UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

The Effects of Household Water Access on Employment in SSA:

Is There a Gender Difference?

Abstract

Several of the Millennium Development Goals stress the importance of working for improved drinking water, gender equality and employment to improve the life of people as well as creating growth in developing countries. Literature suggests investments in infrastructure increase the probability of employment. This study aims at investigating the relationship between having the major source of water for household use close and the likelihood of being employed. Furthermore, using a Probit model, it investigates whether the effect of water access is larger for women. The results display a positive correlation between water access and employment but cannot conclude a gender difference in the effect. The relationship between water and likelihood of employment provides an incentive for future investments in infrastructure in general and water infrastructure in particular in order to enhance labor force participation overall.

Bachelor Thesis in Development Economics (15 credits)

Department of Economics, Fall 12

Authors: Andrea Berggren

Anders Dankvardt

Supervisor: PhD Ann-Sofie Isaksson

Acknowledgments: We would like to express our gratitude towards our supervisor Ann-Sofie Isaksson. Thank you for guiding us and always being available for questions throughout the process of writing this thesis.

Andrea Berggren Anders Dankvardt

Table of Contents

1. Introduction

The World Development Report 2012 points out two important consequences of gender inequality. First and foremost, gender inequality deprives many women of living the life of their own choice and on the same terms as men. Secondly, gender inequality has a negative impact on economic efficiency as well as for other important development outcomes (WDR, 2012).

Gender gaps in education and employment reduce economic growth. There are several reasons for this, such as barriers that exclude women from labor force participation reduce the pool of talent and skill available for employers. High female unemployment increase fertility levels that in turn have a negative effect on growth. Furthermore international competiveness decrease since women often play a crucial role in export oriented industries. Women with an income of their own also increase their bargain power within the household and tend to save more and invest in education and their children to a larger extent than men do. When working in governance women also show fewer tendencies to engage in activities of corruption (Klasen and Lamanna, 2009).

Several of the Millennium Development Goals (MDG), established by United Nations in 2000, addresses issues interesting out of gender equality and employment perspectives. Two of these are goal 1 to eradicate extreme poverty and hunger and goal 3 to promote gender equality and empower women. One of the targets for goal 1 is "to achieve full and productive employment and decent work for all, including women and young people" (MDGR Sec 2:8). Only one out of three paying jobs outside the agriculture sector goes to women in Sub-Saharan Africa (SSA). These employments are often less secure and with lower wages. The financial crises in 2008 led to increased unemployment in SSA, especially for women. As a result parts of the labor force went from paying jobs to own-account or unpaid family work with lower income and no social security benefits. Women are overrepresented in the informal sector and overwhelmingly underrepresented in top-level jobs. An exception from this and where improvements have been made is within the political sector where women are slowly increasing their presence. Proportion of seats held by women in single or lower houses of national parliaments increased from 9% in 2000 to 18% in 2010. South Africa is thirdly best in the world after 44% of the seats in the lower-house parliament went to women. Only Rwanda and Sweden show higher female presence (MDGR, 2010).

There are several barriers that prevent women from taking part in the labor market in developing countries. In Pakistan, Ibraz and Fatima (1993) find that social structures obstruct women from engaging in market-oriented activities. Education has a positive effect on female labor supply according to Shaheen et al, (2011). In SSA girls however still lag compared to boys in school enrolment. In 2008 91 girls per 100 boys attended primary school. For secondary school the ratio decreased to 79 girls per 100 boys and for tertiary school only 67 girls per 100 boys attended school compared to all developing regions where the ratio is 97 girls per 100 boys (MDGR, 2010). Lack of property rights in the form of weak tenure security may lead to decreased labor supply since one household member (the one with lowest potential wage) is forced to stay at home to protect the house due to risk of expropriation by squatters (Field, 2007). Furthermore Cáceres-Delpiano (2012) argues that fertility affect mother's employment negatively. Research also draws upon AIDS related problems. Chicoine (2012) finds that AIDS related mortality decreased employment for both genders among those with no or little education in South Africa. Both Becker (1965) and Khandker (1987) find in their research that married women's labor supply is negatively correlated with husband´s wages and total household assets.

Traditionally, women in developing countries often get the role of household caregivers. Chores such as cooking, fetching water and collecting fuel etc. often fall on women. These are time-consuming activities that reduce the possibilities for education and/or income generating activities. Men in Africa only provide for one fifth of the water collection (UNICEF and WHO, 2008) while rural women spend on average one hour or more per day on gathering water (Koolwan and van de Walle, 2010). Hence women, to a greater extent than men, suffer from bad infrastructure in form of missing access to water or electricity (WDR, 2012). A study in South Africa shows that investments in electricity increased female employment by 9-9.5 per cent (Dinkelman,2011).

A large part of households in development countries lack direct access to water and are therefore forced to collect water from rivers, community taps or from other sources. Over one billion people worldwide need to travel more than one kilometer to collect improved drinking water and this responsibility often lies on the women of the household (Devoto et al, 2011). As a consequence, women's time allocation may be affected in a way that potentially decreases time

which can be spent on market activities or leisure. Previous research concerning the effect of improved water access on female employment has often focused on rather limited areas (e.g. Tangiers in northern Morocco) and has used limited samples. Furthermore these papers have mainly been conducted in development countries outside of SSA.

The aim in this study is to investigate if differences in access to water affect employment in SSA and also if there is a gender difference in this effect.

To address our aim we use a large and heterogeneous dataset from a survey covering 27,713 respondents across 20 Sub-Saharan African countries. The sample will display lots of variation and regional differences.

Our results show as expected that access to indoor- or in compound water is positively correlated with the probability of being employed in SSA. We cannot establish a significant gender difference in this effect.

The thesis is organized as follows: In section 2 we present previous theory and empirical evidence concerning the relationship between water access, employment and female time allocation. Section 3 describes the data and methodology used. In section 4 we present and discuss our results and in section 5 we will conclude our findings.

2. Water access and employment: theory and empirical evidence

In this section we present previous theory and research concerning time allocation, female employment and the effects of improved infrastructure (e.g. improved access to water) on employment. Against this background we then conclude by presenting our hypotheses.

Gronau (1977) have developed a home-production and time-allocation theory mainly as a tool for analyzing labor supply. This theory has its origin in ”A Theory of the Allocation of Time” by Becker (1965). Here time is spent on work in the market or leisure. Gronau argues that leisure as used by Becker is misguiding and stresses the importance of differentiate between leisure and home production since these respond differently to changes in the socioeconomic environment. Hence time is split into three parts; work in the market, home production, and leisure. This is

more appropriate for our study since much work in SSA is informal. Home production can be compared to similar work performed at the market and products and services produced at home are perfect substitutes for goods produced at the market. Household time should optimally be allocated so the marginal product of home production equals the marginal rate of substitution between consumption time and goods. This would in turn equal the shadow price of time. Time allocation between work at home, leisure and work at the market respond to changes in wage rates, commodity prices, productivity, education, marriage and child birth. Work at home and wage rates are expected to be negatively correlated and Gronau finds that wage rate together with household income are the two most significant determinants for married women´s labor supply in the short run. In the long run changes in education and changes in wage are correlated. Improvements in education might further increase labor supply if these are linked to decreasing fertility. The effect of a wage increase on leisure is less clear and dependent on the substitution- and income effects. Improved productivity in work at home is associated with increased leisure. The effects on work at home and work at the market are harder to establish. Having children is negatively correlated with female participation in the labor force, with a diminishing rate as the child gets older and the profitability of home production connected to the child decrease. Marriage and husband's wages (higher husband's wages relative to wives wages increase incentives for specialization) are both negatively correlated with female employment.

In the SSA context, where a large share of the population earns no cash income (66% in our sample, 73% of female respondents), female wages will probably remain low. Hence improved productivity in home production (e.g. improved water access) might have less effect on labor supply than in countries with large demand of labor. Instead time might be re-allocated to leisure as suggested by Devoto et al, (2011), more about this below. In this study we have chosen to focus on employment. The Afrobarometer data lacks data on time spent on leisure.

Several studies highlight the importance of infrastructure on employment. Investments in other infrastructure than water access might free time from household chores and hence increase the probability of labor force participation. Dinkelman (2011) investigates the effects of rural electrification on employment in South Africa. The article concludes that investments in public infrastructure in the form of electricity significantly increase female employment (9 – 9.5 percent). Electricity increases household productivity and create possibilities for home

production and hence possibilities for micro enterprises. Female wages do, however, decrease which suggests that firm labor demand does not grow. Electrified areas in South Africa have higher population growth than non-electrified areas. Dinkelman argues that this is due to migration from less attractive areas with weaker infrastructure and lower labor demand and stress the importance of taking this into consideration when interpreting employment results in areas where infrastructure investments have been made.

There is limited research concerning how water access affects households. Previous studies have often been conducted in other locations than SSA. Ilahi and Grimard (2000) carry out a study in rural Pakistan. Their focus is on infrastructure and the authors aim to find a relationship between access to water and female time spending patterns. Their findings show that improved access to water affects time use both at household and individual level. Where access to water is limited, due to poor infrastructure, women spend less time in market-oriented activities. This in turn has a negative impact on total household income. Furthermore, good water access leads to a decrease in the total work burden of women. Household investments in indoor water access are not only a way to increase total household income but will also ease the work burden of women. Based on their study, the authors argue that investments in water-supply infrastructure will decrease total work burden of women as well as reallocate female workforce from basic household chores to market oriented activities. Other studies have been unable to show any increase in the likelihood of being employed due to improved water access. Devoto et al, (2011) investigate the demand for household water connections and their effects on a range of household outcomes in the city of Tangiers in northern Morocco. They find that reducing time spent on collecting water via indoor connections has positive effects on time-use, mental welfare, and social activities while having no effect on household health. Their results, however, suggest that installing household water connections has no effect on productive activities. Neither male nor female employment increases. Time saved is spent on leisure only. Furthermore, Koolwal and van de Walle (2010) find no evidence supporting that improved access of water increases time spent on the market for women in rural environments in development countries. In contrast to studies mentioned above, this study includes several countries from SSA (Madagascar, Malawi, Rwanda and Uganda), North Africa, Middle East and South Asia. Problems with endogeneity concerning geographical differences are discussed since different regions show a great variance in female work force participation (ranging from 3 percent in Yemen to 40 percent in Madagascar). Most women in

the data set are working in home-production (e.g. farm work) which is not defined as participating in the labor market. Even if these women contribute to household cash income it is likely that the head of the household who, in most cases is a man, controls the income. Hence the effects of improved water access on home-production are ignored. Their results do not find a causal relationship between female work participation and improved water infrastructure. They do, however, find some positive outcomes such as better school enrollment for children.

Previous research shows some evidence regarding investments in infrastructure and increasing employment. In general, the situation for women overall seem to improve (i.e. better health, ease of work burden, more leisure). How time saved due to better water access is spent differ. Devoto et al, (2011) find that saved time is spent on leisure only while Ilahi and Grimard (2000) find that it is reallocated towards market oriented activities. These studies are very different in character and are not comparable. It is hard to establish a causal relationship between improved infrastructure and employment. Furthermore, if a relationship is found, it may be reverse indicating that a rise in employment increases spending and investments in infrastructure such as improved sources of water for household use. In the light of the above we hypothesize that:

1. Reallocation of time, due to having the household's water source inside the house or compound, will have a positive impact on the likelihood of employment.
2. Impact of water access on employment is greater for women.

3. Methodology

In this section the basis for the different regressions will be presented. We will describe variables of interest and discuss strengths and weaknesses with the construction of our variables based upon our sample.

The aim of the regressions is to investigate whether source of household water is significantly related to employment. Furthermore we will investigate whether there are gender differences in this correlation. We will explicitly focus on women as we argue that the effect of having water inside the house or compound will be larger for women, as stated in hypothesis two.

3.1 Data

Our data is from The Afrobarometer which conducts surveys regarding political, social and economic issues in Africa. We use the latest round of the survey which is conducted in 2008 covering 20 African countries on an individual level (see Appendix 1 for specification of countries). The sample is very large, including 27 713 observations collected in 2008 (Afrobarometer Network, 2007). The large sample will constitute a solid base for our multi-country regression.

Table 1 Sample description

	Mean	**Std. Deviation**	**Min**	**Max**
Age	36	14.5	18	100

	Number of Observations	**Share of sample %**
Female	13 837	49.9
Male	13 876	50.1
Rural	17 192	62.0
Employed	9 358	33.9
Education		
No schooling	4365	15.8
Primary education	10 251	37.1
Secondary education	10 115	36.6
College	1 674	6.1
Higher education	1247	4.5
Water source:		
Inside the house	4 602	16.7
Inside the compound	4 712	17.1
Outside the compound	18 249	66.2
Sample	27 713	100.0

Table 1 shows the key explanatory variables along with our socio-demographic variables. It shows that the majority of the respondents in the sample are without cash income (66%). 34% of

the respondents have their source of water inside the house or inside compound, i.e. "nearwater". A majority of households in the sample fetch water outside of the compound. The sample is divided so that there is roughly the same number of male and female respondents. The majority lives in rural areas (62%).

3.2 Model

As we have a binary response model we are using the Probit model (probabilities $\in [0, 1]$). Using the Probit model will allow us to analyze the marginal effects rather than coefficients. This means the interpretation will be the maximum likelihood of success. In this context, success will be equal to employment (Woolridge, 2008). The following will be our true models on which we will run the regressions.

$$\boldsymbol{Prob(employment = 1|x) = \Phi(\beta_1 female + \beta_2 nearwater_i + \beta_3 G_i + \beta_4 I_i + \beta_5 R_i)}$$

The probability of being employed depends on gender and water access i.e. on the variables female and nearwater. The key explanatory parameters will accordingly be $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$, conditional on socio-demographic variables grouped in $\boldsymbol{G_i}$, infrastructural variables represented as $\boldsymbol{I_i}$ and regional effects, $\boldsymbol{R_i}$, where $\boldsymbol{\Phi}$ is the cumulative density function for standard normal distribution.

We will run five different regressions which will follow our hypotheses.

In the first regression, the explanatory variable will only be source of water for household use; hence the model will be as stated above. For regression 2 and 3 we will use the same model but rather run separate regressions for women and men:

$$\boldsymbol{Prob(employment = 1|x) = \Phi(\beta_1 nearwater_i + \beta_2 G_i + \beta_3 I_i + \beta_4 R_i)}$$

When running regression 4 and 5 we are going to include interaction terms. These are multiplicative terms and are used to analyze the effect of one variable which is dependent on another. In this setting, including interaction terms will enable us to analyze the gender difference in the effect of having water on employment (Woolridge, 2008). For regression 4 we include one interaction term for having water inside the house or compound (*female) which will allow us to see if there is a gender difference in the returns of having in-house or inside compound water on the likelihood of employment. The model for regression 4 will then be:

$$\boldsymbol{Prob(employment = 1|x) = \Phi(\beta_1 female + \beta_2 nearwater_i + \beta_3 nearwater * female + \beta_4 G_i + \beta_5 I_i + \beta_6 R_i)}$$

In regression 5 this effect will be broken down into either having water inside the house or inside the compound meaning we are using two interaction terms:

$$\boldsymbol{Prob(employment = 1|x) = \Phi(\beta_1 female + \beta_2 inhouse_i + \beta_3 incompound + \beta_4 inhouse * female + \beta_5 incompound * female + \beta_6 G_i + \beta_7 I_i + \beta_8 R_i)}$$

where the other (control) variables have the same interpretations.

Table 2 Expected Signs

	Full sample(1)	Female sample (2)	Male sample (3)	Female Interaction term (4)	Two female interaction terms (5)
Female	-			-	-
Near water	+	+	+	+	
In-house water					+
Inside compound					+
Near distance_female				+	
In-house water_female					+
In compound_female					+

We expect the sign on female to be negative as empirics show that being female has a negative effect on employment. As stated in the hypothesis we expect the effects of having water inside or near the house to be positive on the likelihood of employment, i.e. the sign of the marginal effects on the water variables to be positive. For our second hypothesis to hold we furthermore argue that the sign on the interaction terms should be positive to, indicating a greater effect for women.

The validity of our regressions is linked to the construction of the model and the quality of the data set and can be divided into internal and external where the former concerns the credibility of the results in this given situation and the accuracy of our parameters. The latter refers to the ability to generalize to other situations based upon our results (Heckman, 2008).The internal validity should be fairly good as the large data set allows us to control for a range of different factors which should make endogeneity problems less apparent. Since our aim is to investigate if access to water access affects the probability of employment, we would like to establish a causal relationship between these two variables. This is hard since there may be problems with reverse causality meaning we are displaying a situation where employed people have higher income and can afford to invest and/or live in areas where water infrastructure is good. Furthermore there are still some issues with the construction of our variables and these will be discussed below. The external validity is potentially high since we have a very large, heterogeneous data set capturing lots of variation between regions and countries in SSA. Therefore, the results can be applicable in different settings although with similar conditions to those in SSA. We cannot claim these results to be valid in all situations and especially not for developed countries where water infrastructure and employment status may be completely different.

3.3 Variables

3.3.1 Employment variable

Our dependent variable is employment. It is a binary variable taking on values 0 or 1 where 0 equals no employment and 1 equals employment. The survey question was based upon having a cash income and the answers were categorized as having no employment, part-time or fulltime employment (see Appendix 2 for all survey questions). We have constructed the variable where employment equals all categories except no job. Accordingly, the interpretation will be the likelihood of having a job (where job is defined a cash income). This definition of employment implies that people reporting no cash income will be excluded from the employed group. This can in turn be people working on the household farm i.e. are subsistence farmers. This will constitute a weakness to our model. By construction, the variable fails to capture the large share of population working on household farms (unless they earn cash income). As we described in the theory section this is especially true for women. Empirics have shown that women are not as

represented in the formal sector (where cash income is present) as men (ILO, 2012). This does not mean, however, that these women spend all their time on leisure. A wider definition of employment would perhaps catch this share too, but as mentioned previously, our data will limit us to wage employment. Studies have shown that women in SSA have more responsibility for household chores than men which will decrease their time available for formal job (Dinkelmann, 2011). Our construction of the variable will still capture important information regarding having or not having cash income. It may be that, for being defined as unemployed you have to actively look for work which many poor people cannot afford. They rather work informally earning cash income from that sector. If this is the case they would be defined as employed according to our definition of employment. This could potentially capture a larger share of the working population than more narrow definitions of employment (i.e. working in the formal sector).

3.3.2 Gender and water access variables

One of our key explanatory variables is gender, in the sample, 40% of all men are employed and 27 % of all women. This result implies that men are more likely to be employed than women. There is a correlation between gender and employment and therefore we need to control for gender differences in the first regression by including a dummy variable. Furthermore, gender will be a variable of interest in the other regressions where focus is on gender differences within the effects of having water. The base group will be men so the marginal effect will be interpreted as the effect of being female on employment compared to being male. Another explanatory variable is source of water for household use. This variable will be, as stated in the hypothesis, one of the variables of main interest. The survey question gives three alternatives for water source: inside the house, inside the compound or outside the compound. The variable "nearwater" is going to be a dummy variable equal to one if the respondent has water inside the house or inside the compound and zero if the water source is outside of the compound. In regression 5 we will use two dummies where one dummy is for in-house water and one for inside the compound. Outside compound will be left out of all the regressions i.e. our base group which we can compare our estimated effects to.

We have also constructed an interaction term that will capture gender differences in the relationship between water access and the likelihood of being employed conditional on our control variables. The interaction term is the product of the female dummy and the dummy for

water, “nearwater”. In regression 5, there will be two interaction terms: one for female*inhouse water and one for female*insidecompound.

3.3.3 Socio-demographic controls

We include control variables to reduce the chance of biased estimators on our explanatory variables. Optimally, we control for everything that affects the probability of being employed in this setting that is also correlated with our key explanatory variables. All factors that are correlated to employment should be included in our model but in reality all of these are not observable. Based upon our data sample we have on decided three major groups of control variables, namely socio-demographic, infrastructural and regional controls. Controlling for these will decrease the probability of biased estimators on our explanatory variables as we believe these also affect the likelihood of female employment.

All respondents in the data set are over 18 years old. We include an age and an age squared variable to capture the non-linear effect of age on being employed. Non-linear effect means in this setting that age will have a positive impact on the likelihood of being employed but at a diminishing rate. This implies that the older you are, the more likely you are to be employed. The effect however gets smaller and smaller as you approach a certain age and there will be a turning point when in fact the effect gets negative (for example when retiring).

There is a negative (although relatively weak) correlation between having in-house water and receiving no schooling (Shaheen et al, 2011). Education will have a strong impact on the probability of being employed. The correlation between education and employment is positive (as shown in the results) and to minimize the risk of biased estimators we will include level of education in our socio-demographic group of control variables. There are five different dummy variables for education: No school, Primary school, Secondary school, College and Higher education. No school will be the base group and hence left out of the regression.

Rural population, compared to urban, is to a larger extent not employed in the data set. Furthermore, the infrastructure of water is expected to differ between rural and urban areas. Hence we include a dummy taking the value one if you live in a rural area and zero if living in an urban area. The definition of the employment variable is having a cash income and having a cash income might be correlated with living in an urban area. Consequently, living in a rural area or

not will affect likelihood of employment. According to the dataset almost 30% of the respondents living in rural areas have reported cash income. The corresponding number for respondents living in urban areas is 40%.

Due to lack of information in the dataset we cannot construct variables for number of children and years of experience. We do, however, believe that these will have an effect on employment

status and are therefore expected to be part of the error term, u. This may cause a potential bias. Fertility is negatively correlated with female employment (Cáceres-Delpiano, 2012). Empirics have shown that experience, however, is expected to have a positive impact on employment although at a diminishing rate (Woolridge, 2008). Both these variables can, however, be captured to some extent by age since both fertility and experience have a strong correlation with age.

3.3.4 Infrastructural controls

Having water or not is dependent on and/or correlated with other types of infrastructure. For example, as mentioned above, the majority of the sample living in the city has water inside the house or compound (60%) whereas only a fraction of the rural respondents have water access (18%). Even though we do control for rural/urban differences, there may still be some differences that are captured by direct infrastructural controls. By including a group of infrastructural variables other than water we aim to capture the benefits of living in areas where existing infrastructure is better and conversely negative effects of poor infrastructure. We also reduce the likelihood that the variable for water captures effects of other correlated types of infrastructure. Other types of infrastructure than water may also affect the probability of employment. According to Dinkelman (2011), electricity is positively correlated with employment. Hence, we will use a dummy variable for electricity grid representing access to electricity infrastructure to minimize the probability of biased estimators. A dummy variable equal to one if market stalls are present in or nearby the sampling unit (zero otherwise). If markets stalls are present the demand for labor and goods should increase indicating a greater possibility of being employed. Market stalls could also be an indicator of an area with adequate infrastructure.

3.3.5 Regional effects

Local differences in regions affect employment status. Including one dummy per region (in total 245 dummies) will allow us to capture regional differences other than rural/urban and the infrastructure controls our data lets us control for. There may for example be differences in labor demand, geographical factors, presence of conflicts in different regions. Our dependent variable, employment (defined as cash income) will most likely also differ amongst regions as the probability of having cash income will be higher in some regions and lower in others.

4. Results

In this section we will present the results from our regressions. We have carried out five different regressions which all support our first hypothesis that reallocation of time, due to having the household's water source inside the house or compound, will have a positive impact on the likelihood of employment. We have had a harder time finding evidence for the second hypothesis that the impact of water access on employment is greater for women which we will discuss more in depth when discussing and concluding our findings.

4.1 Main Results

Table 3: Employment in SSA (*probit marginal effects*)

Dependent variable: **Employment**	Full sample(1)	Female sample(2)	Male sample (3)	Female Interaction term (4)	Two female interaction terms (5)
Female	-0.128***			-0.128***	-0.128***
	(0.006)			(0.010)	(0.008)
Water source					
In near distance	0.039***	0.041***	0.035***	0.034***	
	(0.008)	(0.011)	(0.012)	(0.010)	
In-house water					0.030**
					(0.014)
Inside compound					0.050***
					(0.012)
Interaction terms					
Near distance_female				-0.0005	
				(0.0126)	
In-house water_female					0.016
					(0.016)
In compound_female					-0.016
					(0.015)
Socio-demographic controls					
Rural	-0.022**	-0.014	-0.030**	-0.022**	-0.022**
	(0.009)	(0.012)	(0.013)	(0.009)	(0.009)
Age	0.042***	0.035***	0.047***	0.042***	0.042***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Age^2	-0.0005***	-0.0004***	-0.0005***	-0.0005***	-0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Primary education	0.055***	0.057***	0.051***	0.055***	0.055***
	(0.011)	(0.014)	(0.017)	(0.011)	(0.008)
Secondary education	0.175***	0.200***	0.150***	0.175***	0.175***
	(0.012)	(0.017)	(0.020)	(0.012)	(0.012)
College education	0.413***	0.493***	0.350***	0.413***	0.413***
	(0.016)	(0.024)	(0.023)	(0.016)	(0.017)
Higher education	0.330***	0.393***	0.282***	0.330***	0.331***
	(0.019)	(0.031)	(0.025)	(0.019)	(0.019)
Infrastructural controls					
Electricity	0.028**	0.024**	0.030**	0.028**	0.028**
	(0.009)	(0.012)	(0.014)	(0.009)	(0.009)
Market Stalls	-0.014*	-0.009	-0.017	-0.014*	-0.014*
	(0.008)	(0.010)	(0.012)	(0.008)	(0.009)
Region dummies	yes	yes	yes	yes	yes
Observations	26 782	13 182	13 434	26 782	26 782

Notes: Standard errors in parentheses; *significant at 10%, **significant at 5%, ***significant at 1%.

1. Full sample: Only dummy for having water inside the house or compound, the water parameter is restricted to be the same for both men and women
2. Female sample: Regression on females only, all variables (if female=1)
3. Male sample: Regression on men only, all variables (if female=0)
4. Female interaction term: Dummy for having water inside the house or compound and interaction term = nearwater*female
5. Two female interaction terms: One dummy for water inside house, one for inside compound and 2 interaction dummies with respect to the two water variables

The results in regression 1 (not taking into account explicit gender differences in the water parameter) shows the marginal effects of having water in the near distance. It shows that having access to water is associated with a 3.9 percentage points higher likelihood of being employed. This is statistically significant at the one percent level. Worth noting also is that being female generates an almost 13 percentage points (significant at one percent level) lower likelihood of being employed compared to being male.

Regression 2 is for females in the sample only. The results suggest that the effect of having water, for women, is positive and statistically significant at the one percent level and is associated with an increase in the likelihood of being employed with 4.1 percentage points. In contrast, regression 3 is for male respondents only. The results for this regression suggest a weaker correlation between having water and employment which is in line with our second hypothesis. More precisely having water is associated with higher probability of employment of 3.5 percentage points (significant at one percent level).

For regression 4 and 5 we have included interaction terms to account for gender differences in the water access on employment. Starting with estimation 4, the results suggest that having water nearby is linked to a higher likelihood of being employed of 3.4 percentage points (significant at the one percent level). The interaction term, however, is neither statistically nor economically significant.

In regression 5 the water variable is broken down into two: one for having the water source inside the house and one for having the source of water inside the compound (both are compared to

having the water source outside the compound). Correspondingly, there are two included interaction-terms. The results suggest that having its major source of household water inside the house is associated with an increase in the probability of being employed with 3 percentage points (significant at the five percent level). Furthermore, having the water source inside the compound is associated with a 5 percentage points higher likelihood of employment and this marginal effect is statistically significant at the one percent level. When it comes to the interaction terms the results are a little less clear cut. The marginal effect of in-house water suggests that the likelihood of employment is 1.6 percentage points larger for women (i.e. 0.03+0.016=0.046). Having water inside the house is linked to a 4.6 percentage points higher probability of being employed in total when being female. Even though the total marginal effects is greater for women, the parameter on the interaction term, β_3, is not statistically significant. Thus, we cannot conclude a gender difference in this effect. The same goes for the marginal effect of the interaction term for having water inside the compound. The result here suggests that the probability of employment when being a woman is related to a 1.6 percentage points decrease in the probability. The relationship is still positive (0.05-0.016=4.4), so the total effect is a 4.4 percentage points higher likelihood of female employment which is slightly lower than for men. The interaction term is not statistically significant so even if there is a positive effect of having water, once again, a gender difference cannot be established.

Overall the sign on having water was expected to be positive which these results support. For the interaction terms, the signs differ but the marginal effects are statistically insignificant. The signs on our control variables are as expected. Being female and living in a rural area has a negative relationship with the likelihood of employment. Age and education affects the probability of employment positively.

4.2 Discussion

The result supports the first of our hypotheses, that reallocation of time, due to having the household's water source inside the house or compound, will have a positive impact on the likelihood of employment. The second part however, that this impact on employment is larger for women cannot be confirmed based upon these results. There is no statistically significant gender difference in the effects of having water on employment.

As literature shows, fetching water is a time-consuming task. Having access to water inside the house or compound will release time that can be spent generating income to the household. Therefore the estimated marginal effect on employment of having water nearby compared to having water outside the compound is positive and linked to an increase in the probability of employment with 3.4-5 percentage points. It is statistically significant at the one percent level. Worth noting is that even though the marginal effect is positive the effect seems to be larger for women then for men overall. The difference between regression 2 and 3 suggests that the effect of having close access to water overall is larger for women than for men. For women, water access is estimated to increase in probability of being employed is 4.1 percentage points. The equivalent figure for men is 3.4 percentage points. The marginal effect of the female dummy suggests that women are 13 percentage points less likely to be employed compared to men. It is statistically significant at the one percent level. Overall, we find that an indoor or inside compound household source of water has a positive and significant impact on employment, as stated in our first hypothesis.

For the second part of our hypothesis, concluding a larger effect for women is more difficult. The focus will be on the interaction terms as these tell us if there is a gender difference in the effect of having water on employment, conditional on all other factors being held constant (Woolridge, 2008). The interaction term in regression 4 is the product of having the water source nearby and being female. It is neither economically nor statistically significant. Accordingly, we cannot conclude a gender difference based upon this outcome.

For regression 5 the water source is divided between having water inside the house or inside the compound. Both marginal effects of the water variables are positive. This indicates a higher probability of being employed when having *one of these* water sources in comparison to having its water source outside of the compound. When taking into account the interaction terms the effects becomes slightly different. The total effect of having water inside the house is larger for women than for men. Being female and having in-house water increase the probability of being employed with 4.6 percentage points in total. In contrast, the total impact of having the water source inside the compound is slightly smaller for women compared to men. The total effect is that having this water source predicts a 4.4 percentage points increase in the likelihood of employment for women. The same effect is predicted to increase the likelihood of being

employed with 5 percentage points for men. However, the difference between men and women in these marginal effects is not statistically significant at any conventional level. Hence we cannot conclude a gender difference in the effect of having better water access on the probability of employment based upon these results. If a large share of women in the sample is self-contained it may be that the time saved from having water is rather spent on increased work on the farm. This would not make them employed according to our definition of employment (unless the eased activity in farming resulted in a cash income).

To conclude that source of water drives female employment is a very hard task.

The different regressions estimate a positive effect of water access when other variables are controlled for. We have included a range of different variables of various types. There are socio-demographic, infrastructural and regional controls and our aim is that these capture most of the variation in employment that is not explained by water access. As we mentioned earlier we lack data on fertility, experience and health status. These factors will most likely affect the probability of having a cash income so to omit these from our model is a limitation. As noted in the methodology section we could display a case with reverse causality meaning that when a household earn cash income, that particular household might be more likely to afford a house with better water access. The cost of these houses or of installing water might be too high if the household lacks income. The interaction terms however might not be as effected by reverse causality since they show the combined effect of gender and access to water (although these are not statically significant). If reverse causality is present it would decrease the internal validity of our regressions. We cannot exclude this as an alternative explanation for the relationship between the dependent and independent variable. In order to be able to conclude causality we would need to conduct experiments, randomly choosing households and only improve their access to water. This would give us exogenous variation in water only and we would not have to worry about problems with endogeneity and more easily conclude causality. This type of data set would be much smaller than our existing (which contains almost 28,000 observations) and would also be a drawback to the validity of our results when aiming to capture an area as large as SSA. Another way of solving the causality problem could be to use an instrumental variable approach. The analysis builds upon finding a relevant instrument that is correlated with our explanatory (causal) variable (access to water) and uncorrelated with the residual (exclusion restriction)

(Wooldridge, 2008). The first assumption is testable but the second assumption is harder to justify. In our case we would for example think that health status is included in the residual causing biased estimators. We could then find an instrument that is correlated with access to water but not with health status. There are problems with this approach to however, as mentioned it is hard to justify the exclusion restriction that requires the correlation between the instrument and the residual to be zero since the latter is unobserved (Wooldridge, 2008).

Furthermore, there are most likely some variables that are omitted, i.e. an omitted variable bias. These may be variables that are correlated with both access to water and employment. This could for example be ability, if you have high ability you are more likely to have a cash income and live in an area where the (water) infrastructure is good. We do however control for some infrastructural differences and regions so the probability of having this type of bias decreases. The problems mentioned above will decrease the internal validity of our research. We do however argue that there is some validity to our results because we still control for lots of different factors. This relieves some of the burden of endogeneity although there are still some issues. As mentioned in the methodology section the external validity is higher because of the large sample. In contrast to previous research in the area we have a very large number of respondents and a heterogeneous data set which allows us to generalize more from our results. This is especially true for developing countries similar to those included countries that are included in the Afrobarometer data set (see appendix 1 for a specification of included countries).

Our findings show similarity to those made in previous research and theory. Improved access to water free time from home production which could be used either for leisure or market activities. Whether or not this gained time increase employment seems to vary between different settings. Devoto et al. (2011) was not able to find that investments in private water taps had any positive effects on labour market participation. Our results on the other hand shows that household source of water do affect the probability of being employed. Dinkelman (2011) found a rather large increase in employment in areas where electricity was made available and also that there was a significant and positive gender effect for women. Ilahi and Grimard (2000) found that poor infrastructure i.e. bad water access decreases women engagement in market-oriented work. We were unable to find statistically significant evidence that supports any gender differences. As mentioned before previous studies have been made in narrow, well-defined areas and sample

sizes have accordingly been much smaller and less heterogeneous. Koolwal and van de Walle (2010) do not conclude a causal relationship between water access and female work participation based upon their results and because of endogeneity problems. They include a large set of developing countries in their study but their focus is on rural women. We include both rural and urban households and investigate likelihood of cash income based upon a larger sample size from SSA only. Our data is collected in 2008 when the financial crises began and led to increased unemployment in SSA. A weaker labor demand might underestimate the effect of water access on employment.

5. Conclusion

Inadequate access to water is of interest since it affects a large proportion of people living in developing countries. Furthermore, employment increases economic productivity and will eventually lead to higher growth. The aim of this study has been to investigate if differences in access to water affect employment in SSA and if there is a gender difference in this effect. Accordingly we did formulate hypotheses that related to this aim; in particular we believed that reallocation of time due to greater access to water increase the probability of employment. We also hypothesized a greater effect for women. Our results show a positive correlation between access to water and the likelihood of being employed which supports our first hypothesis. For the second hypothesis however there seems to be a gender difference, i.e. the correlation between women and employment is stronger but we cannot conclude this with statistical significance. This may be due to causality problems such as reverse causality and omitted variables. The possibility of an omitted variable bias is reduced by including control variables of different types. We cannot rule out the possibility of reverse causality. As a consequence we cannot establish a causal relationship between water access and employment. In order to do this we would have to do experiments with randomly chosen household in order to achieve exogenous variation in water only. The sample size of this type of data would be much smaller than the data set used in this study. This would in turn decrease the possibility to generalize from the result (i.e. the external validity), even if it may result in a causal relationship. The existing research regarding the relationship between water and employment is limited. Further research is needed,

especially in SSA, where the relationship between water and employment is relatively unexplored but where problems regarding access to water are most urgent.

Considering challenges ahead with global water shortage and poverty and lack of growth in SSA, this type of research is important. Investing in better water infrastructure would not only improve economic activity but also have a number of other positive effects such as improved health and eased work-burden. Public policies in developing countries should focus on investments in infrastructure to create employment opportunities. Better access to water, electricity and other infrastructure releases time for the household that could be spent in market-oriented activities increasing economic productivity both on micro and macro level. These types of policies are of extra importance from a gender perspective since women to a higher extent than men lack income of their own. Greater gender equality drives economic growth and reduces poverty.

References

Afrobarometer Network (2007). Round 4 Survey Manual, Afrobarometer.

Becker , G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, vol. 75, pp. 493-517.

Caceres-Delpiano, J. (2012). Can We Still Learn Something from the Relationship between Fertility and Mother´s Employment? Evidence from Developing Countries *Demography*, vol. 49, pp. 151-174.

Chicoine, L. (2012). AIDS mortality and its effect on the labor market: Evidence from South Africa. *Journal of Development Economics,* vol. 98, pp. 256-269.

Devoto, F., Chicoine, L., Duflo, E., Dupas, P., Pariente, W. & Pons, V. (2011). Happiness on Tap: Piped water adoption in urban Morocco. *Working Paper 16933*. Cambridge MA: National Bureau of Economic Research.

Dinkelman, T. (2011). The Effects of Rural Electrification on Employment: New Evidence from South Africa. *The American Economic Review*, vol. 101, pp. 3078-3108.

Field, E. (2007). "Entitled to Work: Urban Property Rights and Labor Supply in Peru. *Quarterly Journal of Economics*, vol. 122, pp. 1561-1602.

Gronau, R. (1977). Leisure, Home Production, and Work-The Theory of the Allocation of Time Revisited. *Journal of Political Economy,* vol. 85, pp. 1099-1123.

Heckmann, J. (2008). Econometric Causality*, International Statistical Review*, vol 76, pp. 1-27.

Ibraz, T. S. & Fatima, A. (1993). Uneducated and Unhealthy: The Plight of Women in Pakistan. *The Pakistan Development Review,*Vol 32, pp. 905-915

Ilahi, N. & Grimard, F. (2000). Public Infrastructure and Private Costs: Water Supply and Time Allocation of Women in Rural Pakistan. *Economic Development and Cultural Change*, vol. 49, pp. 45-75.

ILO (2012). *Global Employment Trends for Women*, International Labor Office 2012, Geneva: International Labor Organization

Khandker, S. R. (1987). Labor Market Participation of Married Women in Bangladesh. *The review of Economics and Statistics*, Vol. 69, pp. 536-541

Klasen, S. & Lamanna, F. (2009). The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries. *Feminist Economics,* vol. 15*,* pp. 91-132.

Koolwal, G. & van de Walle, D. (2010). Access to Water, Women´s Work and Child Outcomes. *World Bank Policy Research Working Paper Series*, Vol. , pp.

Millennium Development Goals Report 2010 (2010). New York. United Nations.

Shaheen, S., Sial, M. & Awan, M, S. (2011). Female Labor Force Participation in Pakistan: A Case of Punjab. *Journal of Social and Development Sciences,* vol. 2, pp. 104-110.

UNICEF & WHO (2008). *A Snapshot of Drinking Water and Sanitation in Africa*, UNICEF and World Health Organization 2008, Geneva

Wooldridge, J. (2008) *Introductory Econometrics: A Modern Approach with Economic Application, Data Sets, Students Solutions Manual Printed Access Card.* (4th ed)., South-Western College Pub.

WDR (2012). *Gender Equality and Development*. World Development Report 2012. Washington D.C:The International Bank for Reconstruction and Development / The World Bank,

7. Appendix 1

COUNTRY	Freq.	Percent	Cum.
Benin	1,200	4.33	4.33
Botswana	1,200	4.33	8.66
Burkina Faso	1,200	4.33	12.99
Cape Verde	1,264	4.56	17.55
Ghana	1,200	4.33	21.88
Kenya	1,104	3.98	25.87
Lesotho	1,200	4.33	30.20
Liberia	1,200	4.33	34.53
Madagascar	1,350	4.87	39.40
Malawi	1,200	4.33	43.73
Mali	1,232	4.45	48.17
Mozambique	1,200	4.33	52.50
Namibia	1,200	4.33	56.83
Nigeria	2,324	8.39	65.22
Senegal	1,200	4.33	69.55
South Africa	2,400	8.66	78.21
Tanzania	1,208	4.36	82.57
Uganda	2,431	8.77	91.34
Zambia	1,200	4.33	95.67
Zimbabwe	1,200	4.33	100.00
Total	27,713	100.00	

.

8. Appendix 2

These survey questions come from Round 4 of the Afrobarometer Network (2007).

Employment variable

Question Number: Q94
Question: Do you have a job that pays a cash income? Is it full-time or part-time
Code: 0=No, 1=Yes, part time or full time

Gender and water access variables

Question Number: Q101
Question: Respondent's gender
Code: 0=Male, 1=Female

Question Number: Q93A
Question: Where is your main source of water for household use located?
Variable Label: Source of water for household use
Code: 0=Outside the compound, 1=Inside the house or inside the compound (For regression 5 we have two dummies for water access; one for inside house and one for inside compound)

Socio-demographic controls

Question Number: URBRUR
Variable Label: Urban or Rural Primary Sampling Unit
Code: 0=Rural, 1=Urban

Question Number: Q89
Question: What is the highest level of education you have completed?
Variable Label: Education of respondent
Value Labels: We included 5 separate dummies for education. The respondent answers the highest level of education he or she has received. The alternatives are: No formal schooling, Primary School, Secondary School, College, Higher Education. No formal schooling is the base group (i.e. =0)

Question Number: Q1
Question: How old are you?
Variable Label: Age
Value Lables: 18-110

Infrastructural controls

Question Number: EA_FAC_E
Question: Are the following facilities present in the primary sampling unit/enumeration area, or within easy walking distance: Market stalls (selling groceries and/or clothing)?
Value Labels: 0=No, 1=Yes

Question Number: EA_SVC_A
Question: Are the following services present in the primary sampling unit/enumeration area: Electricity grid that most houses could access?
Value Labels: 0=No, 1=Yes

Regional effects

Question Number: REGION
Question: Region/Province
Value Labels: 100=Alibori, 101=Atacora, 102=Atlantique, 103=Borgou, 104=Collines, 105=Couffo, 106=Donga, 107=Littoral, 108=Mono, 109=OuAmA, 110=Plateau, 111=Zou, 140=Barolong, 141=Central Bobonong, 142=Central Boteti, 143=Central Mahalapye, 144=Central Serowe/Palapye, 145=Central Tutume, 146=Chobe, 147=Francistown, 148=Gaborone, 149=Ghanzi, 150=Jwaneng, 151= Kgalagadi South, 152=Kgalagadi North, 153=Kgatleng, 154=Kweneng East, 155=Kweneng West, 156=Lobatse, 157=Ngamiland East, 158=Ngamiland West, 159=Ngwaketse, 160=Ngwaketse West, 161=North East, 162=Selibe Pikwe, 163=South East, 180=Boucle du Mouhoun, 181=Cascades, 182=Centre, 183=Centre-East, 184=Centre-North, 185=Centre-West, 186=Centre-South, 187=East, 188=Hauts-Bassins, 189=North, 190=Plateau Central, 191=Sahel, 192=South West, 220=Santo Antão, 221=São Vicente, 222=Santiago – Interior, 223=Santiago – Praia, 224=Fogo, 260=Western, 261=Central, 262=Greater Accra, 263=Volta, 264=Eastern, 265=Ashanti, 266=Brong Ahafo, 267=Northern, 268=Upper East, 269=Upper West, 300=Nairobi, 301=Central, 302=Eastern, 303=Rift Valley, 304=Nyanza, 305=Western, 306=North Eastern, 307=Coast, 340=Butha-Buthe, 341=Leribe, 342=Berea, 343=Maseru, 344=Mafeteng, 345=Mohale's Hoek, 346=Quthing, 347=Qacha's Nek, 348=Mokhotlong, 349=Thaba-Tseka, 380=Bomi, 381=Bong, 382=Gbarpolu, 383=Grand Bassa, 384=Grand Cape Mount, 385=Grand Gedeh, 386=Grand Kru, 387=Lofa, 388=Margibi, 389=Maryland, 390=Montserrado, 391=Nimba, 392=Rivercess, 393=River Gee, 394=Sinoe, 420=Antananarivo, 421=Fianarantsoa, 422=Toamasina, 423=Mahajanga, 424=Toliary, 425=Antsiranana, 460=Central, 461=North, 462=South, 501=Kayes, 502=Koulikoro, 503=Sikasso, 504=Ségou, 505=Mopti, 506=Tombouctou , 507=Gao, 508=Kidal, 509=Bamako, 540=Maputo, 541=Maputo City, 542=Gaza, 543=Inhambane, 544=Sofala, 545=Tete, 546=Manica, 547=Zambezia, 548=Nampula, 549=Niassa, 550=Cabo Delgado, 580=Caprivi, 581=Erongo, 582=Hardap, 583=Karas, 584=Kavango, 585=Khomas, 586=Kunene, 587=Ohangwena, 588=Omaheke, 589=Omusati, 590=Oshana, 591=Oshikoto, 592=Otjozondjupa, 620=Abia, 621=Adamawa, 622=Akwa-lbom, 623=Anambra, 624=Bauchi, 625=Bayelsa, 626=Benue, 627=Borno, 628=Cross-River, 629=Delta, 630=Ebonyi, 631=Edo, 632=Ekiti, 633=Enugu, 634=FCT, 635=Gombe, 636=Imo, 637=Jigawa, 638=Kaduna, 639=Kano, 640=Katsina, 641=Kebbi, 642=Kogi, 643=Kwara, 644=Lagos, 645=Nassarawa, 646=Niger, 647=Ogun, 648=Ondo, 649=Osun, 650=Oyo, 651=Plateau, 652=Rivers, 653=Sokoto, 654=Taraba, 655=Yobe, 656=Zamfara, 660=Dakar, 661=Diourbel, 662=Fatick, 663=Kaolack, 664=Kolda, 665=Louga, 666=Matam, 667=Saint-Louis, 668=Tambacounda, 669=Thiès, 670=Ziguinchor, 700=Eastern Cape, 701=Free State, 702=Gauteng, 703=Kwazulu Natal, 704=Limpopo, 705=Mpumalanga, 706=North West, 707=Northern Cape, 708=Western Cape, 740=Dodoma, 741=Arusha, 742=Kilimanjaro, 743=Tanga, 744=Morogoro, 745=Coast (Pwani), 746=Dar es Salaam, 747 =Lindi, 748=Mtwara, 749=Ruvuma, 750=Iringa, 751=Mbeya, 752=Singida, 753=Tabora, 754=Rukwa, 755=Kigoma, 756=Shinyanga, 757=Kagera, 758=Mwanza, 759=Mara, 760=Manyara, 761=North Unguja, 762=South Unguja, 763=Urban West, 764=North Pemba, 765=South Pemba, 780=Central, 781=West, 782=North, 783=East, 784=Kampala, 820=Lusaka, 821=Central, 822=Copperbelt, 823=Eastern, 824=Luapula, 825=Northern, 826=North-Western, 827=Southern, 828=Western, 860=Harare, 861=Bulawayo, 862=Midlands, 863=Masvingo, 864=Mashonaland East, 865=Mashonaland West, 866=Mashonaland Central, 867=Matebeleland South, 868=Matebeleland North, 869=Manicaland,