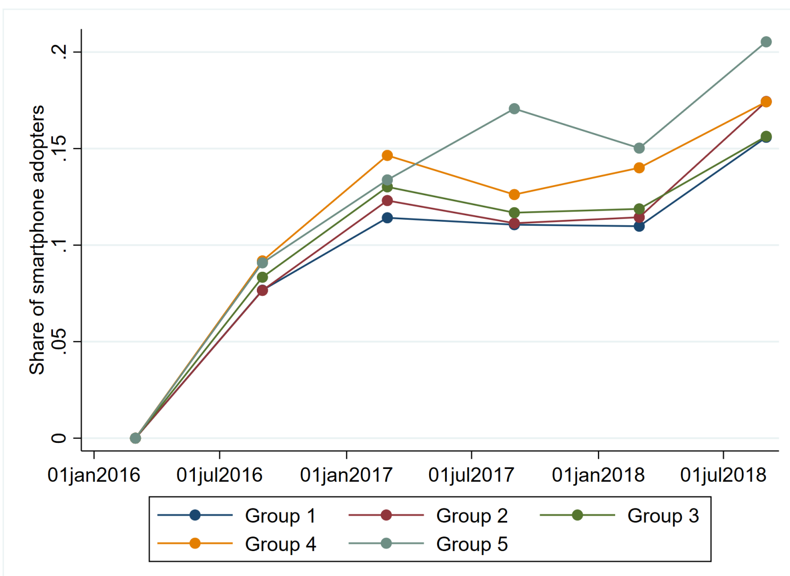
Observations correspond to unique individuals in the full sample.

Figure A.2: Share of smartphones users, per income groups



Computed on the full sample (145,977 observations, 53,110 individuals)

Figure A.3: Share of smartphone adopters, per income groups



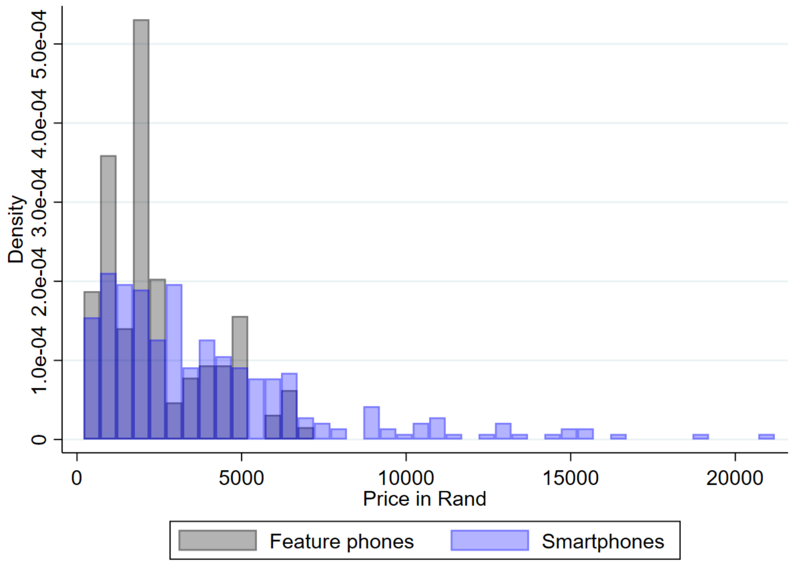
Computed on the estimation sample (65,556 observations, 26,988 individuals)

Table A.3: Characteristics of handsets

Variable	Mean	Std	Min	Max	N
Price in thousands RAND	3.52	3.12	0.24	20.75	413
Average age over the period in years	4.19	3.25	0	18.18	413
Smartphone	0.69	0.46	0	1	413
LTE compatible	0.35	0.48	0	1	413
Height in mm	126.79	19.33	80	188.07	413
Width in mm	63.89	14.66	39	255.5	413
Thickness in mm	11	3.3	5.9	22.5	413
Weight in g	127.15	32.33	59	220	413
Internal memory in GB	11.8	19.89	0	128	413
RAM in GB	0.95	1.2	0	6	413
Camera quality in mpx	6.04	4.79	0	22.57	413
Second camera quality in mpx	2.11	3.37	0	23.79	413
Number of CPUs	0.71	0.64	0	3	413
Number of GPUs	0.26	0.44	0	1	413
Display resolution (Mpix)	0.69	0.88	0.01	4.26	410
Battery power in mAh	1855.78	835.78	400	4100	413
Apple	0.04	0.2	0	1	413
Samsung	0.31	0.46	0	1	413
Android OS	0.51	0.5	0	1	413
Blackberry OS	0.06	0.23	0	1	413
Windows OS	0.05	0.22	0	1	413
Other OS: Symbian, Bada, Tizen	0.03	0.16	0	1	413

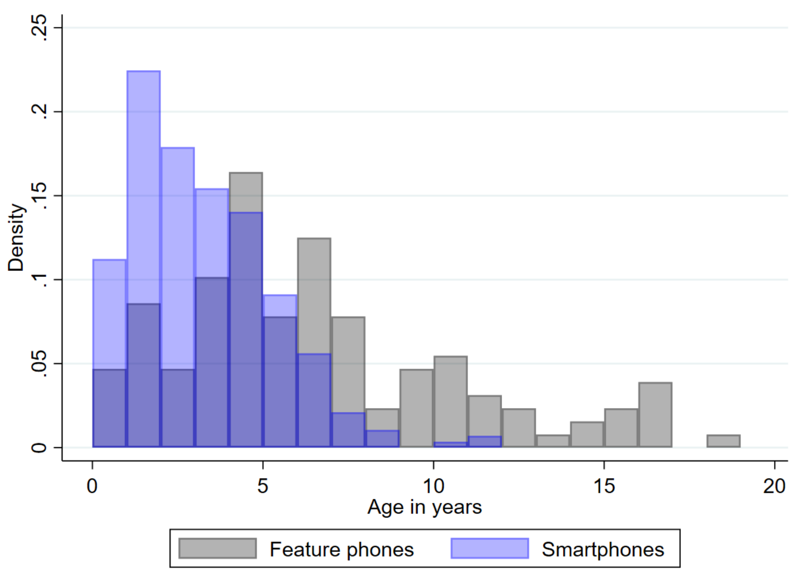
Price and age of phones are computed as average over the period.

Figure A.4: Histogram of handset prices



Average price in RAND over the whole time period for 413 unique models

Figure A.5: Histogram of models age



Average age of models computed over the whole time period for 413 unique models

Figure A.6: Evolution of coverage and data price

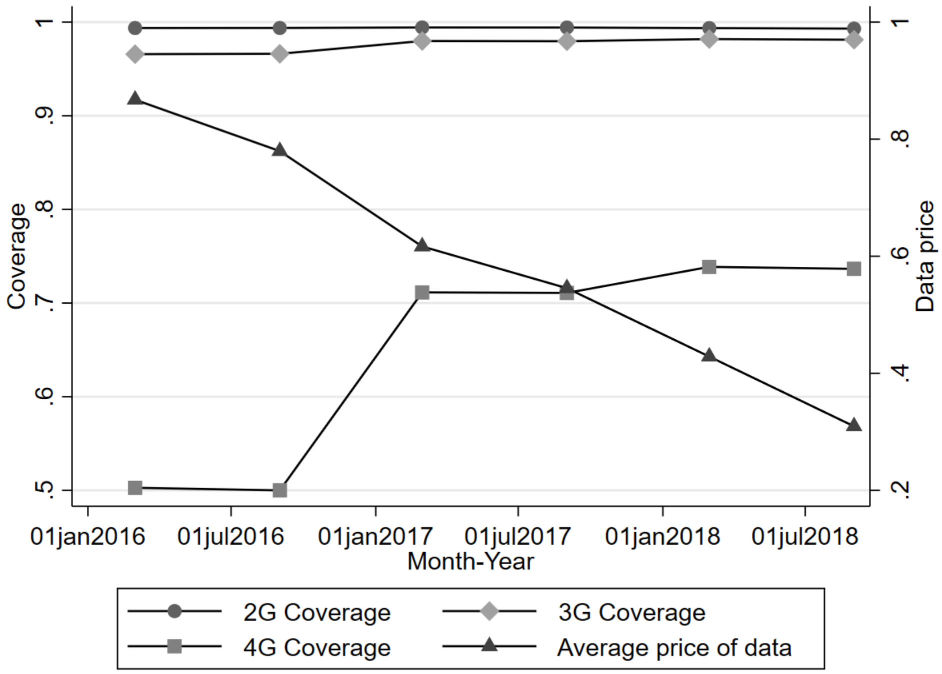
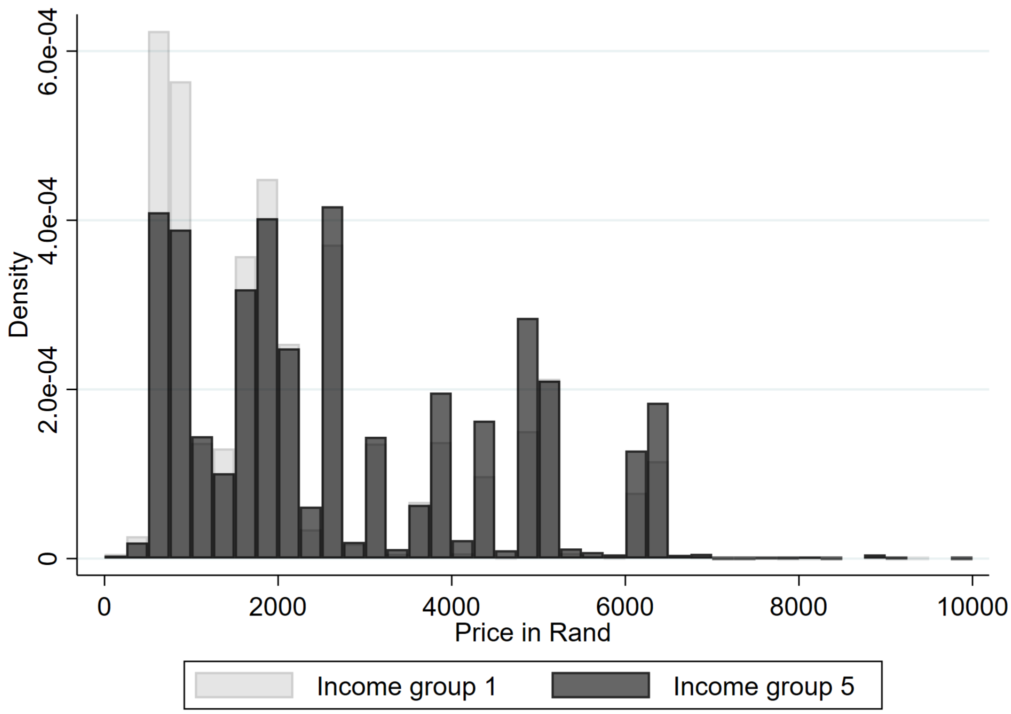


Table A.4: Characteristics of choices, per income group

	Group 1	Group 2	Group 3	Group 4	Group 5
Price in thousands RAND	2.29	2.35	2.38	2.55	2.83
Data usage in MB	72.77	96.75	113.03	196.58	276.14
Voice usage in minutes	230.50	224.46	211.98	199.72	199.09
Text usage in units	14.94	11.33	12.66	12.92	16.02
Age of the model in years	5.02	5.10	5.29	5.83	6.19
Smartphone	0.09	0.09	0.09	0.10	0.11
LTE compatible	0.03	0.04	0.04	0.05	0.06
Height	113.03	112.94	112.39	111.42	110.32
Width	53.70	53.88	53.63	53.40	52.48
Thickness	13.25	13.27	13.40	13.61	13.79
Weight	94.14	94.55	94.46	95.12	93.66
Internal memory	0.84	0.99	0.95	1.37	1.82
RAM	0.15	0.16	0.16	0.19	0.20
Camera quality	0.12	0.17	0.16	0.26	0.34
Second camera quality	0.15	0.16	0.15	0.17	0.17
Number of CPUs	0.02	0.02	0.02	0.03	0.04
Number of GPUs	0.15	0.15	0.15	0.17	0.19
Display resolution	0.15	0.15	0.15	0.17	0.19
Battery power	1216	1220	1216	1213	1188
Apple	0.000	0.001	0.001	0.002	0.005
Samsung	0.19	0.21	0.23	0.27	0.32
AndroidOS	0.07	0.07	0.07	0.09	0.10
BlackberryOS	0.01	0.01	0.01	0.01	0.01
WindowsOS	0.00	0.00	0.00	0.00	0.00
Other OS	0.00	0.01	0.01	0.01	0.00

Computed on 65,556 observations, 26,988 unique individuals

Figure A.7: Histograms of prices for the two extreme income groups



Cut at 10,000 RAND

Table A.5: Number of unique models in the choice set of various income groups

Group	mean	sd	p50	min	max
1	98.25	5.12	99.5	91	103
2	74.4	6.85	76.5	65	80
3	70.75	4.19	69	68	77
4	71.75	6.07	71.5	65	79
5	64.75	11.79	66.5	49	77
Total	76	13.53	74	49	103

A.2 Estimation results

Table A.6: Static demand model

	OLS			GMM		
VARIABLES	All	Poor	Well-off	All	Poor	Well-off
price	-0.027***	-0.038***	-0.021***	-0.078***	-0.130***	-0.066***
	(0.008)	(0.015)	(0.008)	(0.024)	(0.047)	(0.025)
Coverage LTE	2.254***	2.173***	1.998***	2.294***	2.322***	2.060***
	(0.105)	(0.261)	(0.187)	(0.105)	(0.259)	(0.186)
Handset age: 1 year	0.715***	0.934***	0.488***	0.712***	0.933***	0.485***
	(0.077)	(0.132)	(0.088)	(0.077)	(0.134)	(0.088)
Handset age: 2 year	0.708***	0.911***	0.520***	0.713***	0.910***	0.532***
	(0.069)	(0.113)	(0.078)	(0.069)	(0.111)	(0.078)
Handset age: 3 year	0.510***	0.653***	0.379***	0.488***	0.612***	0.359***
	(0.065)	(0.104)	(0.076)	(0.065)	(0.103)	(0.076)
Handset age: 4 year	0.289***	0.379***	0.183***	0.281***	0.367***	0.174***
	(0.058)	(0.090)	(0.067)	(0.058)	(0.088)	(0.067)
CPU = 0	-0.291***	-0.367***	-0.218***	-0.298***	-0.388***	-0.221***
	(0.066)	(0.102)	(0.082)	(0.065)	(0.101)	(0.081)
Height <120mm	0.262***	0.504***	0.053	0.267***	0.534***	0.054
	(0.060)	(0.093)	(0.075)	(0.059)	(0.091)	(0.074)
Width <63mm	0.144**	0.161*	0.108	0.136**	0.157*	0.093
	(0.059)	(0.092)	(0.074)	(0.058)	(0.091)	(0.074)
Weight <120	0.261***	0.319***	0.229***	0.229***	0.254***	0.197***
	(0.055)	(0.086)	(0.065)	(0.057)	(0.091)	(0.067)
Thickness <8mm	0.018	-0.068	0.066	0.187*	0.244	0.222**
	(0.065)	(0.111)	(0.073)	(0.099)	(0.186)	(0.104)
Apple	0.396**	0.468	0.264	0.522***	0.642*	0.369
	(0.179)	(0.291)	(0.216)	(0.195)	(0.349)	(0.232)
Blackberry	0.837***	0.953***	0.678***	0.806***	0.917***	0.643***
	(0.161)	(0.251)	(0.200)	(0.163)	(0.258)	(0.203)
Os Android	0.215	0.290	0.098	0.181	0.267	0.040
	(0.157)	(0.246)	(0.194)	(0.161)	(0.264)	(0.199)
OS Windows	0.378***	0.455**	0.234	0.292*	0.300	0.145
	(0.145)	(0.206)	(0.193)	(0.151)	(0.224)	(0.200)
Constant	-9.707***	-10.006***	-9.112***	-9.576***	-9.837***	-9.000***
	(0.177)	(0.296)	(0.247)	(0.189)	(0.318)	(0.256)
Product dummies	yes	yes	yes	yes	yes	yes
Observations	1,519	691	828	1,519	691	828
R-squared	0.487	0.487	0.403	0.486	0.481	0.400

Table A.7: Dynamic demand model

	OLS			GMM		
VARIABLES	All	Poor	Well-off	All	Poor	Well-off
price	-0.032***	-0.045***	-0.025**	-0.124***	-0.213***	-0.091***
	(0.010)	(0.016)	(0.010)	(0.032)	(0.056)	(0.031)
Coverage LTE	2.098***	1.058**	2.103*	2.176***	1.156**	2.307**
	(0.388)	(0.457)	(1.074)	(0.386)	(0.462)	(1.054)
Handset age: 1 year	1.125***	1.148***	0.879***	1.116***	1.129***	0.870***
	(0.088)	(0.144)	(0.093)	(0.090)	(0.151)	(0.095)
Handset age: 2 year	0.755***	0.818***	0.623***	0.781***	0.850***	0.642***
	(0.082)	(0.126)	(0.087)	(0.085)	(0.132)	(0.090)
Handset age: 3 year	0.482***	0.521***	0.386***	0.456***	0.456***	0.363***
	(0.078)	(0.119)	(0.084)	(0.079)	(0.122)	(0.084)
Handset age: 4 year	0.299***	0.290***	0.214***	0.300***	0.289***	0.205***
	(0.069)	(0.104)	(0.077)	(0.070)	(0.107)	(0.077)
CPU = 0	-0.155*	-0.213*	-0.034	-0.206**	-0.278**	-0.061
	(0.084)	(0.117)	(0.092)	(0.084)	(0.115)	(0.093)
Height <120mm	0.640***	0.902***	0.352***	0.635***	0.946***	0.351***
	(0.067)	(0.099)	(0.074)	(0.068)	(0.101)	(0.076)
Width <63mm	-0.467***	-0.612***	-0.317***	-0.435***	-0.572***	-0.298***
	(0.053)	(0.074)	(0.060)	(0.054)	(0.077)	(0.062)
Weight <120	0.446***	0.479***	0.384***	0.386***	0.340***	0.323***
	(0.066)	(0.098)	(0.074)	(0.069)	(0.106)	(0.077)
Thickness <8mm	-0.001	-0.097	0.014	0.323**	0.479**	0.262**
	(0.080)	(0.114)	(0.086)	(0.131)	(0.207)	(0.132)
Apple	0.496**	0.232	0.492*	0.707***	0.535	0.602**
	(0.222)	(0.282)	(0.252)	(0.254)	(0.413)	(0.278)
Blackberry	0.645***	0.560**	0.632***	0.619***	0.495*	0.582**
	(0.192)	(0.253)	(0.230)	(0.203)	(0.287)	(0.238)
Os Android	0.024	-0.027	0.034	0.004	-0.062	-0.033
	(0.195)	(0.260)	(0.229)	(0.211)	(0.314)	(0.239)
OS Windows	0.093	0.029	0.116	-0.049	-0.294	-0.010
	(0.168)	(0.202)	(0.223)	(0.183)	(0.244)	(0.235)
Constant	-1.113***	-1.338***	-1.075***	-1.064***	-1.285***	-1.019***
	(0.082)	(0.170)	(0.099)	(0.079)	(0.161)	(0.098)
Product dummies	yes	yes	yes	yes	yes	yes
Observations	1,519	691	828	1,519	691	828
R-squared	0.461	0.510	0.401	0.453	0.476	0.392

Table A.8: Static model: simulation of removing 15% VAT

	Penetration					Simulations				
	inc 1	inc 2	inc 3	inc 4	inc 5	inc 1	inc 2	inc 3	inc 4	inc 5
2016q3	41.6%	46.0%	53.0%	63.1%	71.4%	41.7%	46.1%	53.1%	63.2%	71.5%
2017q1	47.7%	51.7%	57.6%	68.0%	75.3%	48.0%	51.9%	57.8%	68.2%	75.4%
2017q3	52.4%	56.2%	61.5%	71.7%	78.8%	52.8%	56.6%	61.7%	71.9%	79.0%
2018q1	57.1%	60.2%	65.1%	75.3%	81.5%	57.7%	60.7%	65.4%	75.5%	81.7%

Table A.9: Static model: simulation full LTE coverage

	Penetration					Simulations				
	inc 1	inc 2	inc 3	inc 4	inc 5	inc 1	inc 2	inc 3	inc 4	inc 5
2016q3	41.6%	46.0%	53.0%	63.1%	71.4%	51.7%	53.3%	56.8%	65.0%	72.0%
2017q1	47.7%	51.7%	57.6%	68.0%	75.3%	62.6%	62.5%	63.5%	70.5%	75.9%
2017q3	52.4%	56.2%	61.5%	71.7%	78.8%	69.9%	69.3%	68.6%	74.5%	79.5%
2018q1	57.1%	60.2%	65.1%	75.3%	81.5%	76.0%	74.6%	73.0%	78.3%	82.2%

Table A.10: Dynamic model: simulation of removing 15% VAT

	Penetration					Simulations				
	inc 1	inc 2	inc 3	inc 4	inc 5	inc 1	inc 2	inc 3	inc 4	inc 5
2016q3	41.7%	46.5%	52.8%	63.2%	70.6%	41.8%	46.6%	52.8%	63.3%	70.7%
2017q1	45.6%	52.1%	58.8%	67.2%	74.1%	45.9%	52.3%	59.0%	67.3%	74.1%
2017q3	48.8%	57.0%	63.8%	70.5%	77.4%	49.3%	57.4%	63.9%	70.7%	77.5%
2018q1	56.2%	61.9%	68.8%	74.6%	79.9%	56.9%	62.4%	69.0%	74.9%	80.0%

Appendix B: Model predictions

The equation which we estimate in dynamic framework for product j is given by:

$$ln\Big(\frac{S_{jt}/S_{1t+1}^{\beta}}{S_{0t}}\Big) = (x_{jt}-\beta x_{1t+1}-x_{Ft})\gamma+\theta(c_t-\beta c_{1t+1})-\alpha(p_{jt}-\beta p_{1t+1}-p_{Ft})-\beta 0.577+\epsilon_{jt}$$

which can be written as:

$$\ln\Big(S_{jt}/S_{0t}\Big) = \beta\ln(S_{1t+1}) + \Delta_{jt}$$

where Δ_{jt} depends on the characteristics of j at time t as well as on the characteristics of arbitrary product denoted by 1 at time $t+1$. We compute Δ_{jt} using data and estimated parameters, where ϵ_{jt} is assumed to be zero-mean. We can further write this equation as:

$$S_{jt}/S_{0t} = \exp(\beta\ln(S_{1t+1}) + \Delta_{jt})$$

and using $S_{0t} = 1 - \sum_{i=1}^{N} S_{it}$ as:

$$\frac{S_{jt}}{1-\sum_{i=1}^{N} S_{it}} = \exp(\beta\ln(S_{1t+1}) + \Delta_{jt})$$

Summing up for all products and solving for the share of all smartphones we get:

$$\sum_{i=1}^{N} S_{it} = \frac{\sum_{i=1}^{N}(\beta\ln(S_{1t+1}) + \Delta_{jt})}{1+\sum_{i=1}^{N}(\beta\ln(S_{1t+1}) + \Delta_{jt})}$$

In the case of static model, the starting equation is:

$$ln(S_{jt}/S_{0t}) = x_{jt}\gamma + \theta c_t - \alpha p_{jt} + \epsilon_{jt} = \delta_{jt} + \epsilon_{jt}$$

where δ_{jt} is computed using data and estimated parameters and assuming that ϵ_{jt} is zero-mean. After taking exponential and summing up for all products we can write this equation as:

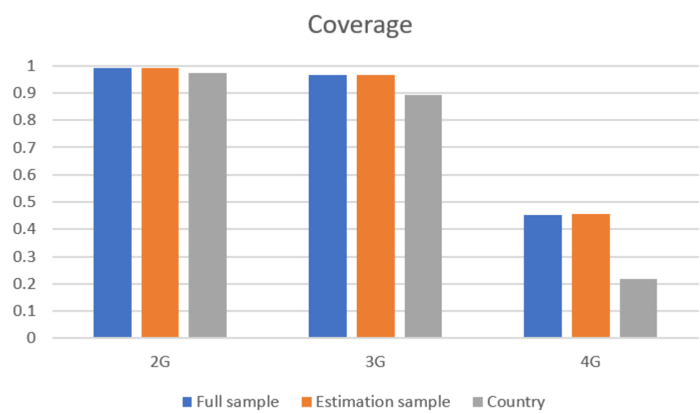
$$\frac{\sum_{j=1}^{N} S_{jt}}{1-\sum_{j=1}^{N} S_{jt}} = \exp(\delta_{jt})$$

which can be solved for the share of all smartphones:

$$\sum_{i=1}^{N} S_{it} = \frac{\sum_{i=1}^{N} \delta_{jt}}{1 + \sum_{i=1}^{N} \delta_{jt}}$$

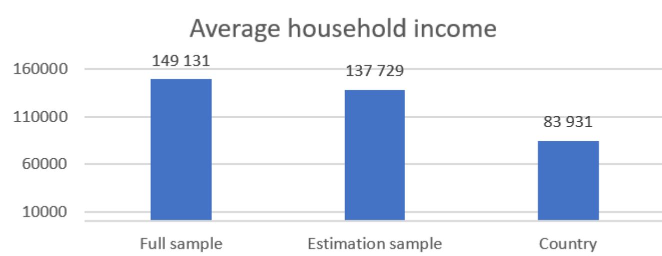
Appendix C: Country vs. sample

Figure C.1: Comparison of coverage



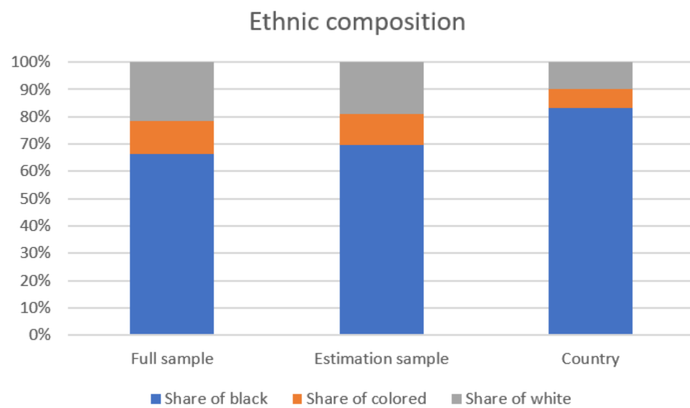
Statistics for main places. Country-level data is based on the coverage data.

Figure C.2: Comparison of average income



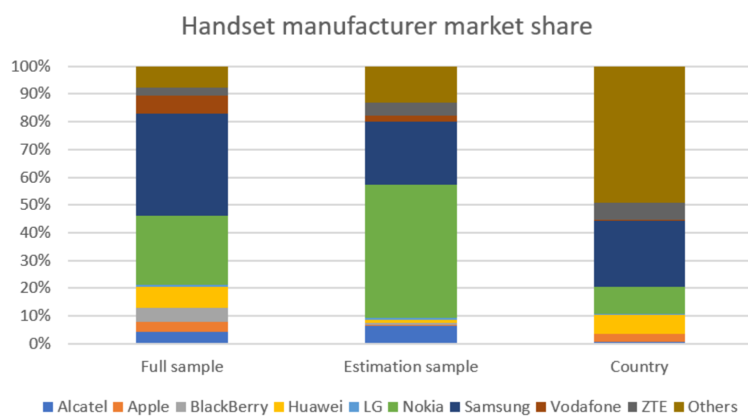
Statistics for small places. Country-level data is based on the census data.

Figure C.3: Comparison of demographics



Statistics for small places. Country-level data is based on the census data.

Figure C.4: Comparison of manufacturer market shares



Country-level data is computed with IDC's data.