Exploring the Walled Garden Theory:
An Empirical Framework to Assess Pricing Effects on Mobile Data Usage

Ava Chen
Princeton University

Nick Feamster
Princeton University

Enrico Calandro
Research ICT Africa

Abstract
This paper performs an exploratory study on mobile usage patterns over a period of three years (2013–2015) in the context of pricing practices such as zero-rating. In recent years, there has been a heated ongoing debate regarding whether offering different pricing plans, such as zero-rated services and applications, might slant user behaviour toward certain content on the Internet. Our study brings real data to address this research question. We shed light on this controversial issue by performing an exploratory analysis on different data plans and connection types on mobile data usage, as well as measuring quantitative and qualitative pricing effects of zero-rating on mobile data usage.

First, we perform a longitudinal exploratory study using data collected from the MySpeedTest application. We analyse differences in usage behaviour for the top five most used applications in the United States (US) versus South Africa (ZA), comparing usage on different connection types (Wi-Fi vs. cellular) as well as for devices on different data plans (prepaid vs. postpaid limited monthly data cap vs. uncapped).

Our findings show that US users consume slightly more cellular data than Wi-Fi data for most of the US top five most used applications, while South African users generally prefer Wi-Fi connections (with the notable exception of Facebook). Further, US users on postpaid plans display much higher average monthly mobile data usage than those on prepaid plans, while South African users on prepaid plans generally display much higher usage than those on postpaid plans.

Next, we perform a deeper analysis into the possible behavioural effects of zero-rating in South Africa. We find in one case that zero-rating WhatsApp on Cell-C's network increases overall usage of the application, regardless of connection type. In the case of zero-rating Twitter on MTN network, we observe increased mobile data usage of the zero-rated application during and immediately after the promotion, but not in the long term. Some of our results yield striking patterns, yet point to the need for richer datasets to confirm these initial results.

Finally, to better understand the rationale behind some of the behaviour that we observe, we implement a mobile-based survey among a randomly selected group of individuals in South Africa and Kenya. We deployed this survey to better understand the user motivations behind our empirical observations. We find that use of zero-rating services is quite low among respondents, and that zero-rating seems to serve more as a popular data conservation option and possible gateway to full Internet usage than as a walled garden discouraging respondents from venturing beyond zero-rated applications.

Keywords: zero-rating, mobile usage, mobile pricing

Introduction

Recent years have seen heated ongoing debate regarding whether offering different pricing plans might slant user behaviour toward certain content on the Internet (A4AI, 2015; Leidel, 2015; Shearsl, 2014; van Schewick, 2015). Of particular interest is the practice of zero-rating, which allows Internet service provider (ISP) customers to access and use certain mobile applications or services without incurring data usage charges. The practice is increasingly popular in both developed and developing countries, but it has stimulated a vociferous debate especially in developing countries (De Guzman, 2014; Bathia, 2016), where the costs of mobile data services are higher relative to per capita incomes (Eisenach, 2015).

Critics of zero-rating cite that the practice may violate the principle of net neutrality, which states that service providers should treat all bits of traffic flow equally, regardless of payment to the service provider (The Centre for Internet & Society, 2014; Drossos, 2015; van Schewick, 2015). A similar argument on the consumer end claims that zero-rating creates a walled garden that discourages users from venturing beyond the limited array of applications dictated by a zero-rating service provider (van Schewick, 2015; Malcom et al., 2016). The other side of the debate, spearheaded by Mark Zuckerberg and Facebook, insists that zero-rating can provide initial connectivity to consumers who otherwise would not have connectivity at all (Shearsl, 2014). According to proponents of this argument, zero-rated applications and services can serve as a gateway to eventual full Internet usage (Nowak, 2015).

This paper offers empirical data to address some of the main concerns on the zero-rating debate by investigating how various pricing-related features affect mobile data usage. The study first investigates behavioural discrepancies between users on different data plans and connection types, as well as taking into account pricing practices such as zero-rating. Second, to understand users' motivations on zero-rating usage, it explores the effects of zero-rating practices on mobile Internet usage for a simple random sample of the South African and Kenyan populations.

Obtaining empirical evidence on the relationship between mobile data usage behaviours and mobile data pricing practices such as zero-rating can benefit all stakeholders involved in the ongoing debate. Our study can help network carriers and regulators determine which data plans elicit different types of data usage, and whether introducing zero-rated services may affect user behaviour. From a policy perspective, better data may dissuade policy makers from making decisions that negatively affect competition in a dynamic telecommunications and Internet services market. Along with network carriers, policy makers and regulators, mobile applications and organizations alike can benefit from analysing possible behavioural effects of zero-rating to plan future pricing strategies to increase connectivity.

Prior and related work

There has been some related work in the areas of investigating data usage practices, and on user behaviour analysis. This paper focuses on three prior studies: two studies analyse behavioural trends on a macro scale, and the third attempts to predict user behaviour at a micro (i.e., per-user) level.

In 2014, the Sandvine Global Internet Phenomena Report (2014) published findings on fixed and mobile Internet access in various continental regions around the world. Most relevant to our topic are mobile access trends prevalent in North America and Africa. The 2014 report claims real-time entertainment and social networking applications dominate peak period mobile traffic in North America. The former accounts for 36.5% of aggregate traffic on the network, and the latter accounts for 26.36%. The high representation of social network traffic, coupled with the fact that social applications typically generate much less traffic than streaming applications, speaks to the popularity of social networking applications among users.

With regard to traffic trends in Africa, the report finds that peak-period mobile traffic is dominated by web-browsing (34.85%) and communications (28.92%) applications. In addition, Africa is the only region in which Opera Mini, a web browser focused on data efficiency, is among the top-10 most popular applications, which may suggest that users in this region are more dedicated to conserving

data usage. The 2015 Sandvine Global Internet Phenomena Report (2015) notes a slight increase in North American real-time entertainment traffic and a slight decrease in social networking traffic. Nevertheless, these two categories remain by far the dominant traffic contributors in North America. In Africa, web browsing and communications still dominate traffic composition. Notably, WhatsApp network traffic has increased by almost 50% and now contributes 10.86% of total network traffic.

In 2015, Mathur et al. (2015) performed a multi-dimensional study on data usage practices in South Africa, a region where data costs are high and usage-based data plans are prevalent. They collected 339 survey responses on mobile data usage and cost management practices from June–July 2014, conducted in-depth interviews with 43 of the survey respondents from June–August 2014, and analysed MySpeedTest data usage logs for 121 unique devices from November 2012–June 2015. The study concluded that mobile users in areas where data is limited, expensive, or both are very cost-conscious; these users frequently adopt a variety of non-trivial strategies in an attempt to optimise their mobile data usage. For example, users often switch off cellular data connections or postpone mobile use until connected to Wi-Fi, in addition to avoiding data-intensive applications and changing settings to disable automatic software updates.

In 2014, Joe-Wong et al. presented results from a time-dependent pricing trial with a commercial ISP. In an effort to optimise network capacity by analysing users' mobile data consumption, their research posits that monthly data caps tend to reduce usage, while time-dependent pricing can increase usage as users consume more data during discounted times.

Contributions of this study

Although the above studies have made significant progress both in exploring Internet access trends and mobile data usage practices and in trialling the effect of mobile data plans on mobile usage, no previous study analyses empirical data to understand how pricing instruments such as zero-rating affect mobile usage.

This study makes two contributions

1) We analyse longitudinal usage data on a diverse range of users to explore relationships between pricing effects and long-term mobile usage patterns; and
2) We collect survey responses for a more focused, comparative analysis on the motivations behind zero-rating usage in South Africa and Kenya.

Our study first explores longitudinal mobile usage patterns in South Africa as compared to the United States, the former as a precursor to our pricing analysis on zero-rating and prepaid data pricing, and the latter as a point of comparison with a high-income country. To analyse zero-rating effects on mobile usage, we focus on South Africa because we have a number of users on MySpeedTest, and zero-rating has been a popular practice in the country for some years (Futter et al., 2015). We investigate different zero-rating practices across the main carriers in South Africa and search for any changes in usage breakdowns as a result of these pricing practices. Finally, to understand users' motivations behind the behavioural patterns that we observe, we perform qualitative comparative analysis on zero-rating between South Africa and Kenya by administering a mobile-based survey to users in the two countries. We introduce Kenya into this portion of the study because of the prevalence of zero-rating data plans in the country (Walubengo, 2016).

Research Methods and Data Sources

This research makes use of two data sources: (1) application usage data from MySpeedTest application (MySpeedTest, 2016); and (2) a mobile survey administered to MySpeedTest users in South Africa and Kenya. MySpeedTest is an Android application that collects active measurements on mobile's network performance (e.g., throughput and latency), and passive measurements on user behaviour (e.g., data consumed by and usage frequency of different applications), and various pieces of metadata (e.g., country, data cap plan, battery life) when the device is on and connected to the Internet. For the purposes of this study, we focus on the relationships between different pieces of metadata associated with each device, and the mobile usage behaviours on that device.

For the longitudinal portion of our study, we analysed usage data measured by MySpeedTest from January 2013 to November 2015. Table 1 below shows the relevant tables in the database and the columns we extracted from them for our purposes, with the columns in bold used to join them together.

The process of collecting, cleaning, and pre-processing the data was an iterative one because the dataset had never before been analysed for this purpose. The MySpeedTest data was stored in a postgres database. We performed the analysis in various iPython notebooks using the Pandas and MatPlotLib libraries for analysis and visualization, respectively. MySpeedTest collects measurements at 15-minute intervals for every active device-application pair, regardless of whether any usage was recorded in a given interval. We aggregated measurements by date to make our queries more efficient, due to the size of the data that was collected. We performed all necessary cleaning, pre-processing, and aggregating directly via SQL queries to extract only the relevant data in Table 1 to store into a new table created in the database. This table was much more manageable in size to directly query. We then queried the database for any desired time period and exported the result into a Pandas dataframe for analysis in iPython.

Table 1: Relevant tables and columns extracted from MySpeedTest database.

Table	Column	Description
application	name	Application name
	package	Package name corresponding to application
application_use	**measurementid**	ID of measurement
	package	Package name corresponding to application used in measurement
	total_sent	Counter for total number of bytes sent by end of measurement
	total_recv	Counter for total number of bytes received by end of measurement
network	**measurementid**	ID of measurement
	connectiontype	Connection type used in measurement (e.g., Wi-Fi, 1G, 2G, etc.)
measurement	**measurementid**	ID of measurement
	deviceid	ID of device reporting usage in measurement
	time	Time of measurement (granularity = 15 minute intervals)
device	**serialnumber**	Serial number of device
	deviceid	ID of device
	networkname	Network carrier to which device belongs
	networkcountry	Country to which device's network carrier belongs
	datacap	Monthly data cap of device (e.g., unlimited, *X* megabytes cap, prepaid)
sim	**serialnumber**	Serial number of device
	operatorname	Network carrier to which device belongs
	networkcountry	Country to which device's network carrier belongs

We also created and distributed a mobile survey to 300 users in South Africa and Kenya (150 respondents in each country) to compare users' perspectives on zero-rating and mobile Internet usage behaviour in the two countries. To simultaneously encourage survey responses and user installs of MySpeedTest, we introduced two new avenues of recruitment for this portion of our study. First, we performed a pay-per-install method of recruitment. A commercial survey company, On Device Research, ensured 300 survey completions across an unbiased distribution of users split evenly across South Africa and Kenya, with MySpeedTest user installs performed at best effort[1]. To encourage

[1] Although On Device Research ensured unbiased distribution of respondents, the survey is not representative of the population of Kenya and South Africa.

users not to immediately delete the application from their phones, we entered users who held the application for a month into a raffle for a new phone.

Data Characterisation and Overview

Our dataset contains a total of 12,277 unique users collectively using 67,821 applications throughout the entire period of study from January 2013 through November 2015. The median number of mobile applications per device is 39, though we observe a right skew in the distribution, with a higher average of 57 applications per device.

Table 2 shows a quarterly breakdown of users by data cap type for the duration of our study. We have grouped together all data plans with monthly limits between 250 MB and 2+ GB into the "Limited" category. Note that Q4 2015 excludes December.

Table 2: Quarterly user breakdown by data cap type.

Quarter	# Users	Unlimited	Limited	Prepaid	Unknown
2013 Q1	4,797	2,150	2,068	0	579
2013 Q2	3,784	1,694	1,664	0	426
2013 Q3	2,610	1,244	1,074	0	291
2013 Q4	2,408	1,217	941	0	250
2014 Q1	1,566	812	592	1	161
2014 Q2	595	362	187	0	46
2014 Q3	1,192	756	355	17	64
2014 Q4	919	593	252	24	50
2015 Q1	522	341	150	6	25
2015 Q2	438	224	147	11	56
2015 Q3	338	175	114	9	40
2015 Q4	262	132	92	7	31

From these results, it appears that 2013 showed more active users, about an order of magnitude more than in 2015. 2014 had around 1,000 active users per quarter, with an aberration in Q2 due to some problems with the MySpeedTest application. 2015 had much lower numbers overall, with a couple hundred active users per quarter. In 2013, we see a comparable amount of users on unlimited plans as users on plans with a monthly limit. This difference in users between the two categories is amplified, as the number of overall users drops through the next two years. By 2015, the majority of users are on unlimited plans, although users on limited plans are still well represented in the data. There are at most a handful of users on prepaid plans in each quarter (throughout 2013 there were no users at all who reported being on prepaid plans). Note that the data cap associated with each device is reported by the respective user. This self-reporting may contain factual errors, as users may be unsure of their data plans and can easily misreport this information[2].

The dataset contains usage information from 1,034 US users, with 403 on unlimited plans, 627 on limited plans, and 4 on prepaid plans; in South Africa, the dataset contains usage data from 249 users, with 63 on unlimited plans, 176 on limited plans, and 10 on prepaid plans[3].

The research is structured in three different yet interlinked components: The first component compares mobile data usage in the United States vs. South Africa, focusing on the top five used apps in each country and measuring the effects on connection type and data cap on mobile usage behaviour; the second component studies the effects of zero-rating pricing on mobile data usage in South Africa; and the third components investigate the motivations on zero-rating usage in South

[2] In fact, South African data plans are predominantly of the prepaid variety, but the MySpeedTest dataset contains also users on unlimited plans (which are considerably less in South Africa).

[3] A new version of MySpeedTest has tried to improve the system of self-reporting data plan by asking users to update their data plan every quarter.

Africa and Kenya.

Comparing Mobile Data Usage in the United States vs. South Africa

The first component of our research analyses usage in the United States and South Africa during a three-year period, from January 2013 through November 2015. By comparing usage behaviours between these two countries, we aim to observe differences between usage on different connection types and under different data cap types across a longitudinal period. The research question we attempt to address is twofold: 1) whether cellular vs. Wi-Fi connections affect data usage in each country, and 2) whether post-paid unlimited vs. limited vs. prepaid data plans affect data usage in each country.

Effects of connection type on mobile data usage on top five popular applications in US vs. ZA

Before we examine any usage breakdowns based on connection types (Wi-Fi vs. cellular) as well as on different data plans, we want to explore which types of applications are most used in each country. For our comparisons, we will only consider the top five mobile applications in each country, ranked by total data usage in bytes throughout the period of study from January 2013 through November 2015.

To calculate an appropriate metric for usage of a given application by a particular category of users over a certain period of time, we first found the average monthly usage of the application across all users in that category. Then we took the median across all months in the period of interest. This made our metric more robust to potential outliers in the dataset while reflecting general usage patterns across our user base.

The top-five used applications from January 2013 through November 2015 are shown in Table 3.

Table 3: Top 5 apps by total data usage for US vs. ZA.

US	ZA
1. Netflix	1. YouTube
2. YouTube	2. Facebook
3. Facebook	3. Chrome
4. Google+	4. Correo
5. Browser	5. Google Play Store

These results are aligned with the Sandvine findings from 2014 and 2015 that North American data usage is dominated by social networking and real-time entertainment applications, as evidenced by the top-four applications in the United States: Netflix, YouTube, Facebook, and Google+. It makes sense that Netflix and YouTube would be the top contributors to mobile usage in the US, because streaming applications typically generate a lot of traffic. Facebook and Google+ are the next most used applications, attesting to the popularity of social media applications in users' mobile experiences.

The Sandvine reports declare web-browsing and communications applications to be the top contributors of African mobile traffic, which is reflected in the presence of Chrome and Correo (a mail app) in our list of South Africa's top five applications. Sandvine's 2015 report shows a significant increase in popularity of real-time entertainment, as well as a small increase in traffic contribution of social networking applications, which can explain the presence of YouTube and Facebook in the list.

In order to understand whether cellular or Wi-Fi connections affect data usage in each country we measure usage breakdown by connection type for the top five apps in each country throughout the entire period of study.

Figure 1: Usage breakdown by connection type for top 5 apps in the US.

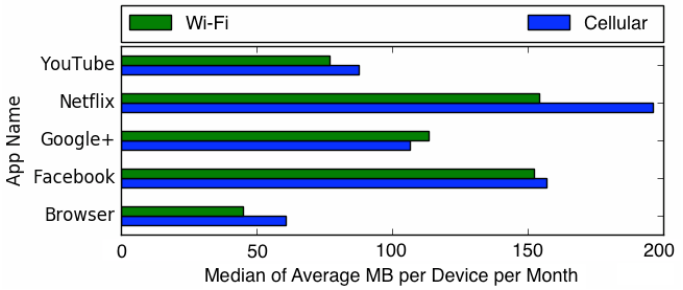


Figure 2: Usage breakdown by connection type for top 5 apps in South Africa.

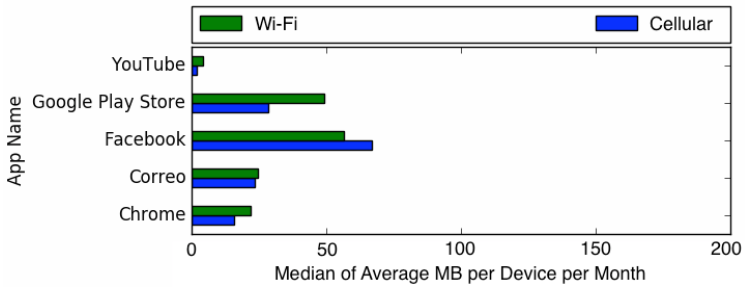


Figure 1 and Figure 2 show that US mobile usage dwarfs South African mobile usage for the top-five apps in each country. In addition, US users generally access the relevant applications on cellular connections more frequently than on Wi-Fi connections, even for data-intensive streaming applications like Netflix and YouTube. In South Africa, on the other hand, users may be wary of cellular data usage, preferring Wi-Fi connections for almost all of the listed applications. This difference may imply that US users are not as cognizant of cellular data usage as South African users, who may take more active measures against using mobile data when not in a Wi-Fi area. It also points to the possibility that mobile data in the US is more affordable than in South Africa, or that perhaps US mobile carriers offer better data plan options that allow for more cellular data usage and are typically not pay-as-you-go.

Interestingly, of the top five applications in the United States, Facebook has the second highest median of average monthly Wi-Fi and cellular usage in the United States, behind only Netflix. In fact, Facebook and Google+ report higher usage than YouTube, even though as a streaming application YouTube generally uses much more data. This finding underscores the popularity of social networking applications among users in the United States. In South Africa, Facebook has the highest median of average monthly Wi-Fi and cellular usage, and is the only application showing higher cellular usage than Wi-Fi usage. The popularity of the application indicates a preference for social media use in South Africa over other Internet activities (Calandro et al, 2012; Stork et al., 2013; Mirani, 2015). YouTube data is relatively low compared to communications and social networking applications, perhaps implying that South African users may be cautious of using data-intensive real-time streaming applications than communications and social networking applications on their mobile devices. Although YouTube was the most used application by number of bytes in South Africa, the discrepancy may be due to extremely high usage by certain outliers that ranked the application very high in terms of total data usage, whereas the metric comparing median of average monthly usage per device mitigates the influence of these outliers and reflects a lower YouTube usage on a typical device. Finally, the fact that Wi-Fi usage of Google Play Store app nearly doubles cellular usage indicates that users adopt various strategies to optimize mobile data usage, including changing settings to disable automatic software updates and postponing use until connected to Wi-Fi, as stipulated by Mathur et al. (2015). These observations all indicate a higher dedication among South African users to conserving data usage when on a cellular connection.

Differentiated effects of data cap and connection type on data usage in US vs. ZA

As depicted in Figure 3 and Figure 4 below, when we analysed usage breakdown by data cap type for the top five apps in South Africa and in the US, we found that in the US, users on limited plans dominate usage, while those on prepaid plans show no appreciable usage. On the other hand, users on prepaid plans in South Africa show much higher usage than those on other types of plans[4].

Figure 3: Usage breakdown by connection type for top 5 apps in South Africa.

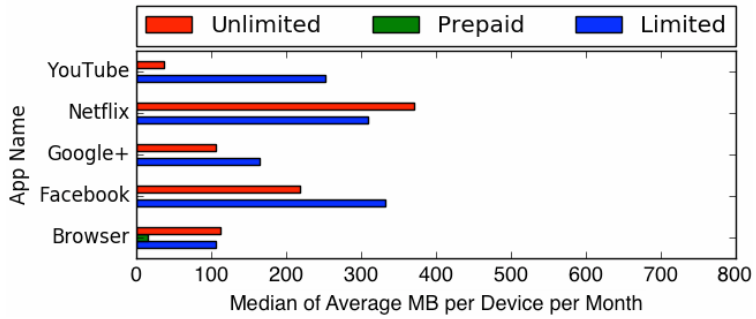
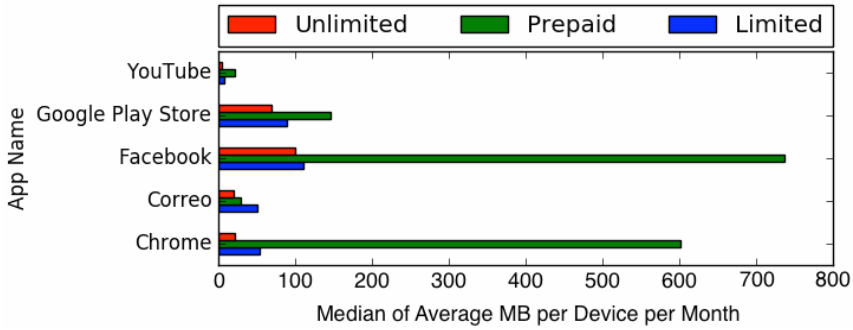


Figure 4: Usage breakdown by connection type for top 5 apps in South Africa.



We now dig deeper into the individual effects of each feature (i.e. mobile data usage on Wi-Fi vs. cellular connections and data usage by users on unlimited vs. limited vs. prepaid plans) on usage behaviour, holding the other feature constant.

In this part of the study, we studied the two applications that were in the top five for both the United States and South Africa: YouTube and Facebook. Table 4 shows the results of the usage breakdown by both features for each application in the two countries, again using median of average monthly usage as our metric of comparison.

Table 4: US vs. ZA usage breakdowns by data cap and connection type for YouTube and Facebook.

App Name	Data Cap	US		ZA	
		Wi-Fi(MB)	Cellular(MB)	Wi-Fi(MB)	Cellular(MB)
YouTube	Unlimited	24.587	10.893	2.444	1.904
	Limited	88.588	133.549	4.331	1.508
	Prepaid	0.049	0.021	8.721	0.034
Facebook	Unlimited	94.597	116.971	49.427	25.399
	Limited	144.927	290.357	42.036	68.579
	Prepaid	0.000	0.000	139.814	443.703

We first explore usage behaviours in the US, recalling that breakdown by only connection type revealed higher cellular usage than Wi-Fi usage. By holding data cap type constant and comparing usage on different connection types, we find that users on limited plans contributed most to this difference for these two applications, using both significantly more on cellular than on Wi-Fi. This finding is consistent with the fact that limited data plans in the US have higher monthly data caps, placing relatively less constraints on cellular usage. Interestingly, users on unlimited plans reported

[4] Due to the low number of users on prepaid plans (only 10) these results may exhibit some skew.

lower Wi-Fi and cellular usage than those on limited plans. Users on prepaid plans contributed almost no appreciable usage, but the small amount of usage on YouTube was higher on Wi-Fi connections than on cellular connections, as users would likely not want to use up their prepaid cellular data on a data-intensive streaming application.

With regard to South African usage behaviours, controlling for data-cap type and comparing usage on different connection types, we see that users on prepaid plans contribute the most overall usage on average. Oddly enough, cellular usage of Facebook is three times that of Wi-Fi usage of Facebook among this category of users, implying that prepaid users make up the category that contributes most to the overall higher cellular than Wi-Fi usage on the app. This finding reflects increasing pervasiveness of the social media application among users in South Africa. Looking at YouTube usage for each data-cap type, we observe that the preference for Wi-Fi over cellular usage for the app is clearly reflected in the table for users on all data cap types, confirming cellular data conservation practices for streaming applications usage.

Zero-Rating effect on mobile data usage in South Africa

In addition to the above exploratory comparisons between mobile usage behaviours in the United States and South Africa, we want to narrow the focus on the effects of zero-rating on mobile data usage, specifically in South Africa. We focus on South Africa because mobile data in the country is relatively expensive in comparison to other African countries (RIA, 2013) and zero-rating services have been offered by the major operators (Chair, 2015).

Zero rating is a relatively new phenomenon in South Africa. While MTN, the dominant operator, was the first to offer Facebook for free, as dominant market player, Cell-C has embraced zero-rated Free Basics and WhatsApp to gain market share in a *de facto* duopoly market dominated by Vodacom and MTN (Futter et al., 2015).

In the context of zero-rating, our study compares data usage on WhatsApp between the four predominant carriers in South Africa: Vodacom, Cell C, MTN, and Telkom[5]. Cell C zero-rated WhatsApp for mobile users on its network from November 19, 2014 to August 31, 2015, before switching to a bundle offer under which users could use the app for 30 days up to a fair usage cap of 1 GB, excluding voice calls, for ZAR 5 (approximately USD 0,36[6]).

We also explore Twitter usage trends, knowing that MTN had zero-rated this application in a temporary promotion from May 1 to July 31, 2014, and again during the ICC Cricket World Cup from February 14 to March 29, 2015. Although the Cricket World Cup ended in March 2015, the offer was maintained and is still active today. With this knowledge, wanted to explore whether temporal usage patterns reflected this time-dependent zero-rating.

The following table summarises the most popular zero-rated practices in South Africa between May 2014 and now.

[5] In addition, Cell C, in collaboration with Facebook, launched Free Basics on its network on July 1, 2015. Although we wanted to conduct a similar comparative study on Free Basics data usage for Cell C vs. the other network carriers in South Africa, upon searching for the app among our entire user base, we found that only a handful of users had the app installed, and they were all on a network called Airtel. Thus, due to the current lack of available data, we are at the moment unable to properly study the effects of launching the zero-rated Free Basics application in South Africa.

[6] Historical exchange rate, first week of September 2015, ZAR/USD=0.0720. Source Oanda.

Table 5: Zero-rating practices among South African carriers.

Carrier	Application	Offer	Duration
Cell C	WhatsApp	No offer	Before 11/19/14
		Zero-rated	11/19/14–08/31/15
		Bundle offer	09/01/15–present
Cell C	Free Basics	No offer	Before 07/01/15
		Zero-rated	07/01/15–present
MTN	Twitter	No offer	Before 05/01/14
		Zero-rated	05/01/14–07/31/14
		No offer	08/01/14–02/13/15
		Zero-rated	02/14/15–present

Effects of zero-rating WhatsApp on Cell C's network

We now analyse the Wi-Fi vs. cellular usage behaviours of mobile users before, during, and after the zero-rating period, when it was replaced with a bundle offer. We compare mobile usage across different mobile operators with a goal of analysing whether mobile usage behaviour reacted to these different types of promotions, transitioning from no offer to completely zero-rating the app to offering a generous amount of monthly usage for a small fee.

Figure 5 below shows WhatsApp's usage breakdown by connection type for the four main South African carriers before the period of zero-rating offer. Overall, WhatsApp usage on Cell C during this period is similar to usage on the other carriers: users exhibit comparable overall median of average monthly usage as well as displaying higher cellular than Wi-Fi usage. Interestingly, Cell C has the lowest Wi-Fi usage on WhatsApp, with a median of around 2.5 MB of average monthly usage per device. Cellular usage on the app is about three times as high, at around 7.5 MB.

Figure 5: WhatsApp usage breakdown for ZA carriers: no offer.

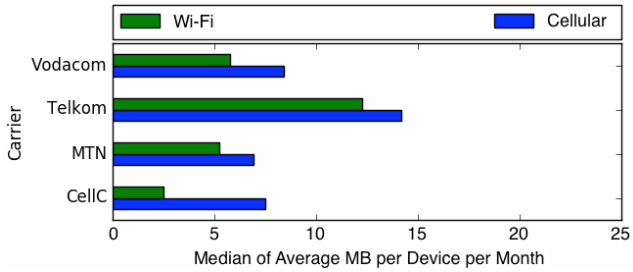
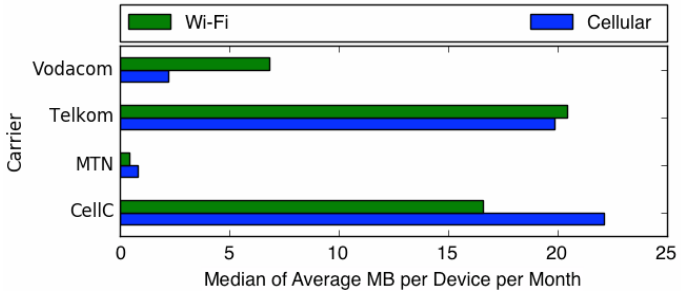


Figure 6 shows WhatsApp's usage breakdown by connection type for the four main South African carriers during Cell C's zero-rating of the application. Cell C usage increased, with Wi-Fi usage increasing almost sevenfold, to a median of around 17 MB of average monthly usage per device; and cellular usage increasing almost threefold to around 22 MB. This result indicates that zero-rating WhatsApp may have had remarkable effects on its usage[7], greatly increasing the application's use not only on cellular connections, but actually even more so on Wi-Fi connections. Overall usage on Cell C increased so dramatically during this zero-rating period, which suggests that perhaps WhatsApp became much more popular overall, regardless of connection type, which may be the result of network effect (Sundararajan, 2013).

[7] Similar conclusions on the effect of free content services on usage are found in Galpaya et al, 2015, even though Galpaya's research approach is based on nationally representative surveys. One of the findings of their nationally representative survey conducted in Myanmar shows that when telcos offer content for free, and users know about it, adoption levels are high.

Figure 6: WhatsApp usage breakdown for South African carriers: Cell C zero-rating offer.



Oddly enough, we also see behavioural changes in WhatsApp usage for the other three carriers during this period. Both Vodacom and MTN overall usage of the app decreased, while overall usage on Telkom increased. Wi-Fi usage on both Vodacom and Telkom became higher than cellular usage during this period as well. Unfortunately, we found that the number of unique users on each carrier during this period may not be high enough to draw conclusive results (see Table 6 below).

Figure 7: WhatsApp usage breakdown for ZA carriers: Cell C bundle offer.

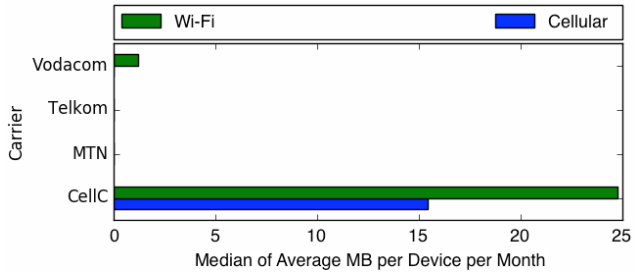


Figure 7 shows WhatsApp's usage breakdown by connection type for the four main South African carriers during its bundle offer period. We observe that the single user on Cell C does display a high monthly median usage on WhatsApp, comparable to usage seen during the zero-rating period, and certainly much higher than usage before any of Cell C's promotional offers. Yet, this finding is not statistically significant; this single data point is merely suggestive and points to the need to gather more data to explore this effect in more detail.

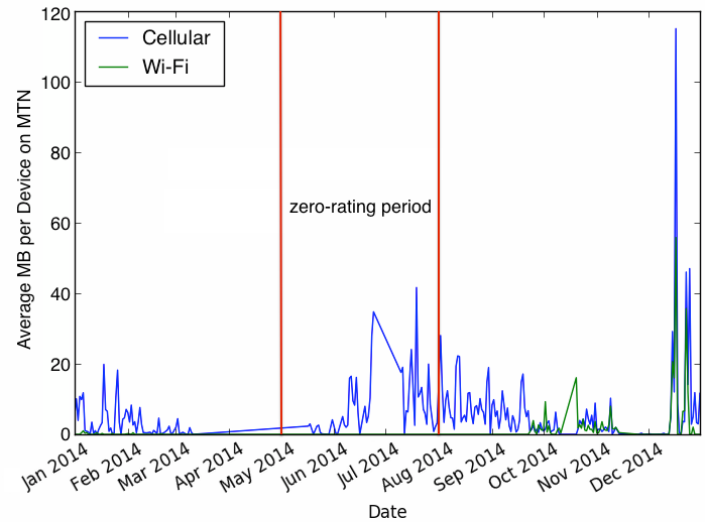
Table 6: Number of WhatsApp users for each carrier and time interval.

Carrier	Before 11/19/14	11/19/14–08/31/15	08/31/15–present
Vodacom	89	10	1
Cell C	46	4	1
MTN	60	4	0
Telkom	14	1	0

Effects of zero-Rating Twitter on MTN's network

This section examines temporal usage patterns that reflected MTN's time-dependent zero-rating of Twitter during promotional periods. Twitter was zero-rated on MTN from May 1 to July 31, 2014, and from the beginning of the ICC Cricket World Cup in February 2015 onward. Figure 8 shows the corresponding time-series plot of average daily Twitter usage per device on Wi-Fi vs. cellular connections.

Figure 8: Time series of average twitter usage per device on MTN.



Our results indicate much higher cellular usage in general than Wi-Fi usage throughout the year, aside from a small bump in Wi-Fi usage over cellular usage in October 2014. The period from the end of March through the beginning of May 2014 reflects the lack of data in the beginning of the second quarter due to a drop in usage resulting from problems with the MySpeedTest application.

The largest spike in both cellular and Wi-Fi data usage seems to correspond to the holiday season in December. This result makes sense, because social media platforms like Twitter are often used during holiday seasons for advertising and marketing, searching for and sharing gift ideas, posting wish lists, and spreading general holiday cheer (Stringfield, 2013).

It is worth noting that the period from June to July 2014 shows the second most significant peak in cellular data usage. This perhaps indicates that MTN's promotional zero-rating Twitter during this period may have had some effect on user behaviour. Cellular usage seemed to begin picking up around early June, reflecting a slight delay in users' response to the announcement of the promotional offer. Usage peaked at a little over 40 MB a day in mid-July, near the end of the event, before tapering back down throughout August and September. The fact that usage remained relatively high throughout the months immediately following the event implies that the promotion did seem to maintain some heightened level of interest and use in the application for a while, but by the end of the year cellular usage of the app had gone back down to pre-promotion levels seen at the beginning of the year. Thus, our results suggest that promotional events like zero-rating Twitter for a certain period of time may not have lasting effects on the app's popularity, though it may increase short-term usage during and immediately following the promotion. Unfortunately, because our dataset contained zero MTN Twitter users after January 2015 in our dataset, we are unable to analyse the zero-rating effects during the Cricket World Cup.

User Motivations and Attitudes Concerning Zero-rating in South Africa and Kenya

To understand user motivations behind zero-rating usage, we administered a mobile-based survey in Kenya and South Africa. We distributed our survey to 150 mobile phone users in each country, and we report and compare our main findings between the two countries. Most of our findings are consistent across both countries.

Mobile Internet usage, access, and conservation practices in South Africa and Kenya

Table 7 shows a comparison of how often respondents in each country claim to perform different activities during online mobile sessions.

Table 7: Online mobile Internet usage distribution of Kenyan and South African respondents.

Activity	Never		Hardly ever		Sometimes		Most of the time		Always	
	SA	Kenya	SA	Kenya	SA	Kenya	SA	Kenya	SA	Kenya
Email	3%	1%	8%	2%	23%	29%	26%	22%	39%	46%
Social networking	0	1%	5%	0	19%	10%	31%	26%	45%	63%
Messaging	1%	1%	2%	1%	10%	8%	22%	21%	65%	69%
Entertainment	6%	3%	7%	7%	43%	41%	19%	29%	25%	21%
Education or work-related	7%	2%	5%	4%	35%	31%	27%	31%	25%	31%

These results show that social networking and messaging applications dominate online mobile usage for both Kenyan and South African respondents. Entertainment and education- or work-related applications show less frequent usage. Overall, respondents in both countries follow similar online mobile usage patterns.

We also explore the variety and extent of data conservation practices among our respondents. To this end, we gathered responses, shown in Table 8, on how often mobile users used various methods to conserve mobile data.

Table 8: Data conservation practices among South African and Kenyan respondents.

Activity	Never		Hardly ever		Sometimes		Most of the time		Always	
	SA	Kenya	SA	Kenya	SA	Kenya	SA	Kenya	SA	Kenya
Turn off device	31%	30%	27%	37%	32%	27%	4%	5%	6%	2%
Turn off mobile data	18%	17%	16%	15%	35%	43%	17%	17%	15%	8%
Close inactive applications	7%	8%	5%	16%	28%	24%	29%	19%	31%	33%
Avoid high data applications	7%	6%	13%	8%	29%	34%	29%	26%	22%	26%
Track data usage	20%	12%	15%	22%	28%	36%	21%	12%	15%	18%

Our results indicate that closing inactive applications and avoiding data-consuming applications are the most popular methods of conserving mobile data among survey respondents in each country. Tracking data usage and turning off mobile data entirely are less commonly used alternative, while fewer respondents opted to turn off their devices entirely to conserve data. Overall, we do see a deliberate effort among respondents in both countries to conserving mobile data using these techniques. These results corroborate Mathur *et al.*'s findings (2015) that both South African and Kenyan users displayed a strong dedication toward active data conservation methods.

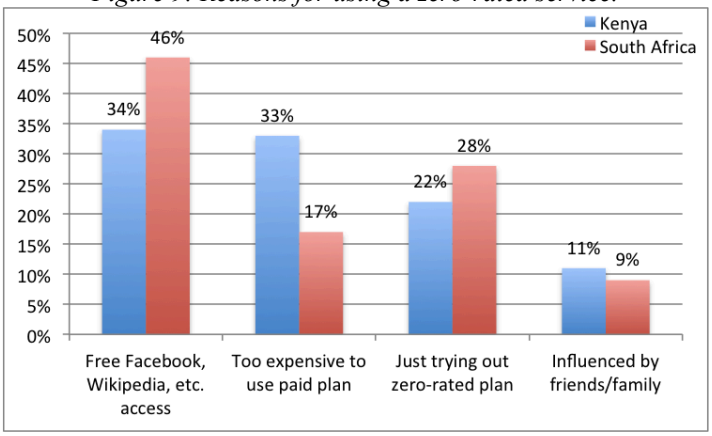
In addition, presence and effects of zero-rating practices among respondents are captured using a set of questions investigating whether users know about zero-rating services, and if so, whether they have used these services, and if they are still using them.

Regarding zero-rating services use, our survey data shows that 39% of Kenyan respondents knew about zero-rated plans offered by their mobile carrier throughout 2015; more than a half (53%) had used a zero-rated plan in the past and one quarter of them (26%) were currently using one. In South Africa, less respondents than in Kenya (23%) knew about zero-rated plans in the last year, with 29% having used one in the past and only 15% currently using one.

Figure 9 shows why respondents currently use a zero-rated plan or have used one in the past. The

largest percentage of respondents in each country claimed they had used or were currently using a zero-rated plan to access popular zero-rated applications like Facebook, Wikipedia, and Twitter for free. Compared to South African respondents, many more Kenyan respondents claimed that paid plans were too expensive, and therefore they found more convenient to use zero-rating plans, indicating that zero-rated applications were used to help conserve mobile data usage.

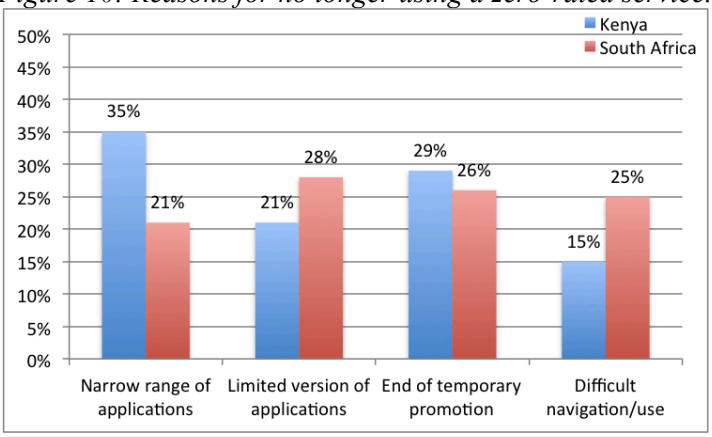
Figure 9: Reasons for using a zero-rated service.



We also observed that 59% of Kenyan respondents claimed to use regular applications and services more often than zero-rated ones, while 30% used both types equally. Only 11% used zero-rated applications more than regular ones. We found a similar distribution in South Africa, where 54% of respondents used regular applications more often than zero-rated ones, 30% used both types equally, and 15% used zero-rated applications more than regular ones. These results imply that while consumers may try to conserve data usage by using zero-rated applications, the majority still use regular applications and services more often than zero-rated ones. In fact, a majority of the respondents in both countries who have used zero-rated plans in the past were no longer using them at the time of the survey.

Figure 10 shows a distribution of various reasons why respondents stopped using a zero-rated service.

Figure 10: Reasons for no longer using a zero-rated service.



The majority of respondents in each country (65% in Kenya and 57% in South Africa) said that given the choice, they would prefer to use a non-zero-rated application with full functionality instead of a zero-rated application with limited functionality. These findings serve as a counterpoint to the walled garden argument that zero-rated applications discourage users from using the full Internet. Our survey results imply that consumers prefer to use full-service applications on paid plans, even with the option of using a limited-service application on a zero-rated plan. Our Kenya survey data includes only two users who had not used the Internet prior to using a zero-rated plan, with one now using a regular paid

plan and one not using a paid plan. In South Africa, there were no users at all for whom zero-rating served as an initial exposure to the Internet. Thus, we do not find strong evidence from our survey data of zero-rating as a gateway to full Internet usage for those who have never used the Internet before. Instead, our results indicate that zero-rated applications and services offer consumers a way to save on mobile data usage in conjunction with other methods of data conservation, an effort toward which both South African and Kenyan users show strong dedication.

Conclusion and Policy Implications

This paper presents the results from three different analyses concerning the effects of mobile data pricing on user behaviour concerning mobile data usage: (1) an exploratory analysis comparing mobile usage patterns between the United States and South Africa; (2) an analysis of the effects of zero-rating WhatsApp and Twitter in South Africa; and (3) a survey-based study on zero-rating attitudes and behaviours in Kenya and South Africa. The first portion of the study is intended to explore how the overall cost of data affects mobile data use; the second part explores how different pricing strategies affect mobile data use; and the last part explores user perceptions and attitudes that can help explain the effects that we observed in the first two parts of the study. We summarize our findings below.

Exploring mobile data usage in the United States vs. South Africa, we find that social networking and entertainment applications are popular in the United States, while the most popular applications in South Africa are more varied, ranging from web-browsing to communications to social networking to entertainment applications. Further investigation of mobile usage behaviour reveals that US users display higher overall mobile usage than South African users, and also exhibit higher cellular than Wi-Fi usage, even for data-intensive applications. On the other hand, South African users seem wary of cellular data usage, generally preferring Wi-Fi usage, which perhaps indicates more dedication to conserving cellular data usage in the latter country. While in general South African users are more cognisant of cellular data usage, Facebook presents an exception to this observation. As one of the most prevalent social media applications in both the United States and in South Africa, it displays higher average cellular than Wi-Fi data usage for both countries.

With regards to mobile usage behaviour for different data caps, US users on limited plans seem to dominate usage, and users on unlimited plans are also widely represented, but users on prepaid plans show virtually no appreciable usage. We observe a different picture in South Africa: users on prepaid plans in general show much higher usage than those on other plans.

In assessing the effects of zero-rating on mobile usage in South Africa, we find that overall usage of WhatsApp increases during a zero-rating period, regardless of connection type. This increase in popularity is also demonstrated in the case of zero-rating Twitter on MTN, where we observe increased usage of a zero-rated application on the cellular network during and immediately following the zero-rating period. However, we find that users may not be willing to continue heightened use of the application in the long term after the promotion has ended.

Analysis of qualitative survey data in Kenya and South Africa shows the popularity of zero-rated services both in Kenya and in South Africa. One of the more interesting findings is that zero-rating users prefer full-service applications on a paid plan to limited-service applications on a zero-rated plan. These findings suggest that rather than creating a walled garden discouraging users from using the full Internet, zero-rating in fact provides a popular method of data conservation.

These findings bring upon important policy implications and recommendations: Banning zero-rated services through regulation in order to preserve the open internet as we know it may limit Internet use for price sensitive users and for those who cannot afford data services otherwise. Rather, in order to reduce some of the harms that could arise from zero-rated services, policy makers should monitor the effects of zero-rating on Internet use in the medium and long terms, for both Internet and non-Internet users. As this study has demonstrated, some of the assumptions on the walled garden theory are withdrawn by simply monitoring mobile usage and how people use different plans and promotions.

Future research
A natural extension of the study would be to analyse the effects of zero-rating usage on people who are not yet using the Internet, to more conclusively differentiate between the "walled garden" and "gateway drug" effects on Internet use. Conducting a nationally representative ICT access and use survey would ensure an unbiased sample and stronger research conclusions.

We must also actively recruit more MySpeedTest users for the longitudinal portion of our study. As the conclusions of this research show, a larger and more diverse sample of users is particularly important in South. We especially want to target recruitment toward users on prepaid plans and across all network operators, to re-run our analysis code to assess pricing effects on mobile usage on a broader dataset.

Finally, it may be interesting to narrow the scope of this research from a broad exploratory analysis to a study on the predictive powers of various pricing features on mobile data usage.

Acknowledgments
We thank Dhanaraj Thakur for his help with the survey design, and Ikram Mirza and Vadim Sablo with OnDevice Research for implementing our pay-per-install recruitment plan. Thanks also to Guilherme Martins for his insights into navigating the MySpeedTest database, improving application functionality, and crosschecking results when needed. This research was partially supported by NSF Award CNS-1539902.

References

Alliance for Affordable Internet (A4AI) (2015). "Measuring the impact of mobile data services in developing countries", 20 October 2015. [Online]. Available: http://a4ai.org/measuring-mobile-data-services/

Alvarez, E. (2015). "Facebook's free Internet service expands to Malawi." EnGadget.com, 14 May 2015. Available: http://www.engadget.com/2015/05/14/facebook-Internet-org-malawi/

Bathia, R. (2016). "The inside story of Facebook's biggest setback". The long read. The Guardian. May 12, 2016. Available: https://www.theguardian.com/technology/2016/may/12/facebook-free-basics-india-zuckerberg.

Bronkhorst, Q. (2015). "Sa mobile subscribers: Vodacom vs MTN vs Cell C vs Telkom." BusinessTech, 23 April 2015. Available: http://businesstech.co.za/news/mobile/85752/sa-mobile-subscribers-vodacom-vs-mtn-vs-cell-c-vs-telkom/

Calandro, E., Gillwald, A., Deen-Swarray, M., and Stork, C. (2012). "Mobile Usage at the Base of the Pyramid in South Africa". December 2012. infoDev, The World Bank. Available: https://www.infodev.org/infodev-files/final_south_africa_bop_study_web.pdf

Chair, C. (2015). "Africa supply side assessment of Zero rating". Presentation at IGF, November 10, 2015. Available at http://www.researchictafrica.net/presentations/Presentations/2015_Chair_Zero%20Rating%20presentation_IGF2015.pdf

De Guzman, N. F. (2014). "Zero rating: enabling or restricting Internet access?". Internet Society. September 24, 2014. Available: https://www.Internetsociety.org/blog/asia-pacific-bureau/2014/09/zero-rating-enabling-or-restricting-Internet-access

Drossos, A. (2015). "The Real Threat to the Open Internet Is Zero-Rated Content," World Wide Web Foundation. Available: http://webfoundation.org/wp-content/uploads/2015/02/The-_real-_threat_open_Internet_zerorating_logos.pdf.

Eisenach, J. A. (2015)/ The Economics of Zero Rating. NERA economic consulting. March 2015. Available: http://www.nera.com/content/dam/nera/publications/2015/EconomicsofZeroRating.pdf

Futter, A., and Gillwald, A. (2015). "Zero-rated Internet services: What is to be done?" Policy Paper 2015: Broadband 4 Africa September 2015. Available: http://www.researchictafrica.net/publications/Other_publications/2015_RIA_Facebookzerorating_policy_paper.pdf

Galpaya, H., Zainudeen, A., and Suthaharan, P. (2015). A baseline survey of ICT and knowledge access in Myanmar. August 2015. Available: http://lirneasia.net/wp-content/uploads/2015/07/LIRNEasia_MyanmarBaselineSurvey_DescriptiveStats_V1.pdf

Internet.org (2016). "Free Basics by Facebook". Available at https://info.Internet.org/en/story/free-basics-from-Internet-org/

Joe-Wong, C., Ha, S., Sen, S., and Chiang, M. [2014]. Do Mobile Data Plans Affect Usage? Results from a Pricing Trial with ISP Customers. Available: http://www.princeton.edu/~cjoe/MTA_PAM.pdf

Leidel, S. (2015). "Zero rating: Why are people using a half-baked Internet." DW Akademie, 12 February 2015. Available: http://www.dw.com/en/zero-rating-why-are-people-using-a-half-baked-Internet/a-18887956

Malcom, J., McSherry, C., Walsh, K. (2016). "Zero rating: What it is and why you should care." Electronic Frontier Foundation, 19 February 2016. Available: https://www.eff.org/deeplinks/2016/02/zero-rating-what-it-is-why-you-should-care.

Mathur, A., Schlotfeldt, B., and Chetty, M. (2015). "A mixed-methods study of mobile users' data usage practices in South Africa" in Proceedings of the 17th ACM International Conference on Ubiquitous Computing , ser. UbiComp '15. ACM, 2015, pp. 1209–1220. Available: http://dblp.uni-trier.de/db/conf/huc/ubicomp2015.html#MathurSC15

Mirani, L. (2015). "Millions of Facebook users have no idea they're using the Internet" Quartz, February 9, 2015. Available: http://qz.com/333313/milliions-of-facebook-users-have-no-idea-theyre-using-the-Internet/

Mutegi, L. (2015). "Ghana: Facebook, airtel partner to bring Internet.org app to Ghana." AllAfrica (Nairobi), 26 January 2015. Available: http://allafrica.com/stories/201501270091.html

MySpeedTest (2016). MySpeedTest by GTNoise. Available at https://play.google.com/store/apps/details?id=com.num&hl=en

Nowak, P. (2015). "Why 'zero rating' is the new battleground in net neutrality debate." CBC News: Business, 07 April 2015. Available: http://www.cbc.ca/news/business/why-zero-rating-is-the-new-battleground-in-net-neutrality-debate-1.3015070.

Research ICT Africa (RIA) (2013). "How do mobile and fixed broadband stack up in SA?" RIA Policy brief South Africa No 2, July 2013. Available at http://www.researchictafrica.net/polbrf/Research_ICT_Africa_Policy_Briefs/2013_Policy_Brief_2_How_do_mobile_and_fixed_broadband_stack_up_in_SA.pdf

Sandvine Intelligent Broadband Networks (2014). "Global Internet phenomena: 2h 2014.", 21

November 2014. [Online]. Available: https://www.sandvine.com/downloads/general/global-Internet-phenomena/2014/2h-2014-global-Internet-phenomena-report.pdf

Sandvine Intelligent Broadband Networks (2015). "Global Internet phenomena report: Africa, middle east & north america.", 04 December 2015. [Online]. Available: https://www.sandvine.com/downloads/general/global-Internet-phenomena/2015/global-Internet-phenomena-africa-middle-east-and-north-america.pdf

Shearsl, M. (2014) "No. 208 net neutrality, zero-rating & development: What's the data?" Center fo Democracy and Technology, 2014. Available: https://www.intgovforum.org/cms/wks2014/index.php/proposal/view_public/208

Smith, D. (2014). "Internet use on mobile phones in africa predicted to increase 20-fold." The Guardian, 05 June 2014. Available: http://www.theguardian.com/world/2014/jun/05/Internet-use-mobile-phones-africa-predicted-increase-20-fold

Stork, C., Calandro, E., and Gillwald, A. (2013). "Internet going mobile: Internet access and use in 11 African countries", info, Vol. 15 Iss: 5, pp.34 – 51.

Stringfield, J. (2013). "Study: Twitter's influence on holiday shopping." Twitter, 07 October 2013. Available: https://blog.twitter.com/2013/study-twitters-influence-on-holiday-shopping

Sundararajan, A. (2013). "Network Effects". Available: http://oz.stern.nyu.edu/io/network.html

The Centre for Internet & Society (2014). "What is net neutrality and why it is important", 20 January 2014. Available: http://cis-india.org/news/times-of-india-january-20-2014-what-is-net-neutrality-and-why-is-it-important

van Schewick, B. (2015). "Net Neutrality and Zero-rating", filing at Federal Communications Commission. Available: http://apps.fcc.gov/ecfs/document/view?id=60001031582

Walubengo, J. (2016). Zero-rating of mobile services in Kenya. Under publication