CONFERENCE

Strategies to Overcome Poverty & Inequality

"Towards Carnegie III"

University of Cape Town, 3 - 7 September 2012

Tweeting out of poverty: Access to information and communication technologies as a pathway from poverty

Julian May[1]

1 Introduction

The information deficiencies concerning the impact of ICT on poverty reduction have raised concerns among policy makers who are being repeatedly urged to invest a substantial part of the national budget in ICT infrastructure on the basis of an incomplete evidence base. In this situation, it is tempting to question whether investments in ICT represent a worthwhile option for poor communities. A common mistake made by many existing studies is the collection of data at a level that is too general leading to an over investigation of macro- over micro-level trends. Because changes in the well-being of individuals and households are not necessarily directly linked to changes in economic output at the national level, it is important to extend the analysis of the ICT-poverty nexus beyond the national and toward the household level. Using data from an East African case study, this paper seeks to make a contribution to this discussion. I use three multivariate approaches to examine the likely links between changes in poverty status and changes in access to ICT.

2 Conceptual Framework and Methodology

3 Whatever dimension of welfare change is being considered, the direction of its causal link to ICT is contentious. Problems of reverse causality and spurious correlation that apply to the relationship between any investment in infrastructure and increasing output are equally of relevance to the analysis of the ICT/poverty nexus. Recognising this, the research question for this paper can be stated as follows: "what change in poverty status results from a change in ICTs usage, taking into account confounding factors such as socio-economic and demographic characteristics." In formal notation this can be represented as:

$$\Delta P^{\alpha} = \Delta\theta + K + B + \varepsilon \quad (1)$$

where P^{α} represents measures of well-being outcomes, θ is a measure of ICT access, K are the assets possessed by the household, B represents the observable socio-economic and demographic characteristics of the unit of analysis (households and individuals), and ε is an error term that includes unobservable characteristics. The explanatory variables producing a change in θ will need to be broken down as will the impact indicators implicit in P^{α}.

To do this I define θ_t as the ICT access gained by a household at end of the study period, t. Following a similar approach to that used for other forms of impact assessment, the research question can then be framed as the impact of the change in ICT status on the change in household's well-being level, controlling for possible confounders in the initial period ($t-1$)

[1] The contribution of Vaughan Dutton and Louis Manyakazi in previous drafts of this paper are acknowledged as is funding received from the IDRC.

where I define this impact as:

$$P_t^{\propto} - P_{t-1}^{\propto} = (\theta_t - \theta_{t-1}) + K_{t-1} + B_{t-1} + \varepsilon \tag{2}$$

ICT access will be proxied by the reported change in the availability of new forms of ICT (mobile phones, SIM cards and email addresses) to the household between the initial (*t-1*) and the subsequent period. I make use of the asset-vulnerability framework and the notion of a five-fold asset portfolio developed by Moser (1998) to identify the outcome variable and the household assets which must be controlled for. Mindful of its limitations, I adopt a money-metric indicator of well-being and use monthly *per capita* household expenditure adjusted for domestic inflation rates and purchasing power parity (PPP) as the proxy. Here deprivation is conceptualised as the inability to attain an absolute minimum standard of living reflected by a quantifiable and absolute indicator applied to a constant threshold that separates the poor from the non-poor (Ravallion 1995).

While many methods exist by which to calculate national poverty lines, I adopt the commonly used upper and lower international poverty lines of PPP$2.50 dollars and PPP$1.25.[2] The monthly expenditure per adult (PCE) of the sampled households was first adjusted by domestic inflation rates to 2005 prices and the converted into PPP$. I also generated a poverty score based on the upper bound poverty line which was then normalised against household size. Thus a score of 1 is a household whose PCE is equal to the poverty line, while 0.5 would be a household whose PCE is equal to half the poverty line.

A growing literature stresses the importance of the underlying economic stocks that determine income flows (Reardon and Vosti 1995; Carter and May 2001; Carter and Barrett 2006). I take account of this and include four forms of capital for which I have measures: economic, human, physical and social.[3] Due to data limitations the indicator for these stocks is restricted to the number of assets owned by each household. Physical capital reflects inadequate access to essential services and is largely derived from a basic needs approach to development. Indicators relating to housing and access to essential services have been combined into a single index donating access to physical capital (Alampay 2006).[4] Following Fiadzo et al (2001), de Vos (2005) and others I have chosen not to use principal components or factor analysis to develop these indices, and instead opt for an approach that theorises a structural relationship between the components of each of the uni-dimensional measures of housing quality and networks. The variables to be used are selected by assessing their inter-correlations, item-rest correlations, calculating Cronbach's *α,* and then excluding components that increase α (and which I assumed to be measuring other dimensions of deprivation). This proxy ranges from 0 (the dwelling is constructed of impermanent materials and no services are provided) to 5 (the dwelling is constructed of permanent materials and all services are provided). Human capital is measured by the presence of at least one household member who has completed their secondary education. Finally, the absence of social capital is recognised as a dimension of deprivation, and for the

---

2 The most recent round of PPP price data was collected in 2005 by the International Comparison Program. See http://siteresources.worldbank.org/ICPEXT/Resources/ICP_2011.html for details concerning methodology and application.

3 The five-fold asset portfolio usually includes natural capital.

4 The indicators are: bricks or blocks are used for walls; floors are cement or cement plus a covering; there is access to electricity; there is access to a protected water source; and there is access to a flush or improved toilet.

purposes of this study I focus on participation in social institutions through group membership, as well as participation in governance processes (Coleman 1988).

In this paper, unless indicated otherwise, 1508 and 8049 are the sample sizes (n) from the PICTURE-Africa data used in all tables and figures referring to households or household members in the first wave of data collection (called Cross Section 1, or CS1) while 1092 households and 5783 members are the sample sizes used when referring to the second wave of data collection (Cross Section 2 or CS2), and the matched sample. The difference in the sample size reflects attrition arising both from the failure to find households surveyed in the first wave (loss to follow-up) as well as from fieldwork errors that prevented surveyed households from being identified and linked in both surveys.

It is important that I considered the likely effects of attrition given the panel study research design. When working with panel data, attrition is almost always present and refers to the circumstance in which a proportion of the respondents from the first cross section are not present in the second cross section. This may be due to a large number of reasons: respondents may have moved out of the area; they may have passed away; they may have been away from the household at the time of visit; and so on. Attrition is not necessarily a problem and, at times, can be ignored. At other times, however, it has the ability to bias findings to a significant degree, introducing false patterns into data. In this circumstance, in which continuing members differ systematically from those who drop out, the sample of continuing members is no longer representative of the original population. This may result in erroneous conclusions being drawn and bad recommendations being made. In this sense, a pattern in the attrition refers to whether there is a correlation between attrition and a variable of interest. In this circumstance, attrition would be biasing the variable. Following Baulch and Quisumbing (2010) I chose to deal with attrition by means of a three phase methodology. The first phase involved identifying whether attrition was indeed present; if so, which variables it was related to, if any; and third whether the relation between these variables and attrition was random (in which case no correction would be required). Finally, if the attrition was not found to be random, the calculation of weights would be required.

The first phase required that the panel be examined for the presence of attrition and reveals that 35% of respondents included in Cross Section 1 (CS1) were not present in the second wave, Cross Section 2 (CS2), or which were successfully interviewed but could not be matched to CS1. This indicates that attrition was present in the sample and therefore requires analysis.

Having established that attrition is indeed present in the data, it is necessary to identify whether or not there is a pattern to the attrition which would cause a bias to be introduced to the second cross section. This requires, as a first step, the identification of which variables correlated with attrition. Thereafter, the degree of randomness between attrition and the variables of interest was assessed. If attrition is found to be random, no corrective measures are required. If, however, it is found to be non-random, weighting procedures would have to be implemented as a corrective.

In order to identify whether attrition is related to particular variables, and if so which ones, I ran a logistic regression. Attrition was set as the dependent (0=did not attrite; 1=did attrite). Independents included all the variables of interest (the six dimensions of poverty) as well as variables which were not of theoretical interest but which may nonetheless affect attrition [I used age, location (rural or urban), and household size].

Table 1*: *Variables in the equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.647	.029	509.812	1	.000	1.910
		Chi-square		df			Sig.
Step 1	Step	346.332		10			.000
	Block	346.332		10			.000
	Model	346.332		10			.000
Step	-2 Log likelihood		Cox & Snell R^2			Nagelkerke R^2	
1	6600.075(a)		.062			.086	

Table 5 indicates that the constant was significant and, therefore, that the null model should be rejected (i.e., the regression was valid). The model coefficients indicate how well the model fits the data and show significance, which implies that the inclusion of the independent variables into the model (beyond the mere constant) is justified. Finally the result indicates that the Nagelkerk R^2 is very low (0.086).

This is an important output, because it indicates that the baseline variables and attrition explain about 9% of the panel attrition between 2007 and 2010. The Nagelkerke R square can be taken as an indication of the degree of randomness of attrition. In this case, the attrition is 91% random. This finding therefore indicates that corrective measures for attrition such as weighting the data are not necessary in this instance. This report therefore makes use of unweighted matched data from wave one and two of data collection.

4 Poverty dynamics 2007/8-2010

I start by comparing the standard measures of money-metric poverty and inequality for the sample in the two cross-sections in Table 2.

Table 2: Per capita expenditure and financial poverty status

	2007/8	2010
Headcount (P^0)	55.0	58.4
Gap (P^1)	23.5	27.6
Severity (P^2)	13.2	16.5
Tanzania P^0	57.9	53.5
Kenya P^0	46.1	55.9
Rwanda P^0	53.3	61.1
Uganda P^0	60.7	62.7
Urban P^0	38.7	45.9
Rural P^0	68.3	66.4
Gini coefficient	0.48	0.48
n	1476	1086

Although there is no change in inequality (as shown by the Gini coefficient), the data reveal that the households surveyed in CS1 and CS2 experienced an increase in poverty. For these, poverty rose from 55% of those surveyed in 2007/8 to 58% in 2010. This is true for all measures of poverty and particularly for the poverty gap, confirming that on average, poor households had slipped further below the poverty line by 2010. If I restrict the analysis to the matched sample only (households surveyed and linked in both waves), the headcount is marginally lower in 2010 than in 2007/8 but the Gap and Severity measures reveal the same trend of increasing poverty.

I can represent these changes using the poverty score in a cumulative distribution function (CDF) in which the poverty score is represented on the horizontal axis and the percentage of the matched sample is shown on the vertical (Figure 1). A score of one on the horizontal axis represents the upper bound poverty line of PPP$2.50. This allows us to compare changes in well-being over the full distribution of per capita expenditure.

Figure 1: CDF of household poverty score

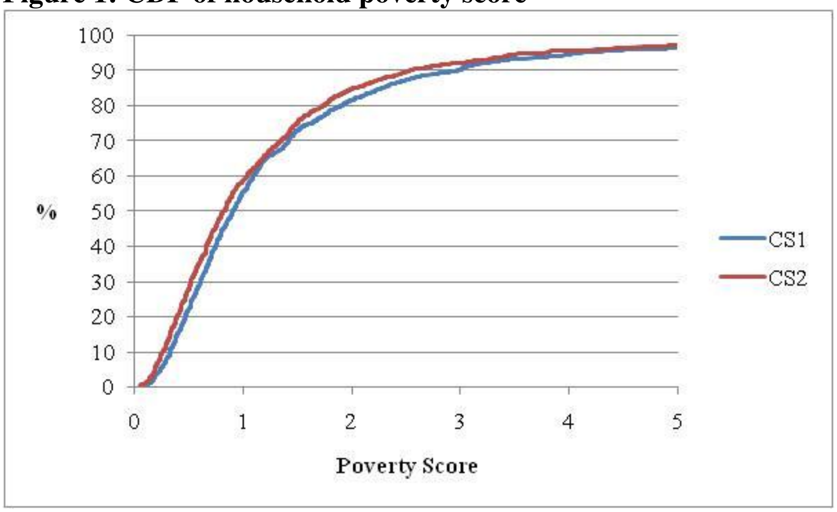


Figure 1 shows that up until four times the international poverty line, households were poorer in 2010 than in 2007.[5]

Table 1 also shows that these changes in poverty did not affect all households equally and poverty rates increased in urban areas and decreased in rural areas. There were also differences between the four countries. Only Tanzania experienced a modest decline in poverty while in Kenya poverty rates increased by almost 10 percentage points. East Africa was not the only region where poverty increased during this period. The World Bank's Global Monitoring Report for 2010 estimates that the economic crisis left 50 million more people in extreme poverty in 2009 worldwide, with an additional 64 million expected to fall into extreme poverty by the end of 2010 relative to pre-crisis trends (World Bank, 2010). These data suggested that the impact of the crisis fell most heavily upon the urban population and perhaps on those regions more integrated into the global economy. The Tanzania data suggests that, at least during the survey period, households may have been initially buffered from the crisis.

Finally, the poverty dynamics of households surveyed in both waves can be seen in a transition matrix which reveals the extent of chronic poverty (households observed to be poor in both

[5] Levene's Test for Equality of Variance and the T-Test for Equality of Means confirmed that the means and distribution of households in 2007 and 2010 were statistically different for the poverty score and PCE. Poverty deficit and poverty severity curves (not shown) confirm the observed increase in poverty between waves.

waves) as well as movements into and out of poverty. This is shown in Table 3 in which the cells represent the percentage of households that had a per capita expenditure of less than half the poverty line (PL), between half and the PL, one to two times the PL and more than twice the PL. The diagonal line shown in italics reflect households who were found in the same PL band in both 2007/8 and 2010, while those to the right of this are households whose position had improved.

Table 3: Transition Matrix 2007/8-2010

2010 2007/8	< 0.5 PL	0.5-<1PL	1PL-<2PL	2Pl +	n
< 0.5 PL	***53.7***	28.3	14.3	3.7	244
0.5-<1PL	29.6	***37.4***	25.6	7.5	348
1PL-<2PL	15.8	32.0	***28.7***	23.5	247
2Pl +	8.1	15.6	38.1	***38.1***	160

Table 3 suggests considerable mobility and that despite the general increase in poverty during the survey around 29% of all households managed to improve their 2007/8 position by 2010, while a similar share fell into a lower poverty line category in 2010. Such 'churning' has been observed in many studies making use of panel data (Baulch and Hoddinott 2000). However 40% of the sample were poor in both periods and are potentially caught in a poverty trap.

5 ICT dynamics 2007/8-2010

Attention can now be turned to the access and use of ICT. In this report, 'access' is taken to mean ownership, although access can also be obtained through using ICT that are owned by others or through public access points such as internet cafes or public telephone booths. Table 4 shows the ownership of all forms of ICT for the four countries in 2010 and compares the results from 2007/8 and 2010.

Table 4: ICT ownership of households (%)

ICT	2007/8	2010	Tanzania	Kenya	Rwanda	Uganda
Radio	70.2	72.9	65.3	79.2	76.6	70.5
TV	24.1	24.5	24.5	39.3	23.8	10.5
VCR	13.2	16.1	15.1	24.8	20.9	3.6
Landline	1.5	0.5	0.0	0.3	1.4	0.4
Email	17.8	15.2	7.4	20.1	25.9	7.6
Mobile phone	61.7	67.8	63.7	82.5	71.6	53.5
Computer	4.2	5.0	4.0	7.1	8.0	0.4
Internet connection	0.7	0.5	0.0	1.2	0.9	0.0

Access was highest in the Kenya sample for the radio, television, VCR and the mobiles but had some striking exceptions such as access to the land line, computers, internet connections and e-mail addresses (in other words most of the digital ICTs). Rwanda was better off than other countries in terms of computers, land lines and email addresses. Access to most forms of ICT had increased, with the exception of land lines, internet connections and email addresses, perhaps due to attrition of younger members of the surveyed households.

Radios are the most commonly owned ICT asset but this is closely followed by mobile phones which range from 54% of sampled households in Uganda to 83% in Kenya. The very low percentage of households with a land line or internet connection and computers is indicative of a major constraint to the delivery of privately owned internet access through conventional technologies common to most countries in Africa. The low level of access to the internet may also be due to the low levels of education in general and skills in using computer based ICT in particular among the majority of the sampled households.

Table 5 confirms that combined access to most forms of ICT increased between the two waves of the survey. This includes new forms of ICT, notably mobile phones and SIM cards. This gain was not, however, geographically equal.

Table 5: Changes in ICT status 2007/8-2010

	Total	Rural	Urban	Tanzania	Kenya	Rwanda	Uganda
Gained ICT	48.3	44.3	59.5	38.1	57.7	48.8	50.6
Lost ICT	18.1	24.4	11.5	13.0	4.5	4.6	45.5

Almost 60% of households in Kenya that did not have access to new forms of ICT in 2007/8 had gained such access in 2010, compared to less than 40% of this group in Tanzania. Households in Kenya also reported the smallest loss in ICT with just 4% of those households that had ICT access in 2007/8 losing this in 2010, while households in Uganda appear to have experienced the greatest change, with an almost equal percentage of households changing their ICT status. In terms of socio-economic characteristics, households that gained access to ICT tended to be urban, not poor, and to have at least one household member with secondary education. In comparison to those that lost ICT by 2010, they tended to be rural, poor, and without an educated family member. All differences are statistically significant. These changes in access to the new forms of ICT are strongly associated with changes in poverty status (Table 6).

Table 6: ICT dynamics 2007/8-2010

ICT Status	Change in PPP$ p/m CS1-CS2	PCE $PPP p/m 2007/8	PCE $PPP p/m 2010	Change in Poverty Score	% Poor 2007/8	% Poor 2010
Never had ICT	-$1.86	$57.66	$56.13	-0.02	81.9	79.6
Always had ICT	-$9.45	$125.57	$118.94	-0.12	45.5	45.8
Gained ICT	$20.76	$72.98	$92.84	0.27	72.5	62.6
Lost ICT	-$6.58	$72.19	$65.85	-0.09	64.3	70.3

On average, only those that gained access to ICT experienced a real gain in their *per capita* monthly expenditure between CS1 and CS2, with their incomes increasing by just less than $21 per month, or by almost 30% on their CS1 monthly PCE. This is even more apparent when looking at the mean change in the poverty score (which is also the proxy for financial capital) which shows that while all other groups experienced a decline in their poverty status (thus became more poor), those gaining ICT experienced a gain of 28% on the poverty score (thus improved their wealth). This is striking since 50% of this group were categorised as poor in CS1 and CS2 and would otherwise be potentially trapped in poverty. Those that reported no access to ICT in both CS1 and CS2 were the poorest of the four groups in both waves, with around 68% of this group found below the $2.50 per day poverty line on both periods. This is in marked contrast to those that had ICT in both periods of whom 30% were poor in both periods. Although this group experienced the largest loss in income during the survey period, this amounted to only 7% of their PCE in 2007/8 compared to the 9% drop experienced by those who lost access to ICT. These differences can also be observed in some of the other dimensions of poverty. As provocative as these findings are, they do not demonstrate causality - it is possible that the increase in income caused households to gain access to ICT rather than the reverse. A more carefully constructed counterfactual is required.

6 Impact of ICT access

6.1 Impact of poverty on ICT access

The effect of improved income on gaining access to ICT seems the most intuitive causal link between ICT and poverty and this will be the point of departure. I thus first consider the association between household ICT access in both waves and their poverty score, along with a number of control variables and various dimensions of deprivation using a binary logistic regression. The results are shown for the sample in 2007/8 and 2010 in Table 7.[6]

Table 7: Predictors of household ICT access[7]

	2007/8			2010		
	B	**Wald**	**Exp(B)**	**B**	**Wald**	**Exp(B)**
Location (Urban)	**.557**	11.419	1.746	**.483**	5.238	1.620
country dummy1	**.677**	9.838	1.969	**.819**	8.336	2.268
country dummy2	.0.18	0.009	1.018	**.927**	14.807	2.527
country dummy3	**.987**	23.274	2.684	-.076	.101	.927
Logged poverty score	**.801**	13.426	2.228	**.996**	16.192	2.707
Social capital	0.065	.602	1.068	**.247**	6.753	1.280
Economic capital	**.372**	33.761	1.450	**.326**	8.207	1.386
Physical capital	**.281**	24.696	1.325	**.386**	26.924	1.471
Human capital	**1.490**	104.811	4.435	**1.349**	52.419	3.854
Constant	-1.829	87.629	.131	-1.585	36.247	.205
-2 Log likelihood	1381.892			919.515		
Nagelkerke R Square	.320			.280		
Cox & Snell R Square	.435			.393		
n	1462			1003		

[6] The outcome variable was found to be skewed and a log transformation was carried out to normalise it to an acceptable level for our analysis. All variables pass a null hypothesis test that they do not significantly increase our ability to predict ICT access when entered. Further the pseudo R^2 statistics suggest that the predictions of the model are reasonably robust. The Wald Chi-square tests the unique contribution of each predictor holding all other predictors constant. All predictors meet the conventional 0.05 level of statistical significance with the exception of the country dummy variables which is expected, and social capital which suggests that this proxy is weak.

[7] Coefficients that are significant at the 0.05 confidence interval are shown in bold.

As anticipated the model shows positive and significant associations between all the dimensions of poverty and ICT access except for social capital in 2007/8. The odds of gaining access to ICT are more than doubled by a unit improvement in the logged poverty score. Stronger results are found for the proxy for human capital (I used education). The odds of a household containing a member with secondary education are around 4 times those of households without this asset. The odds of having ICT in urban areas are just over 1.5 times those in rural settlements. Excluding the country fixed effects does not change the sign or significance of any of the predictors but does increase the contribution to the model that is made by differences in *per capita* expenditure. This reflects the differences in the poverty incidence of the four countries.[8]

6.2 Impact of ICT access on poverty

Having shown that a higher household poverty score improve the odds of having ICT in both years I shift attention toward investigating any casual link between ICT and the underlying analytical model. In the context of both my focus and the available data, this presents several challenges. Firstly the pathways through which ICT influence changes in poverty status remain under-theorised and offer little guidance in terms of variables that should be included, excluded or which used as instruments. Second, I have neither an experimental design nor a clear counterfactual. Finally, I have already demonstrated the strong reverse causation whereby higher incomes are shown to increase the odds of access to ICT. Mindful of these limitations, I focus on two approaches. I chose to make use of lagged access to ICT as a predictor of the change in the poverty score. Here I hypothesise that prior access of ICT cannot have been determined by the rate of change in the poverty score between the two survey periods. In a methodology analogous to Granger testing used in time-series analysis, I test for the statistical significance of the coefficients of the exogenous variable (ICT access in 2008) and the poverty score in 2010 when I include the poverty score in 2008 as an explanatory variable (Menard 1991). The logic of this method holds that x causes y if I am better able to predict y with all the possible causes than with all possible causes minus x (Granger 1969). The causal variable of interest, x, is therefore evaluated alongside a range of other possible causes rather than on its own - as might be the case in a simple correlation. The test therefore requires the estimation of two models: one with the independent of interest (unrestricted model) and one without (restricted model). The restricted and unrestricted models are then compared in order to yield an estimate of the effect of the restriction (which is then conceptualised as the strength of causal influence exerted by the excluded variable). In the data, the causal variable is ICT access and the dependent is the logged poverty score. Other possible causal variables include a number of household variables. The unrestricted model therefore includes a range of household variables as well as ICT access; while the unrestricted includes only the household variables without ICT access. Should I be able to better predict expenditure with ICT access included in the model, I will be able to conclude that ICT access *causes* changes in the poverty score in at least a statistical sense. It must be cautioned that this is not equivalent to absolute causality, and should rather be seen as a 'smoking gun' that might indicate the presence of causation.[9]

To manage some of these difficulties I extended the analysis to incorporate a first differences

[8] The Wald chi square for the income predictor more than doubles to 35.4 while the odds ratio increases to 3.2.

[9] We are, however, unable to fully account for auto-correlation in which the values of our variables of interest in the second wave of data collection depend on values of the same variables in previous periods (Hood et al., 2008).

model (Liker et al, 1985). The logic underpinning this approach is to estimate the difference between the outcome measure, financial poverty, at CS1 and CS2 for those who experienced an improvement in their ICT status and those who did not. These are the treatment and control groups. I then compare the difference between the two groups, controlling for other possible prior determinants of this change. This should ensure that any variables that remained constant between waves, but which are unobserved and which are correlated with the decision to acquire additional ICT capability and with the poverty score will not bias the estimated effect of the treatment (Ashenfelter 1978; Ashenfelter and Card 1985). To minimise the risk of a selection bias error, I first use propensity score matching to identify a sub-sample of households that are matched in terms of the socio-economic predictor and control variables (Rosenbaum and Rubin).

Although the first differences model has the advantage of relative simplicity given the availability of panel data, it requires that I make the crucial identifying assumption that the counterfactual trend (what would have happened in the absence of ICT access) is the same for the treated and non-treated units. This is necessary to obtain impact estimates of the gains resulting from acquiring ICT between the two waves of the survey. This is a potentially contentious assumption, since it is still possible that the gain in income was the reason why the ICT was acquired. I could better take account of this if three waves of data were available, in which case I would have been able to more clearly identify the prior trend of the treated group. This would have allowed for a difference-in-difference analysis. Without this third cross section, I maintain that consistent findings from my combination of the lagged approach and differences model will offset this potential weakness in my methodology.

6.3 Lagged access model

In practice, the lagged variable test is a regression, in which the dependent variable is predicted using all the predictors and then using all predictors minus the independent of interest. The independent of interest (ICT access) is lagged in order to take advantage of the temporal dimension in the data. The intention is that prior ICT access cannot be affected by current levels of *per capita* expenditure (PCE), nor should prior ICT access be affected by the growth in PCE between waves. Thus, any relationship that is shown is more likely to reflect the outcome of prior ICT access on PCE. The difference between the corresponding model residual sum-of-squares (-), and the associated degrees-of-freedom ($T - 2p - 1$), make up the test (SAS 2009; Hood et al. 2008; Hurli and Venet 2003):

$$S_1 = \frac{(RSS_0 - RSS_1)/p}{RSS_1/(T - 2p - 1)} \sim F_{p,\,T-2p-1} \qquad (3)$$

An asymptotic version of the test is:

$$S_1 = \frac{T(RSS_0 - RSS_1)}{RSS_1} \sim \chi^2(p) \qquad (4)$$

The F-test is a common instrument to evaluate the significance of a parameter restriction in an ANOVA (linear models). The null and alternative hypotheses are:

$$H_0\colon \beta_1^{(1)} = 0$$
$$H_\alpha\colon \beta_1^{(1)} \neq 0$$

or

H_0: independent did not cause the % change in PCE in 2010

H_α: not H_0

The rejection of the null hypothesis provides evidence of causality of one on the other. Failure to reject means that the independent of interest does not statistically cause changes in the dependent.

In the analysis, the independent of interest was ICT access (γ) and the dependent was logged per capita expenditure in national currency *(y)* adjusted for domestic inflation.[10] In sum, the variable might be described as the 'log of total expenditure per capita in national currency units adjusted for inflation'. Further, the independent variable of interest (representing ICT access) was constructed as an index. The construction of an index was considered pragmatic because it allowed the inclusion of the greatest amount of information (or conversely, the loss of the least amount of information). A principal components analysis (PCA) was employed in order to create the scale. The data contain many candidate variables with which to construct the multidimensional index of ICT access. The following variables were collected in both waves: *per capita* mobile access, *per capita* email access, landline, television, and internet. These variables, however, have some natural redundancy; that is the variables are not uncorrelated with one another. This makes the use and interpretation of ICT access very difficult. The PCA allows creating indices, commonly referred to as principal components (PC) that account for most of the variation in the data. As opposed to the original variables, the new created variables (PC_1--PC_n) are uncorrelated. Because the PCA is a data reduction technique, it is usually the first index (or first principal component, PC_1) that is subject to a particular interpretation. In my case, the first PC explains 35% of the total variation and the second, 18 percent.

When I compute the correlation between the original (correlated) variables with the first two new non-correlated PCs I find that all the first indices have a high and positive correlation with *per capita* mobile and *per capita* email regardless of the variable that is dropped from the PCA and that interpretation associated with the index of ICT access should place emphasis on *per capita* mobile use and *per capita* email more so than the other access variables.

I also include all variables from the available dataset that are commonly theorised to have an effect on money-metric poverty. These include geographic location (urban/rural), maximum education in the household (HH), the average education of the HH, the gender of the head of the HH, the age of the HH head and the asset indices. Four models are estimated. These are:

1. Response is *Poverty Score for 2010* (y_t) (the level of per capita expenditure divided by the poverty line in 2010)

    1.1. *Unrestricted* model with <u>no</u> interactions vs. *Restricted* model with <u>no</u> interactions

---

[10]Unlike the previous analysis we do not convert the dependent variable into PPP$ and then into a poverty score as the conversions are potentially sensitive to the surge in prices that was experienced between CS1 and CS2. Since we are attempting to compare actual levels of expenditure and are concerned rather with change in PCE, this is acceptable.

1.2. *Unrestricted* model with interactions vs. *Restricted* model with interactions

2. Response is *the Log of the Differences in Per Capita Expenditure 2007-2010* ($\Delta y = y_t - y_{t-1}$) (the change in income between CS1 and CS2)

2.1. *Unrestricted* model with no interactions vs. *Restricted* model with no interactions

2.2. *Unrestricted* model with interactions vs. *Restricted* model with interactions

The unrestricted model refers to a model in which, in addition to the other factors, the following variables were included: 1) *Principal Component Scores* for every HH. This variable is lagged (γ_{t-1}) and thus constructed from the elements of year 2007 and 2) per capita expenditure in 2007/8 (y_{t-1}) which is introduced as a control variable since I already know that PCE 2007/8 is correlated with ICT access in 2007/8 (May et al 2010). Missing values are assumed to be missing at random (MAR) and no attempt to impute them was made. The missing values approximate 26% of the total number of observations. The results for the four unrestricted models are shown in Table 8.

Table 8: ICT and per capita expenditure

	Model 1 PCE 2010 No Interactions	**Model 2 PCE 2010 Interactions**	**Model 3 PCE Difference No Interactions**	**Model 4 PCE Difference Interactions**
R^2	0.4216	0.463	0.3506	0.4112
N	754	754	754	754
PCE2007 p test	0.0005	0.0001	0.0001	0.0001
ICT p test	0.0149	0.0021	0.1914	0.0377
ICT coefficient	11.006	13.663	0.0017	0.0367

In the restricted model, the variable 'Principal Component Scores' is omitted. All the other factors (linear and interactions) remain the same. Inclusion of 2-way interaction terms was based on their contribution to the overall fit (= 0.05). In general then, the model can be expressed as:

$$y_{2,2} = \alpha_2 + \gamma^{(1)} y_{2,1} + \beta_1^{(1)} x_{2,1} + \epsilon_{2,2} \qquad (5)$$

In this equation, the regressors (on the right hand side) were lagged value variable, x_{21}, of the independent variable and lagged value, $y_{2,1}$ of the dependent. The first index was for the cross-section identification, and the second index indicates the time period; $\gamma^{(1)}$is the autoregressive coefficient. The coefficient α_2 represents all the model fixed effects. This was a model in which the *change in PCE* is $y_{2,2}$ and its lag value was $y_{2,1}$. The fixed effects examine group differences in intercepts, assuming the same slopes and constant variance across groups.

Because I used a measure of "change", the underlying assumption is that the autoregressive coefficient is 1. Equation (6) can be written as:

$$y_{2,2} - y_{2,1} = \alpha_2 + \beta_1^{(1)} x_{2,1} + \epsilon_{2,2} \qquad (6)$$

The analyses indicated that the null hypothesis must be rejected in favor of the alternative when using both in-sample-F-test (p-value=0.013) and the asymptotic chi-square test (p-value=0.012). Model 1 indicates that the 2007 ICT-index contains new information that predict (log) difference the *PCE* (p-value=0.038) above and beyond the information contained in past values of the (log)% change in per capita expenditure (value<.0001).

The inclusion of the per capita expenditure in 2007 (a quasi-lag of the response), in the same model as the ICT index, usually makes it harder to find the significant relation between technology index and change in per capita expenditure. Results show that the connection is non negligible (slope= is 3.7% with a p-value of 0.038). Therefore the appropriate conclusion is that the relationship between the two truly exists. This indicates that I have enough evidence to conclude that the "ICT index" *statistically causes* change in the poverty score. The interpretation is that two individuals with 1 unit difference in their ICT indexes are expected to differ in log poverty score change by an amount equal to the corresponding slope. This implies that for one unit increase in ICT access, the expected difference in the logged PCE is 3.7 percent.

6.4 First Differences model

The starting point for the first differences analysis is identifying a matched sample to avoid possible selection bias. Two options for my choice of dependent variable are available for this analysis: change from having no access to new forms of ICT to having such access; and change from a lower level of ICT access to a higher level. In Table 9 I compare the results of two logistic regression models used in the first step of PSM to predict the probabilities of change on which matching will take place.

Table 9: Predicting change in ICT status[11]

Predictors	**Model 1 New ICTS only**			**Model 2 All ICTs**		
	B	**Wald**	**Exp(B)**	**B**	**Wald**	**Exp(B)**
Survey location (Urban)	**-.576**	5.642	.562	-.354	2.825	.702
Kenya_dummy	**-.676**	4.181	.509	-.126	.212	.882
Rwanda_dummy	-.007	.001	.993	.212	.919	1.236
Uganda_dummy	-.524	2.846	.592	.479	3.426	1.615
Log change in poverty score	**8.221**	8.609	3718.963	**5.203**	5.830	181.821
Sex of Head (Female)	.341	2.382	1.406	.275	2.195	1.316
Human capital in CS1	**-.540**	6.473	.583	-.031	.029	.969
Social capital in CS1	.060	.009	1.062	.910	2.701	2.484
Economic capital in CS1	**-1.714**	8.193	.180	-.604	1.558	.547
Physical capital in CS1	**-1.782**	21.481	.168	-.313	.941	.731
Constant	-8.962	6.802	.000	-6.887	6.704	.001
-2 Log likelihood	662.597			889.844		
Cox & Snell R Square	.149			.046		
Nagelkerke R Square	.214			.064		
Hosmer and Lemeshow test	.502			0.258		

[11] Coefficients that are significant at the 0.05 confidence interval are shown in bold.

Although my preference would be to match using all forms of ICT, comparison of the coefficients, pseudo r^2, and the Hosmer and Lemeshow tests confirm that the model better predicts changes in access to new forms of ICT only. While there are several possible explanations for this, it seems likely that access to radios, TVs and ICT appliances may be more closely linked to intra-household decision making processes and are perhaps less likely to change in the short period between the two waves of data collection.

Having chosen the most appropriate model with which to estimate probabilities of ICT change, propensity score matching is used to identify a sub-sample of households with similar characteristics. This results in 554 eligible cases, with 204 'treatment' households which had gained access to new forms of ICT between the waves of the survey. These are used to estimate the coefficient for gaining ICT (the treatment) on logged change in poverty status between CS1 and CS2 (the outcome) using an ordinary least squares regression (OLS). The result is shown in Table 10.

Table 10: Change in ICT status and change in poverty status

Predictors	B	sig
(Constant)	1.212	.000
survey location (Urban)	-.008	.317
country dummy	-.014	.139
country dummy	-.017	.060
country dummy	-.009	.399
Physical capital CS1	.005	.500
Economic capital CS1	.009	.276
Social capital CS1	.007	.518
Human capital CS1	.009	.264
Treatment (Gained ICT)	.025	.003

The results support the lagged model both in terms of the impact and its scale. Controlling for all other dimensions of poverty, the results show that gaining access to ICT is associated with a 2.5% improvement in poverty status between 2007 and 2010. Inter-acting the four capitals does not improve result and removing the country dummy controls results in the treatment effect becoming insignificant suggesting that country specific attributes are an important influence of this association.

6.5 Is ICT access pro-poor?

Having demonstrated likely causality between a reduction in ICT access and a reduction in money-metric, the next question is whether this impact is pro-poor and thus a potential resource with which households can eventually escape from poverty. In other words do those less financially resourced benefit from improved access to ICT more so than those that are better endowed? In this section, *per capita* expenditure (PCE) was split into two groups of 'Poor' and 'Very Poor' based on the percentile in the distribution of (untransformed) PCE. A series of analyses of covariance (ANCOVA) are performed to quantify and test the slopes' parallelism at each percentile. The full model contains most of the elements presented previously (see model 1), in addition to the factors "*Status*" (defining *Poor* and *Very Poor)*, the "*ICIndex2007"* (i.e., first principal component), and their interaction. At each percentile level, estimable functions are used to extract from the ANCOVA the parameters of intercepts and slopes from the interaction (see table below).

The hypotheses tests are:

1. Are the slopes equal to zero?
   $H_0 = \beta_{poor} = \beta_{very\ poor} = 0$ (7)
   H_1 not H_0

2. Are the slopes parallel?
   $H_0 = \beta_{poor} = \beta_{very\ poor} = \beta$ (8)
   H_1 not H_0

The results are presented in the following table along with the sample sizes and the significant tests. Only percentiles from 25 to 90% are shown for reason of space but the full range of points includes 5, 10, 25, 50, 75, 90, and 95 percent. These point estimates allow calculating the speed with which the groups evolve across percentiles.

Table 11: Threshold between the poor and the very poor

	25%			50%			75%			90%		
	Int	**Slope**	**N**	**Int**	**Slope**	**N**	**Int**	**Slope**	**N**	**Int**	**Slope**	**N**
Poor	2.03	0.0198	558	2.12	0.005	378	2.25	-0.017	197	2.38	-0.10	61
Very Poor	1.58	-0.0407	196	1.71	0.010	376	1.82	0.003	557	1.83	0.023	693

[a] α = 0.001, [b] α = 0.10, *Interc* is intercept, (N_{cell}) is the number of observations in each group

The results from the analysis indicate that the slopes are different from zero especially for the poor group at percentiles equal to 75% or less. However, the slopes diverge significantly at the extreme values, generally at significant level equal to 0.10. At low percentiles, the slopes are positive for the poor and negative for the very poor. After the 75 percentile, the very poor have positive slopes and the slopes of the poor become negative.

Moreover it can be seen that the slopes of the very poor keep rising until it reaches a significant level (=10 percent) whereas the slopes of the poor tend to decline as the level of the percentile poverty line increases. The speed of "recovery" for the very poor is positive (=0.14 percent) and is significant at =0.01. The slope of "decline" for the poor is negative (=-0.29) and is unchanged (=0.05 percent). All the intercepts are significantly different from zero and from each other (p< 0.001). These values represent the predicted 2010% change in the (log) per capita expenditure of individuals who don't have any form of ICT (i.e, total absence of ICT or the ICT index is zero)

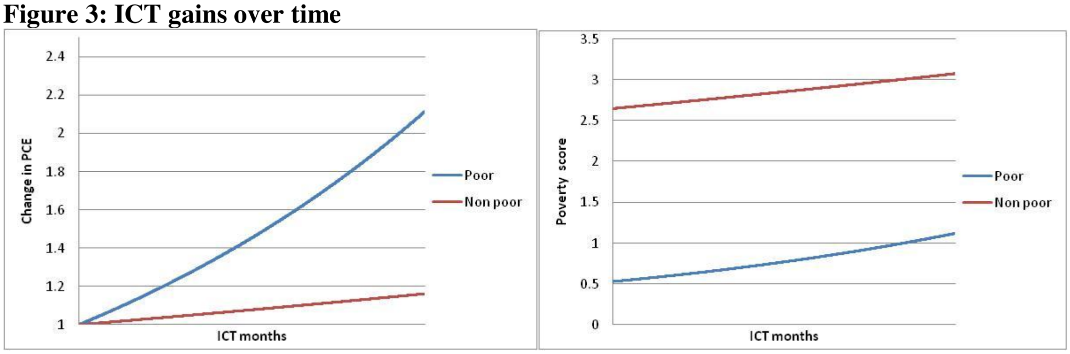
In order to assess the sensitivity of the analyses, I consider the following: first I investigated changes of the correlations $c_{11} * \sqrt{l_1}$ brought by the removal of one variable at a time in the PCA equation (5). The assumption is that if the relationships remain unchanged, then the highly correlated original variables are truly the ones to include in the ICT-index, and therefore the content (and interpretation) of the index should be confined to those variables. Regardless of the removed variable, I find both "Per Capita Mobile Use" and "Per Capita Email" have, for the most part, the most dominant correlation with the first PC. Loosely speaking, variations in the "output" (correlations) were not affected by variations in the "input" (removal of variables) of the first PC.

Second, the first PC is used in the model (1) to assess causality in the presence of PCE in 2007. One may then monitor whether changes in the input variables to the PCA equation create major inconsistencies in the subsequent tests of parallelism defined in (4). I further split the PCE 2007 into two groups (poor and very poor) under the name "status". The interaction between status and the first PC is the test of parallelism shown in (4). The splits are based on different percentiles, p, and result in different sample sizes. The percentiles are 5, 25, 50, 75, 90, and 95% of the data is allocated to one group and the other group has (1-p)% of the observations.

It is easy to lose sight of the importance and the implication of intercepts. What I see is that in all levels of the percentiles (25, 50, 75, and 90 percent), the intercepts of the very poor group are significantly lower than those of the poor. The finding together with the positive slope of the very poor over the percentile 50% implies that the ICT-index is pro-'very poor'. The technology index is for most part represented by the per-capita-mobile phone use variable.

This result can be depicted in Figure 3 in which the horizontal axis shows the number of ICT months (the number of 'technology units' available to a household over time) and the vertical axis depicts gains made to the poverty score of poor and non-poor households.

Figure 3: ICT gains over time



Over a ten year time horizon the modest additional gains from ICT can be seen to disproportionately benefit the very poor (those below median PCE) compared to the rest of the sample. This results in a slow but steady convergence between the two groups. Another way of interpreting this finding is that the gains resulting from ICT access for the most poor are twice that for the non-poor.

7 Conclusion

The information deficiencies concerning the impact of ICT on poverty reduction have raised concerns among policy makers who are being repeatedly urged to invest a substantial part of the national budget in ICT infrastructure on the basis of an incomplete evidence base. In this situation, it is tempting to question whether investments in ICT represent a worthwhile option for poor communities. A common mistake made by many existing studies is the collection of data at a level that is too general, thereby neglecting micro-level data that is required for the interpretation of macro-level trends. It is important to go beyond national level growth and development and analyse the role and impacts of ICT on poverty reduction at the micro-level, since changes in the well-being of individuals and households are not necessarily directly linked to changes in economic output at the national level.

The two waves of PICTURE Africa have attempted to address these concerns and are informative about inequalities in ICT access in Eastern Africa as well as the obstacles that hindered better and more equitable access. The odds of gaining ICT access was shown to improve by more than 100% relative to improvements in the poverty line and an additional year of education was shown to increase the odds of having ICT access by around 30 percent. Rural living, on the other hand, was found to significantly reduce the chances of having ICT access (by about 50 percent). Being female was also found to reduce the odds of ICT access by 50 percent. These relationships did not change significantly across the two waves of the study, suggesting some intransigence in terms of who benefits from ICT access.

Both waves of data therefore paint a picture of a heterogeneous ICT landscape. A multivariate analysis confirms the importance of formal education, but unsurprisingly suggests that there is also an interaction between education and income, and that this enhances ICT access. The determinants of individual access to ICT are largely similar in both waves, although there is a clearer gendered distribution of ICT access, with women 1.5 times less likely than men to have a mobile phone or email address, controlling for income and education.

The analysis of the casual link between ICT and poverty reduction indicates that there is a small but positive ICT benefit to the very poor group compared to the poor group. This is evident in the rates observed over the range of percentile levels (confirmed, once again, the 'poor nature' of the entire sample and more importantly the role of ICT in helping the very poor). The availability of mobile phones in particular is therefore a potentially valuable tool to improve the livelihood of the very poor over the medium term (6-10 years) and enable their escape from poverty traps produced by low stocks of assets and limited opportunities for their use. A panel study of more than two waves is needed to confirm the finding and would allow us to move beyond reliance on testing for statistical causality and instead make use of a more reliable difference-in-difference approach.

References

Alampay, EA., (2006). "Beyond access to ICTs: Measuring capabilities in the information society." *International Journal of Education and Development using Information and Communication Technology* 2(3): 4-22.

Ampah, M., Camos, D., Briceno-Garmendia, Minges, M., Shkratan, M and Willams, M., (2009). Information and Communication Technology in sub-Saharan Africa: A sector review, Africa Infrastructure Country Diagnostic Background Paper 10, World Bank, Washington DC.

Arunachalam, S., (2002). "Reaching the unreached: how can we use information and communication technologies to empower the rural poor in the developing world through enhanced access to relevant information?" *Journal of information Science* 28(6): 513-522.

Ashenfelter, O., 1978. "Estimating the Effect of Training Programs on Earnings", *The Review of Economics and Statistics*, 60 (1): 47-57.

Ashenfelter, O., and Card, D., 1985. "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs", *The Review of Economics and Statistics*, 67 (4): 648-660.

Athreya, V., (2004). "Dimensions of Poverty and Use of Technology to Achieve Poverty Reduction". GKP South Asia Regional Meeting.

Barrentes, R., (2005). "Analysis of ICT demand, in Regional Dialogue on the Information Society". *Digital Poverty: Latin American and Caribbean Perspectives*. H. Galperin and J. Mariscal. Ottawa, IDRC.

Baulch, R. and Hoddinott, J., (2000). "Economic mobility and poverty dynamics in developing countries: Special issue." *Journal of Development Studies* 36(6).

Bollou, F. and Ngwenyama, O., (2008). "Are ICT Investments Paying Off in Africa? An Analysis of

Total Factor Productivity in Six West African Countries from 1995 to 2002." *Information Technology for Development*, **14**(4): 294-307.

Carter, MR. and Barrett, C., (2006). "The economics of poverty traps and persistent poverty: An asset-based approach." *Journal of Development Studies* 42(2): 178-199.

Carter, MR. and May, JD., (2001). "One kind of freedom: Poverty dynamics in post-apartheid South Africa." *World Development* 29(12): 198-2006.

Coleman, JS., (1988). “Social capital in the creation of human capital.” *American Journal of Sociology* 44: S95-S120.

De Vos, S., (2005). "Indicating Socioeconomic Status among Elderly People in Developing Societies: An Example from Brazil." *Social Indicators Research* 73(1): 87-108.

Fiadzo, E., Houston, J., and Godwin DD., (2001). "Housing Quality for Poverty and Development Policy Analysis: CWIQ in Ghana." *Social Indicators Research* 53(2): 137-162.

Granger, CWJ., (1969). “Investigating causal relations by econometric models and cross-spectral methods”, *Econometrica*, 37 (3)L: 424-438.

Hood, MV., Kidd, Q, Morris, IL., (2008). “Two sides of the same coin? Employing Granger causality tests in a time series cross-section framework”. *Political Analysis*, 16, 324-344.

Hurlin, C and Venet, B., 2003. “Granger causality tests in panel data models with fixed coefficients”. Working paper, University of Orleans.

ITU (2011). *Measuring the information society: The ICT development index*. Geneva, International Telecommunication Union.

Jensen, R., (2007). "The Digital Provide: Information (Technology), market performance, and welfare in the South Indian fisheries sector." *The Quarterly Journal of Economics* 121(3): 879-924.

Kirkman, G., (1999). It’s More Than Just Being Connected. A Discussion of Some Issues of Information Technology and International Development. Development ECommerce Workshop. The Media Laboratory at the Massachusetts Institute of Technology. Cambridge, Massachusetts.

Liker, JK., Augustyniak, S. and Duncan, GJ., (1985). “Panel Data and Models of Change: A Comparison of First Difference and Conventional Two-Wave Models”, *Social Science Research* 14, 80-101 (1985)

Marker, P., McNamara, K. and Wallace, L., (2002). *The significance of information and communication technologies for reducing poverty*. London, DFID.

May, J., Dutton, V. and Mascarenhas, O., 2010. Poverty & Information and Communications Technology in Urban and Rural Eastern Africa (PICTURE-AFRICA): Results from cross-sectional data, School of Development Studies Research Report, University of KwaZulu-Natal, Durban.

Menard, S., (1991). *Longitudinal research*. Thousand Oaks, CA, US, Sage Publications, Inc.

Moser, CON., (1998). "The asset vulnerability framework: Reassessing urban poverty reduction strategies." *World Development* 26(1): 1-19.

Nulens, G. and van Audenhove, L., (1999). "The African information society: an analysis of the information and communication technology policy of the World Bank, ITU and ECA." *Communicatio,* 24(1/2): 28-41.

Ravallion, M., (1995). *Poverty Comparisons*. Chur, Harwood.

Reardon, T. and Vosti, SA., (1995). "Links between rural poverty and the environment in development countries: Asset categories and investment poverty." *World Development* 23(9): 1495-1506.

Röller, L.-H. and Waverman, L., (2001). "Telecommunications Infrastructure and Economic Development: A Simultaneous Approach." *The American Economic Review*, **91**(4): 909-923.

Rosenbaum, PR. and Rubin, DB., 1983."The central role of the propensity score in observational studies for causal effects", *Biometrika* 70(1): 41-55.

Singh, J., (1999). *Leapfrogging development? The political economy of telecommunications restructuring Albany*, State University of New York Press.

Sutherland, E., (2011). "Coltan, the Congo and Your Cell Phone." *SSRN eLibrary*.

Torero, M. and von Braun, J., (2006). *Information and Communication Technologies for Development and Poverty Reduction: The Potential of Telecommunications*. Washington, Johns Hopkins University Press and IFPRI.

Warschauer, M., (2004). *Technology and social inclusion: Rethinking the digital divide*. Massachusetts, The MIT Press.

World Bank, (2010). *Global Monitoring Report: The MDGs after the crisis*, World Bank, Washington DC.