Power for empowerment: the impact of electricity on women and children in Subsahara Africa

Emile Tenezakis*
Paris School of Economics

Ahmed Tritah†
Le Mans Université, Chaire Energie et Prospérité

June 2019

Abstract

Is electricity a vehicle of children and women empowerment in poor rural areas of the developing world? In this paper we approach the question through the human capital and labor market transmission channels using a rich and detailed micro-level household survey data from Rwanda. Our analysis considers effects occurring directly at the household level and indirectly at the village economy level. We address endogeneity raised by the non-random distribution of grid access across villages and household selection into treatment within villages using a combination of inverse probability weighting methods and instrumental variable strategy. Reweighting allows us to carefully choose the control group in a multilevel treatment context in order to uncover the direct and indirect effect of electrification, while the instrument deal with the unobserved selection bias at the village or household level. As a whole, we find little impact of electricity on children employment. A distinction by gender shows however that spillover effects bring boys into wage employment, while the direct household electrification exerts an effect in the opposite direction. Additional time devoted to work does not result in any negative impact on schooling. Women participation to paid employment is not affected by household connection or village access to the grid. Instead, women that are already likely to participate to paid employment, work longer hours as wage worker and in independent businesses. Looking at inequalities between spouses, women position relative to their male spouse is improved in terms of employment participation due to their greater involvement into family businesses.

Keywords: Labor, Child-labor, Energy, Poverty, Human Capital, Sub-Sahara Africa, Treatment Effects

*Contact: emile.tenezakis@polytechnique.edu
†Contact: ahmed.tritah@univ-lemans.fr

We gratefully acknowledge the Chaire Energie and Prospérité for its support, and particularly Anna Creti and Jean-Pierre Ponssard for their encouragement. We are grateful to participants at the conference Energie Climate and Development at Université Paris Dauphine and particularly to Jörg Peters for their insightful comments.

1 Introduction : rural electrification and development

It is widely acknowledged that access to modern energy services is a prerequisite for the economic, and social development of populations in Africa, especially in rural areas. As of 2014, 87% of the population worldwide had access to electricity, whereas only 43% Sub-Saharan Africans were able to use electricity at home. For the particular case of Rwanda only roughly 20% of its inhabitants have access to electricity. The lack of energy is even more acute in rural areas : in 2014, 17% Rural Sub-Saharan Africans were connected to the grid, and 9% in Rwanda. Because of the low connection rates and their economic status, developing countries have the highest potential for energy consumption growth. As households rise out of poverty they purchase new assets, many of which use substantial amounts of energy: e.g. refrigerators, TVs (Wolfram et al., 2012). The channels through which access to modern form of energy impacts rural households are numerous and still not well understood and quantified. In this paper we focus on the human capital and labor market channel.

An intensive movement of rural electrification projects is underway,[1] across all regions of the developing world, and more recently in Sub-Saharan Africa (SSA). Since 2010, for Africa alone, 38 projects related to rural electrification were approved and supported by the World Bank with a commitment amount of approximately 4 billion USD. Despite these efforts, knowledge on the consequences, and economics underpinnings of electrification in poor rural areas still need to be scanned. A "natural" answer to this, which sounds as a truism, is that of course, populations need electricity to develop. This is true. Hadn't industrialized countries had access to modern form of energies, they wouldn't be industrialized today. But this does not predict at all the *potential consequences* of electrification and the channel of transmissions, in particular in poor rural areas.

At a macroeconomic level, access to electricity is a strong vector of productivity jumps ; factories that acquire electrified capital may decrease their production costs which would eventually decrease prices and enhance manufactured goods consumption (Rud, 2012). At the micro-economic level, mechanisms are fuzzier. Empirically, for several outcomes such as labor or education, empirical results are not conclusive and highly heterogenous. Lee et al. (2017) sheds light on the lack of micro-founded studies that assess the different impact mechanisms. Jimenez (2017) highlights this and points out the problem of the external validity of each of these works produced in very different contexts. One reason for this is that there is a plenty of theoretical channels through which access to electricity may impact microeconomic outcomes in rural areas. For example, one can expect that if electrified capital is acquired in a given area as a consequence of electrification, and the demand for low-skill labor increases, households trade-off between child education and work will be modified and child labor may increase as a consequence. Bernard (2012) and Lee et al. (2018) respectively, summarize the situation quite well. According to the first "No one doubts that rural electrification (RE) positively affects household well-being. In addition, if RE is not necessarily a sufficient condition to long-term development of rural areas, it is probably a necessary one." Quite in contrast, the second authors pointed out that "connecting rural households today is not necessarily an economically productive and high return activity in the worlds poorest countries".

This paper aims at improving the stock of available knowledge with a specific highlight on children

[1] We will make use of RE as an abbreviation for rural electrification

and women time use, which we relate to their education and employment outcomes. Our contribution is twofold. First, our focus on both children education and employment, and women employment outcomes is to our knowledge absent in the literature[2]. Impact on these two demographic groups, which we call the marginal workers, is of a particular importance for development policy. First they are largely over-represented in unpaid farm family businesses, and domestic household production. Several studies point out the importance of women empowerment, through access to paid employment, on growth and economic development (Duflo, 2012).[3] For instance, under certain circumstances, money that ends up in the hand of mothers rather than fathers increases expenditures on children human capital (Doepke and Tertilt, 2018). Therefore we could expect some spillover effects from women employment to their children human capital. Finally, the opportunity costs of working are socially very high for children as this comes at the expense of human capital investment which may hinder future welfare and growth prospects. Our second contribution is a more careful consideration of the channels of transmission and impact levels of rural electrification. Identifying the causal impact of electricity at the household level is challenging. Indeed, the treatment maybe considered twofold. A household may be treated because it is in a village with access to electricity, where the connection of its neighbors affect its own economic outcomes. This indirect treatment effect of neighborhood connections has largely been overlooked in the literature on impact of rural electrification. A household may be also treated because, given that its village is connected to the grid, it decides to use electricity in its dwelling. This treatment effect combines both the direct effect of access to electricity in the dwelling and the indirect effect that works through spillover effects and general equilibrium effects within the village economy.[4] For this reason, endogeneity with respect to outcomes of interest arises at two levels. First grid connections are not randomly distributed across villages, some villages are more likely to be connected than others according to characteristics related to their potential economic outcomes. Second, given village access, households within the village decide whether to connect or not. This decision depends on household characteristics, of which some are endogenous to employment and human capital decisions. For instance workers with high wage potential may be more likely to connect than others and at the same time these workers are more likely to engage in employment and for longer hours.[5]

We tackle the issue of multi-level effect of electrification by distinguishing direct and indirect treatments effects. To do so, we discern outcomes for which we suspect village access to be endogenous, these are outcomes measured at the extensive margins, such as labor market participation and school attendance, from outcomes for which we suspect, conditional on the village having access to the grid, that the source of endogeneity occurs at the household level. These are outcomes measured conditional on participation such as hours worked in different type of activities and grade repetition for children attending school.

We deal with the endogeneity occurring at the village level with the use of an instrumental variable

[2]Our approach is related to Salmon and Tanguy (2016) who study jointly wife and husband labor outcome in Nigeria.

[3]Female empowerment figures as a priority in the United Nations' Millennium Development Goals, which postulate: "putting resources into poor women's hands while promoting gender equality in the household and society results in large development payoffs. Expanding women's opportunities [. . .] accelerates economic growth."

[4]In the context of a cash transfer program evaluated using a large RCT in Bangladesh, Bandiera et al. (2017) shows that not taking into account village level spillover leads to a large underestimatation of the impacts, that affect directly some individuals and indirectly others. They suggest that the correct level of analysis in that case is the village economy.

[5]We are assuming here that in the context of poor countries, substitution effects dominate income effects, therefore an increase in wage potential raises labor supply.

strategy. Our argument follows a cost-approach, and considers potential altitude variation within a village as a predictor of village connection cost. We predict this altitude variation by a proxy for altitude level of a village. We show that this altitude is highly correlated with the main source of water used by all the households in the village. Therefore we instrument village access with a binary variable equal to one if all households are using spring water in a village. We provide geographical evidence to support our approach which shows that the use of spring water is very frequent in high altitude areas, while it is nevertheless little correlated with male employment.

We assume that sources of endogeneity regarding the decision on hours of work or educational outcome, *conditional on participation* is mostly determined within the "household economy". This may be due to household specific role gender preferences, intra-household bargaining, or household specific assets, which we will all assume independent of the grid connection status of the village. Moreover, under the assumption that village-level spillover effects are of second order with respect to the time allocated to different activities (working or studying), we could instrument the decision to connect with the access to the grid status of the village. Throughout all the analysis, we control for a large set of household observable characteristics, village remoteness and sub-regional fixed effects.

Finally, since we are in a context of multilevel treatment we use inverse-probability weighting method to benchmark the estimated coefficients for each level of treatment with respect to the characteristics of connected household since these are the ones that benefit from both the direct and indirect treatment effects. Evaluating each treatment effect across the same set of characteristics allows to take into account discrepancies of results between both level of treatments that may be due to heterogeneous treatment effects, and allows to uncover the total effect of electrification by adding the estimated direct and indirect effects.

Our findings are as follows. As whole, we find little impact of electricity on children employment. However a distinction by gender shows however that spillover effect pull boys into wage employment, this effect is totally compensated by the direct of electrification within household which instead reduces employment among boys. Surprisingly this additional time devoted to work does not result in any negative impact on schooling. As for Women, they are more likely to engage in paid employment, though the main beneficiaries seems to be the already working women. Indeed, participation is little affected affected by direct household connection or village access to the grid. Instead, already working women, work longer hours both as a dependant and independent worker. Lately, looking at inequalities between spouses, we find that women position relative to their male spouse improves in terms of employment participation, due to their greater involvement into independent business. This effect is nevertheless more than compensated by a relatively larger increase in hours worked among male spouse.

The rest of the paper is organized as follows. In the second section we provide macroeconomic stylized facts and present microeconomic mechanisms that we wish to investigate. In the third section we present our data and variables of interest. In the fourth we discuss identification issues and chosen empirical strategy. In the fifth we present the main results. The sixth section concludes. Additional results and robustness checks are provided in a separate appendix.

2 Macroeconomic stylized facts and microeconomic theory

Our motivation is twofold. On the one hand, in SSA the situation in terms of human capital and labor outcomes is of particular interest for children and women. Child labor of any kind is still pervasive, and women participation to remunerated labor activities, in the form of independent (business-like) or wage labor, is still marginal. On the other hand, electrification comes with many promises, including that of allowing children to study more instead of working, and allowing women to work independently for a remuneration. Nevertheless, one must have in mind that electricity is an *enabling energy.* Households or firms don't use electricity as an input for consumption or production. Therefore, outcomes depend on the grid access quality and ultimately on the availability of complementary energy-using assets (Squires, 2015; Wolfram et al., 2012), for both households and firms.

As for households, for instance in India (World Bank, 2004), for electricity to allow people to save time they must decide to acquire electric stoves to replace their wood-based cooking appliances and thus avoid having to look for wood everyday. Impacts on households differ whether people buy TVs or electric stoves as a first asset. TV may improve marginal value of leisure time, electric stoves may improve household chores productivity. If household faces financial constraints - which is systematic in SSA where credit market is underdeveloped - it may be impossible to acquire first assets. For the specific case of Rwanda Lenz et al. (2017) show that there is little productive usage of electricity among connected households. A finding we confirm in our large household sample survey. An evidence of costs constraint in the acquisition of complementary assets in Africa is provided by the surge of import from China of cheap electrified assets following the accession of the country to WTO in 2001. This represented for SSA countries a positive supply shock, if the constraints on asset acquisitions were the availability of electricity we should not have observed important increase in imports. Instead, data suggests a strong reaction in several countries, for instance in Kenya where refrigerators imports increased dramatically after 2002.[6]. This suggests that cost constraints on complementary assets are probably an important barrier for the productive usage of electricity, and therefore also for electricity take-up in grid catchment areas.

Second, at the firm-level, if companies do connect to the grid and acquire electrified capital requiring labor to be used, then labor demand in the locality should increase. The literature on firm-side effect in poor countries is scarce (Rud, 2012). Beside a low access rate, SSA countries have also a low quality of electricity connection : there is a high variance among countries with respect to number of outages and those with the highest number are typically center-SSA countries.[7] In these conditions, as shown by Alby et al. (2013), firms may decide not to invest in electrical features from which they get low expected returns given the quality of access.

Third, provided that energy is used by both firms and households, the household-level causality chain is complex. Intuitive effects of electric appliances use are numerous and there is a growing body of literature unequally investigating each path. In developing countries rural areas men and women allocate their time among labor market work, home business, household chores, leisure, and education-devoted time for children. The use of appliances has effect on all these activities returns and costs.

[6] Some empirical evidence, obtained from the COMTRADE database is available from the authors upon request
[7] World Bank enterprise survey data

With electricity usage, home production in developing countries is exposed to a positive technology shock (Grogan and Sadanand, 2009). For example, electric stoves or refrigerator allow individuals to save time in home activities. This may either reduce or increase the time allocated to chores : there is a substitution effect because of household chores higher productivity and an endowment effect because of a higher overall time resources (Dinkelman, 2011). The use of modern form of energy like electricity, would plausibly decrease time allocated to chores as the home produced goods are bounded from above (Kohlin et al., 2011). In addition, we expect in one hand that, following home production shock, women are freed from some chores, allowing them to go work in the labor market or open micro-enterprises and on the other hand this relationship may also reflect a demand for labor effect from firms instead of a supply effect from households. In any case interpreting such a relationship requires a careful consideration at the micro level. For instance, Salmon and Tanguy (2016) found effect only for men when taking into account dependence of decision within households. As for home business opportunities, Dinkelman (2011); Kirubi et al. (2009) showed that they are indeed stimulated with electricity access. Time allocation to human capital acquisition may also be impacted. With electric appliances, returns to educational investment may be higher : the daytime is extended with light, the quality of studying is improved with communication technologies, school education is of higher quality. But two other phenomena may occur: a substitution of chores from women to children, or an increasing opportunity cost of education if low-skilled labor demand increases. Again there is no consensus. Some authors do find that education outcomes improve after electrification, some find no effect, and others find that children are worse off with respect to their education outcomes Jimenez (2017).

3 Data and variables definition

Our data source is the EICV4, the cross-sample of the integrated household living conditions survey provided by the national institute of statistics of Rwanda. It is the fourth round of a household survey conducted every five years, across two years. The EICV4 was implemented in 2013-2014. It contains a wide range of socio-economic information including labor and education for 66'081 individuals in 14'419 households distributed across 1'230 villages covering 30 districts. There are four levels of information: district, village, household, and individual. We restrict our analysis to rural households composed of monogamous couple and at least one children. The reason we focus on rural households is threefold: first, the outcomes are of particular interest in rural areas where child labor is still prevalent and women participation to remunerated labor too marginal, see Table 2. Second, developing access in rural areas is more costly and therefore knowledge of impacts is necessary. Third, in urban areas electrification rates are very high and an exogenous variation in access is more plausible in rural areas; this is particularly the case in the Rwandese context where the government is engaged in a large national roll-out electrification program (Lenz et al., 2017). Grid extensions prioritization are determined by costs criteria which, to some extent, are arguably exogenous to village level job creation potentials and household level employment decisions.[8] The reason we focus on monogamous couples is the following: we are interested in outcomes for children and women that are interrelated and are close substitute for

[8]The national electrification strategy of the actual government has a target to electrify 100% of household by the year 2020, given the actual state of progress, it seem that the target is too ambitious

various time consuming tasks in domestic production at the household unitary-level. Therefore our household-level analysis will be limited to those households, and our individual outcomes analysis will be limited to the female spouses, whether they are household head or not, and to children of these households. We loose 6'415 households and 22'533 individuals who are not in these type of households - note that those lost households were mostly single-female-headed households. We further loose 6'711 individuals and 1'174 households from in-rural restriction. Our final sample is composed of 36'798 individuals in 6'830 households distributed across 1'013 villages.

We suggested that employment outcomes effect of electrification may occur through higher home-production efficiency : but what if one is in a situation where households are even too poor to buy electrified assets such as new stoves or fridges? As we shall see, this is precisely the case of rural Rwanda.

In Table 1 we present various household-level summary statistics related to equipment and energy usage, and composition, wealth and education characteristics, across electrification status. We refer in the text to "connected" households as those that were in a village with access to the grid and that decided to connect. "Access" households are those households that had the choice to connect but did not. "Non-access" households are those that are in village without access and therefore, did not have the choice to connect. In the first part of the table, we observe that almost no household own time-saving assets such as refrigerator or cooker, whatever electrification status. In addition, no household uses electricity as a main cooking fuel. We observe that once household are connected, they don't buy time-saving assets but rather leisure-valuing assets, such as radios, phones, and most importantly, TVs and DVD players. For the latter we are more certain, because, when comparing access but non-connected and non access households, there is almost no difference: the difference arises once only one households are connected. Furthermore, while almost all non-connected household, what ever their access status, cook with wood, this proportion drops among connected households to 75%. At the same time, very few non connected households are using charcoal for cooking, instead 24% of connected household use this form of combustible. Finally, we observe a strong positive difference in wood and coal purchases once household connects in villages covered by the grid. The rise is substantially bigger than that seen among non connected households in villages with access and non-access to the grid. Household also spend as much as twice more time foraging for woods where they are not connected. This amount of time, around 5 hours, is similar among non connected households whatever their village connection status. Altogether, this very preliminary evidence suggests that any effect of electrification on employment outcome is not mediated through the acquisition of time-saving and more efficient assets in household production. An alternative plausible mechanism may be at stake: the fact that all connected households use electric grid as an efficient and cheap lighting source as shown in the second part of Table 1. The effect of grid-electrification on electric light usage has already been demonstrated by (Bensch et al., 2011). This represents a shock in terms of light time endowment for households, allowing their members to better organize their different activities throughout the day. These gains in time allow them to provide more work outside household production, to have access to cash to buy goods that they would otherwise have to produce.

We turn to household characteristics. There is no observable difference in the gender of household and its age, mostly because of our sample restriction to monogamous couples. The composition of

Table 1: Household summary statistics by connection status

	Access to the grid		No access to the grid
	Connected	Not Connected	
	(1)	(2)	(3)
	Equipment and energy usage		
purchased wood	0.5530	0.3550	0.2520
purchased coal	0.5293	0.1756	0.0772
hours spent foraging wood	2.7272	5.9464	6.3413
primary fuel for cooking is firewood	0.7570	0.9390	0.9801
primary fuel for cooking is charcoal	0.2402	0.0487	0.0109
primary fuel for cooking is electricity	0.0000	0.0000	0.0000
owns a radio	0.8370	0.6410	0.6460
owns a mobile phone	0.9560	0.6850	0.6020
owns a TV	0.3780	0.0149	0.0050
owns a satellite dish	0.0263	0.0000	0.0006
owns a DVD player	0.2780	0.0094	0.0038
owns a cooker	0.0069	0.00118	0.0002
owns an electric fan	0.0015	0.0000	0.0000
owns a refrigerator	0.0193	0.0010	0.0000
owns an electric burner	0.0018	0.0000	0.0000
Observations	605	1413	4806
	Composition, education and wealth		
HH is male	0.999	0.997	0.997
HH age	38.73	39.97	40.04
HH is literate	0.889	0.692	0.689
size of household	4.744	4.767	4.726
size of household (extended)	5.717	5.403	5.351
number of literate (extended)	3.283	2.413	2.371
nb children of HH below 6 yo	1.209	1.228	1.215
nb children of HH above 6 and below 12 yo	0.964	0.986	0.974
nb children of HH above 12 and below 18 yo	0.658	0.633	0.628
nb (+18 and -60) persons (extended)	2.454	2.267	2.282
nb of (+60) persons (extended)	0.0371	0.0773	0.0679
uses piped water (home or public)	0.517	0.368	0.182
uses electricity for light	0.998	0.487	0.667
cement walls	0.563	0.194	0.150
distance to closest road (meters)	198.1	356.4	901.6
all households use spring water in the village	0.0598	0.0870	0.204
Observations	604	1411	4774

Source: Authors computation from EICV4 Data.

Notes. Weighted proportions are reported for each sample using the population sampling weights available in the EICV4. In column (1) we refer to the sample of households in villages that have access to the grid and that are connected. In column (2) we refer to households in villages that are reached by the grid but are not connected. The column (3) refers to the sample of households in villages outside the grid catchment area. The term "HH" is a shortcut for "Household head". The term "Extended" refers to *all* individuals reported as member of the household, not only the couple and their kids. Electricity usage for light can be from the grid or not (generators, battery, solar panels among others).

the household is quite similar, with nevertheless slightly higher average number of young children and slightly lower average number of old persons in connected households. The education characteristics are different: 89% of connected households have a household head who can read or write, versus approximately 68% for non connected households whatever their village connection status. Similar difference is observed when looking at number of literate in the household. This underline that connected household tend have higher human capital which may explain their potentially greater demand for leisure diversification; suggesting an issue of *household self-selection into treatment* in connected areas. Conditional on having access, more educated, and probably better-off households are more likely to connect. Similar observation can be made when looking at indicators of wealth, such as the proportion of households with dwelling walls made in cement: 50% have such walls among connected, versus 19% among non-connected within the grid catchment area and 15% for non connected in villages without access to the grid. When we look at variables that depend on village infrastructures, we have some evidence of "*village program placement*". On average, households in villages with access to the grid are closer to the main road, and this is also true for connected households compared to non-connected in this villages. However, as for most characteristics, there is much more disparities between connected and non-connected households within villages with access, than between villages among non-connected households. Indeed, if the households in the village are on average far away from the road, it may be costly to connect that village because electric grid is often located close to the road. This spatial characteristic is actually often used as an instrument for village access (Salmon and Tanguy, 2016). Our instrument for village access is a binary variable indicating whether or not all the households in the village use springs as a main source for water, it turns out that in electrified villages 6% of connected households do so against 9% among non-connected access, but 20% of households without access to the grid are of this type. Further support for this instrument is provided in section 4.2.

Overall Table 1 highlights also an interesting feature of our sample. Non-connected households in electrified villages are very similar in their characteristics to households in non-electrified villages. This suggests that differences in electricity related outcomes among these households are more likely to be driven by differences between their village characteristics, rather than their household level characteristics. In particular, the fact that non-connected household in villages with access to the grid benefit from spillover and general equilibrium effects in the village economy. This is important because village-level electricity impact can be identified only by exploiting outcome differences among non-electrified households, which indeed are the only ones for which village connection status varies. We will exploit this fact to identify spillover impacts at the village level.

To further highlight the last point, we supplement this presentation with the Figure 1. As detailed in the section 4.1 and in appendix, we implement an inverse-probability weighting using the estimated propensity score to balance the observable characteristics between connected households in one hand, and non-connected households, whatever their village access status, on the other hand. The propensity scores (estimated conditional probabilities) of being connected *conditional on having such a choice* are shown in the two panel for the three treatment status of interest: the connected, the non-connected with access, and those with no access. In the left hand-side panel of Figure 1 we present predicted probabilities estimated using a set of household and village level characteristics, without reweighting, and in right hand-side panel with reweighting. As expected, the distribution of the estimated propensity score is right skewed for connected households. Indeed, as shown in Table 1 these households have specific characteristics that makes them more likely to connect. More interestingly, the distribution of the propensity scores among the non-connected household is highly skewed to the left and indeed very similar whatever the village access status along the whole support of the propensity score distribution. This makes a clear case that among these households there is probably limited scope of self-selection into village treatment status. Instead this selection is much more likely among households within the grid catchment areas. Lately, the right hand panel, present the distribution of the score after reweighting to insure a balancing of the characteristics between households. We re-balance non-connected household characteristics to mimic those of connected households. We can see that the propensity to connect is now much more similar across the three groups across the whole support of the propensity score distribution. In the next section we will further discuss to what extent this balancing of characteristics using the propensity score could allow us to uncover the direct and indirect effect of connection on household level outcome.

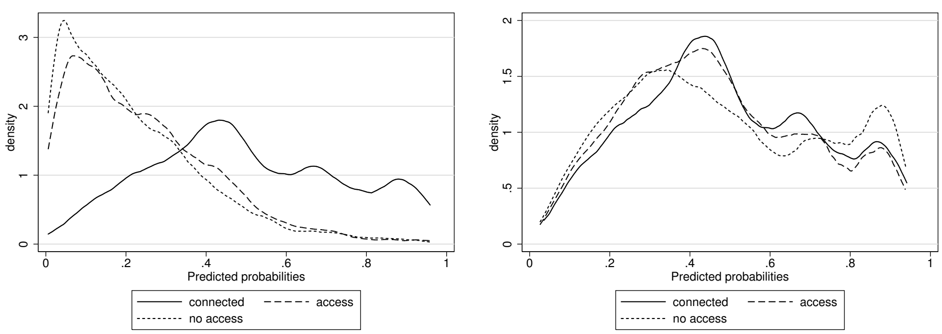

Figure 1: Propensity scores distribution across connection status without reweighting (left) and with reweighting (right)

We are interested in labor outcomes for female spouses of each household and labor and education outcomes of children of the household head couple. We define the *female spouse* of the household as the female member of the household head couple.[9] We define a *children* as being one of the child of the household head couple, therefore we do not consider, for example, nephews or domestic workers. Our data on labor consists in a job-level file, where each individual may have multiple jobs of different types, and for each of these jobs the number of hours of activity in the previous 7 days is reported. We distinguish six types of labor: (1) family non-paid workers, (2) independent workers, or (3) wage workers. For each of this three category, the labor can be performed in-farm or out-of-farm. Overall this allows us to define 6 different type of labor. We refer to any type of labor as "labor", to remunerated independent or wage labor as "paid", and to any non-remunerated family activity as "unpaid". Our goal is to evaluate to what extent electrification allowed women to be released from unpaid labor, and to participate into paid labor in the form of micro-enterprise that we call "independent" or wage labor that we call "wage". Therefore we will mostly focus on participation and hours worked in "unpaid" on one hand, and participation and hours worked into "paid" on the other hand. We will investigate to what extent access and connection to electricity triggered a transfer of activity between these categories of labor at both intensive and extensive margins. For children, we aim to evaluate

[9]Homosexual couples are anecdotal in this data.

the capacity of electrification to incentivize families to send them to school or increase the intensity of studies rather than sending them to do any kind of labor, be it unpaid or paid. Therefore we are interested in "labor" for them.

Table 2 provides individual summary statistics for children, women, and their male spouses, for several individual characteristics and outcomes of interest.

We are interested in outcomes for children of more than 6 years old, for which there is a trade-off between school and labor in the context of rural areas where farm labor is important. We observe that in our sample, around 9% of the children are in some form of employment, this share is almost twice higher among non connected household (10%-11%) as compared to connected households (6%). This is mostly unpaid; but also, marginally, paid labor (4%-5%) in non-connected areas, in connected areas children paid labor is almost existent. When they do participate to labor, they work around 20 hours in unconnected households, against 16 hours for the connected. In non connected households when children exert a paid job they work a substantial amount of hours (20 hours-25 hours). Notice that unpaid labor and paid labor doesn't sum up to all labor: they may overlap within an individual because some individuals including children report multiple jobs. Overall this result suggest that child employment is more prevalent among non connected household, and the difference is relatively independent of the village connection status, suggested limited spillover effect in connected villages. We define attendance as a binary indicating whether the children attended school in previous 12 months. While school attendance is rather high in connected areas (94%) it drops to around 85% among the non-connected whatever is village connection status. There is also substantial difference in the quality of education the quality. We measure this quality with grade repetition across two subsequent schooling years. We form a binary variable indicating whether, in the sample of children who where in primary or secondary-level class in 2012, the individual is in the same class in 2013. We find that as much as 18% of the children repeated class among connected households, and this share jump to 26% among non-connected households. To refine our measure of quality of studies, we define across-year drop-out as a binary equal to one if we observe that, in the sample of children who were in primary or secondary level in 2012 and not in the last class of either level (to rule out level completion), they are not in class anymore in 2013. Only 3% of children dropped-out, which is quite marginal.

The second bottom panel of Table 2 provides summary statistics of female spouses along the three treatment sample. Female spouses have quite a low educational background ; among connected household 90% of them have attended school in their life, and 84% are literate, though a majority has not completed primary education and only 13% have at least a secondary education diploma. Again non connected household have similar and substantially lower lever of education than the connected. Participation into any form of employment is relatively high, above 80% and higher among the non connected, though only half of spouses are in a paid employment in access area, this share is only 44% outside the grid catchment areas. A majority of spouses in non connected areas are instead in unpaid employment, which as high as 66% outside grid catchment area. This share is substantially lower among connected households (48%). The prevalence of independent labor is substantial, since almost a third of female spouses participate are independent workers, which we consider as "micro-enterprises", this share is higher for connected household (38% against 30% among the non connected). As for children wage employment is more frequent in non connected areas, at around 25%-22% while it is only 16% for connected households. Lately when female spouses do participate to labor, they work on average more in connected areas in all form of labor. The difference in working hours with non connected household is particularly important for wage labor since they work on average 36 hours a week, which is almost twice the hours worked among non-connected. Therefore although they are less likely to work when they are connected the quality of these jobs measures by hours of work seem to be higher among the connected. One should notice some important contrast with male spouses, for whom paid labor, unlike for women, is systematic whatever the connection status (above 90%) and unpaid labor is scarce while it is pervasive among female spouses. We can see that wage employment appear as a channel through which non connected household may benefit from village electrification spillover. Indeed, for both male and female this form of labor is more frequent among non-connected households in access areas. Lately when they are in employment, male work longer hours than female spouses, except in unpaid employment. This gender differences are actually the "*empowerment*" motivation of this paper and we will provide further insights on this in section 5.

Table 2: Individual summary statistics by connection status

	Children					
	Access to the grid				No access to the grid	
	Connected		Not Connected			
	(1)		(2)		(3)	
	Mean	N	Mean	N	Mean	N
age	11.25	1087	11.11	2549	11.10	8539
literate	0.81	684	0.66	1558	0.65	5161
labor	0.06	1087	0.10	2549	0.11	8539
paid	0.01	1087	0.04	2549	0.05	8539
unpaid	0.06	1087	0.07	2549	0.08	8539
labor^h	16.47	63	22.32	248	19.72	962
paid^h	12.35	9	24.43	108	19.26	390
unpaid^h	14.97	60	16.56	176	15.81	718
attended school	0.94	1087	0.86	2549	0.84	8539
repeated	0.18	816	0.26	1706	0.26	5539
dropped-out	0.02	766	0.03	1629	0.03	5371
Observations	1087		2549		8539	
	Female spouses					
age	35.05	605	35.61	1413	35.79	4806
ever been to school	0.90	605	0.81	1413	0.80	4806
literate	0.84	604	0.65	1413	0.64	4805
primary diploma	0.35	605	0.25	1413	0.25	4806
secondary diploma	0.13	605	0.19	1413	0.21	4806
labor	0.84	605	0.87	1413	0.92	4806
paid	0.51	605	0.50	1413	0.44	4806
unpaid	0.48	605	0.56	1413	0.66	4806
independent	0.38	605	0.30	1413	0.28	4806
wage	0.16	605	0.25	1413	0.22	4806
labor^h	31.81	516	27.22	1242	26.34	4420
paid^h	30.89	310	23.42	705	21.74	2138
unpaid^h	23.35	293	21.54	805	21.96	3185
independent^h	26.99	229	22.29	434	20.65	1357
wage^h	36.61	97	19.79	352	17.66	1044
Observations	605		1413		4806	

	Male spouses					
	connected		access		no access	
	Mean	N	Mean	N	Mean	N
age	38.73	605	39.99	1413	40.05	4806
ever been to school	0.94	605	0.81	1413	0.81	4806
literate	0.89	605	0.69	1413	0.69	4805
primary diploma	0.32	605	0.28	1413	0.26	4806
secondary diploma	0.11	605	0.19	1413	0.20	4806
labor	0.92	605	0.91	1413	0.91	4806
paid	0.92	605	0.91	1413	0.91	4806
unpaid	0.03	605	0.02	1413	0.03	4806
independent	0.64	605	0.61	1413	0.70	4806
wage	0.42	605	0.50	1413	0.43	4806
labor^h	43.02	554	35.69	1281	33.37	4361
paid^h	42.80	549	35.70	1271	33.16	4334
unpaid^h	16.85	20	9.15	35	13.89	125
independent^h	33.37	384	26.68	864	24.92	3350
wage^h	42.58	259	32.84	704	29.96	2070
Observations	605		1413		4806	

Source: Authors computation from EICV4 Data.

Notes. Weighted proportions and means are reported for each sample using the population sampling weights available in the EICV4. The number of hours workers in each category of employment is computed conditional on working positive hours in this category. The upper-script h above variables indicates number of hours worked over the last 7 days in the employment category. Individuals can hold multiple jobs.

To give a more synthetic picture of employment disparities across different treatment status, we provide in figure 2 estimation of work hours distribution for female spouses and children, for each connection status, *conditional on participating to the given type of labor.* We report weekly hours distribution estimation for paid labor and unpaid labor for female spouses, and the same distribution for general labor for children. We observe differences in distributions across connection statuses. Hours worked are more right skewed for paid hours for female spouses in connected dwellings, compared to nob-connected household, for which there is very little differences. As of unpaid labor, the difference is much less obvious, distribution of hours are only slightly right skewed among the non connected. For children unpaid hours, the observation is similar but the difference is stronger.

To complete the picture of employment differences across households in relation to their connection and access status; we provide a set of simple linear regressions of labor and education outcomes on household electrification status and village electrification status in Table 3, without controlling for anything. They should not be given a causal causal interpretation and are here to motivate next empirical exercises and the double treatment that we refer to in 4.2. Indeed, non electrified households in connected areas may witness a change in their employment outcomes due to changes in local labor demand and not from a direct change in household production efficiency brought by electricity. The effect of the village electricity status on employment outcome is identified by comparing non connected households in villages with access to the grid to households in villages outside the grid catchment area. The direct effect of being connected to electricity is measured in these tables by the coefficient associated to the variable $E^{household}$. This coefficient is an estimate of mean differences in outcome between connected households and non connected household in access areas, because $E^{village}$ is controlled for. The "indirect" effect is measured by the coefficient associated to the variable $E^{village}$, the coefficient is an estimate of the mean difference in outcome between non-connected household across villages with access and without access to the grid, because $E^{household}$ is controlled for. In principle if both treatment status (access and connection status) where randomly assigned across households then adding up the two coefficient will provide an estimate of the total effect of electrification. However as forcefully shown in the previous descriptive statistics this is clearly not the case.

In general this table suggest a relationship between village connection and amount of labor participation, whereas the relationship with the intensive margin labor materializes mostly at the household level. Note, importantly, that there is much less difference in observable characteristics between access and non-access non connected households than there are between connected and non-connected households in access areas. Therefore, the coefficients associated to household electrification are likely to be more biased than that for village status electrification.

For children, coefficient on attendance is positive, strong for household connection and weak for access, and negative on repetition, strong for connection and insignificant for access. These results concord with what we are expecting for children: they work less, and study more and better. Although what we provide here are just mean differences across groups, these results suggest that the potential for electrification to improve educational outcome and reduce children employment occurs only if household connects to the grid. Otherwise there is at that stage no evidence that labor demand pulls non-connected children out of school in villages that have access to electricity.

For women, we expect paid labor participation to be affected by access more than connection, and naïve OLS agrees with this. It indicates that access increases the probability of participation by 5.7%. In contrast, connection affects strongly hours and increases them by 11, conditional on having access to the grid in the village, whereas village access increases them by only 2.8. Unpaid labor participation seems to be strongly affected by both connection and access ; independent paid labor is not affected in participation but in hours, positively. Wage labor participation seems to be affected positively by access but negatively by connection, and hours of wage labor seems to be strongly positively affected by connection and barely affected by access. The high and quite implausible result on wage hours tend to indicate a sample selection bias that we will discuss further.

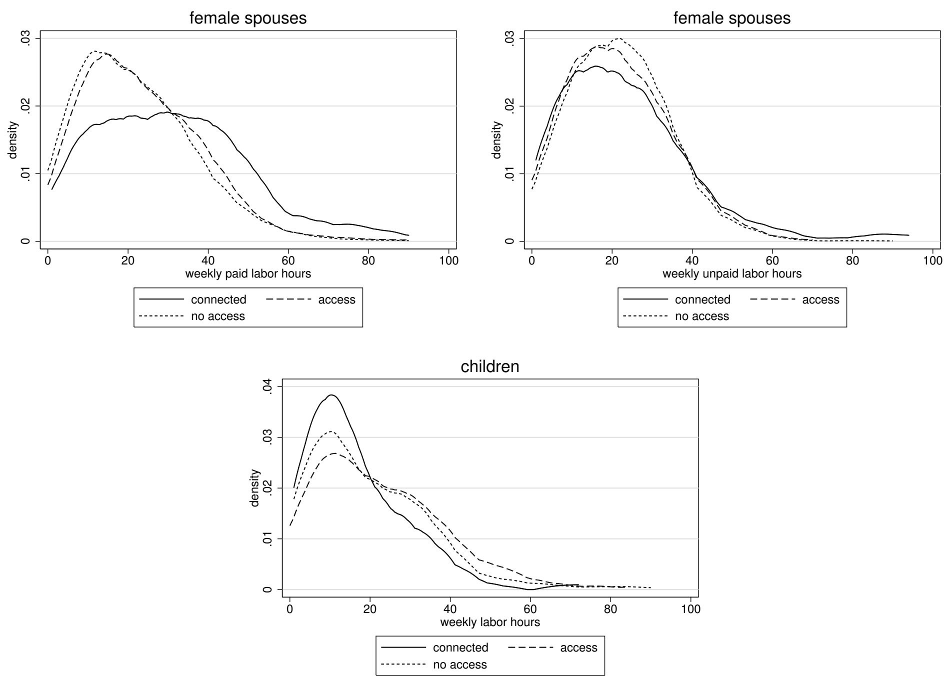

Figure 2: Estimation of work hours distribution over various labor type

Table 3: Simple OLS without covariates on individual outcomes of interest

	Female spouses							
	paid		unpaid		independent		wage	
	Extensive	Intensive	Extensive	Intensive	Extensive	Intensive	Extensive	Intensive
$E^{household}$	0.0109	7.472***	-0.0875***	1.805	0.0725***	4.697***	-0.0912***	16.82***
	(0.0260)	(1.321)	(0.0259)	(1.143)	(0.0247)	(1.624)	(0.0196)	(1.964)
$E^{village}$	0.0561***	1.679***	-0.0987***	-0.420	0.0218	1.634**	0.0329**	2.132***
	(0.0156)	(0.644)	(0.0154)	(0.510)	(0.0142)	(0.827)	(0.0135)	(0.804)
N	6824	3153	6824	4283	6824	2020	6824	1493
	Children							
	labor		schooling					
	Extensive	Intensive	attendance	repetition				
$E^{household}$	-0.0395***	-5.851***	0.0723***	-0.0768***				
	(0.00986)	(1.860)	(0.0110)	(0.0179)				
$E^{village}$	-0.0142**	2.594**	0.0206**	-0.00167				
	(0.00710)	(1.128)	(0.00830)	(0.0124)				
N	12175	1273	12175	8061				

Statistical significance are * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors are reported in parenthesis under coefficient estimates. The columns label "paid" refers to any paid employment; those label "unpaid" refer to any form of unpaid employment, those label "independent" refer to self-employment, and those label "wage" refer to wage employment. Extensive margin is the participation decision and include all the eligible sample, the variable take the value 1 if at least one hour is supplied during the last 7 days in this form of employment. The extensive margin refer to the number of hours worked conditional on supplying at least one hour. $E^{village}$ and $E^{household}$ are a binary variables taking respectively the value of one if the village is connected to the grid, and the value of one if the household dwelling is connected.

4 Identification issues and empirical strategy

4.1 Control methods and inverse probability weighting using the propensity score

In the basic OLS specification, we control for household wealth, composition and education characteristics variables that are presented in section 3. We also include results of regressions analyzing the effect of connection and access on time allocation to various household chores, that are presented in appendix. We make use of a Tobit model to account for the 0-hours of chores.

Since we are in the context of a multi-level treatment, we make use inverse-probability weighting (IPW) as both an additional control method, and a method to obtain comparable groups. A first advantage is that we avoid any misspecification bias coming from wrong functional form assumption. As shown by (Imbens and Wooldridge, 2008) specification bias may be important in the context of high heterogeneity between treated and control groups, as it is the case in our context.[10] A second and important advantage in our context, is that IPW, because it allows us to obtain comparable groups, allows us to infer a *total* effect of electrification when summing up direct household effect and indirect village level spillover effect. IPW method lies in the class of matching methods, for which practical guidance is provided in Caliendo and Kopeinig (2005). We estimate a probability of connection within the sample of households which have access to electricity, with a logistic model. We then estimate the probability of connection on the whole sample. We call these probabilities *propensity scores*. Then, we include these scores as weights in our regressions. The purpose is to construct counterfactuals, in order to be sure that we compare comparable households. Formally, we rely on the Roy-Rubin model. Let $E_{jk} = \{0, 1\}$ denote the connection status of household j in village k. Let Y_{ijk} be our outcome of interest for individual i. $Y_{ijk}(E_{jk})$ is the potential outcome for individual i given the connection status of his household. Then, the treatment effect for individual i is $\tau_{ijk} = Y_{ijk}(1) - Y_{ijk}(0)$. The fundamental problem is that we do not observe both potential outcomes, but only one. Our parameter of interest is the average treatment effect on the treated :

$$\tau_{ATT} = \mathbb{E}(Y_{ijk}(1)|E_{jk} = 1) - \mathbb{E}(Y_{ijk}(0)|E_{jk} = 1) \tag{1}$$

We don't observe the second term, so we must construct a counterfactual. Using the expectation for the untreated is not a good idea because of self-selection bias:

$$\mathbb{E}(Y_{ijk}(1)|E_{jk} = 1) - \mathbb{E}(Y_{ijk}(0)|E_{jk} = 0) = \tau_{ATT} + B \tag{2}$$

We have to assume that we have a set of observables X_{jk} that explains E_{jk}. We can then invoke the following two assumptions to reach an identification strategy:

[10] Our computations of standardized mean differences across connected and non-connected households, as suggested by (Imbens and Wooldridge, 2008), show that there is a large disparities across these groups. For most variables values are above the 0.25 threshold suggested by (Imbens and Wooldridge, 2008) as a rule of thumb to detect misspecification threat in OLS estimations

$$Y_{ijk}(0), Y_{ijk}(1) \perp\!\!\!\perp E | X_{jk} \textbf{(Unconfoundedness assumption)}$$

$$0 < \mathbb{P}(E_{jk} = 1 | X_{jk}) < 1 \textbf{(Common support)}$$

Given X_{jk} a vector of observable covariates, potential outcomes are independent of a person treatment status. Controlling in a flexible manner, without constraining functional form assumptions such as a linear one, for the whole X_{jk} vector is difficult. One may control only for a uni-dimensional function of this vector. The *Propensity score* is one function:

$$\mathbb{P}(E_{jk} = 1 | X_{jk}) = P(X_{jk}) \quad \textbf{(Propensity score)} \tag{3}$$

The intuition is that we build a "balanced" sample of individuals. We try to find good counterfactuals of households with $E_{jk} = 1$ given X_{jk}. There are several ways to do so. One way is household to household matching: constructing pairs of households that have a similar propensity score. But this approach implies substantial data sacrifices. One intermediary way is to simply weight our regression with our propensity score. In our case, we fit the model using households in access villages :

$$\mathbb{P}(E_{jk} = 1 | X_{jk}, V_k = 1) = \Phi(X_{jk}^{'}\beta) \tag{4}$$

With $V_k = \{0, 1\}$ the dummy for village status. The vector of parameters β are estimated with maximum likelihood. We then predict probabilities for all the households in the sample:

$$\hat{p}_{jk} = \hat{\mathbb{P}}(E_{jk} = 1 | X_{jk}) = \Phi(X_{jk}^{'}\hat{\beta}) \tag{5}$$

And the weights in the regressions are the following :

$$\begin{cases} \frac{\hat{p}_{jk}}{1-\hat{p}_{jk}} \quad if \quad E_{jk} = 0 \\ 1 \quad if \quad E_{jk} = 1 \end{cases}$$

That is, among non-connected households, the household has a weight which is proportional to its predicted probability to be connected, if he had the chance the be connected. Our reweighting approach share some similarity with Lenz et al. (2017) who model household connection decision using the sample of connected and non-connected households in grid catchment area, which they apply to form a control group in the sample outside the grid catchment area of households with relatively high predicted probability to connect if they had the choice of doing so. Our strategy differs, since we use the whole sample of households outside electrified villages (including those with low probability to connect), which is useful to predict impact on non-connected in electrified villages.

In practice, we only keep households that respect the common support condition, and proceed with trimming with a rejection window of -0.03 to avoid very high propensity scores that would affect the regression through the weights. Propensity score estimation details are presented in the appendix.

4.2 Identifying direct and indirect treatment effects

We need to confront two main intertwined identification challenges: a treatment that occurs at a multi-level and the subsequent multi-level endogeneity issues. A household may be treated because it is itself connected to the grid, or because its neighbors are connected to it, i.e because its village has access to the grid.

The issue is typically overlooked in the literature. In general the two effects are not distinguished and simultaneously evaluated. When authors control for spillover effects, they do not identify and estimate it but make sure their estimate is not ”biased” by this spillover effect, for example see (Bensch et al., 2011). We are in the situation of nested data structure with multi-level treatment. For a household to be supplied by the grid, and therefore be electrified, her village needs to be electrified. But a household can remain non-electrified while her village is connected to the grid. In that case, is that household different with respect to labor market outcome or education, from a household in a non-electrified village, all other things being equal? We expect the answer to be positive. Indeed, the labor market effect of electrification can be separated into a labor demand effect that must arise at the village-level - neighbors in the village get electrified, business develop and local labor demand rises - and a labor supply effect that arises at the household level : thanks to lighting the members have more potential light time available and they may supply more hours of work. We can actually distinguish three treatments effects, as shown in Table 4: (1) Relatively to household in non-electrified villages, electrified households in electrified villages (first line) benefit from both, households and village level treatment effects; (2) Relatively to non-electrified households in electrified villages, electrified household benefit only from household level treatment effect (second line) - indeed in connected villages all households take benefit from spillover effects so that, in principle these two groups of household differ with respect to household level treatment effect. (3) Finally, relatively to households in non-connected villages non-electrified households within electrified areas benefit only from village level spillover effects (line 3);

This nested treatment structure magnifies identification issues we will highlight hereafter. In line three, we may conclude that labor demand is higher in connected localities just because of selective program placement - authorities choose more dynamic villages, villages already endowed with better infrastructure for grid expansion. In line two, we may conclude that labor supply is higher in connected households just because those households have preferences that make them both working more and more likely to connect, or simply because, given their other observable characteristics they have higher wage potential and/or lower reservation wages.

Table 4: Double treatment issue

Treated	Control	Variation level
$H = 1$ and $V = 1$	$H = 0$ and $V = 0$	village and household
$H = 1$ and $V = 1$	$H = 0$ and $V = 1$	household
$H = 0$ and $V = 1$	$H = 0$ and $V = 0$	village

Note. H denote the household level treatment status (=1 if household is electrified), V denotes the village level treatment status (=1 if the village is connected to the grid)

Therefore, in practice, we need to find two instruments: one variable that randomizes, with respect to labor and education outcomes, village access probability, and one instrument that randomizes household connection probability conditional on being in a connected village. The ideal case would be to have two such instruments that allow us to investigate spillover effects in all our outcomes. In our case, we have three potential instruments. One is the distance from household to the main road. This has been used in the literature already (Salmon and Tanguy, 2016). Nevertheless, for two reasons we do not use it. First, it is not very clear at which level it purges endogeneity. It is clear that if the village is on average far from the road, the probability of village access is low because of grid expansion cost, the grid being usually located around the road. It is less clear that it deals with household self-selection into treatment: is the remoteness of the household, given village access, a strong determinant of the connection? However the main drawback in using such an instrument is the potential violation of the exclusion restriction. Indeed, it is likely that villages and households within village closer to a road face a more dynamic local labor market and business opportunities. In general, households located close to the road open all sort of business activities (phone services, barbers, food services, reparations, etc.) and are more likely to meet with potential customers and engage in trading. In our context we better view distance to the road as a control variable rather than an exogenous determinant of connection.

To instrument village access, we choose to use an instrument that is based on the same reasoning as in Dinkelman (2011). She instruments village connection with average land gradient, expecting that a high gradient increases village connection cost. In Rwanda, such a reasoning is relevant: the average land gradient in the country is especially high (see 3). In particular, some districts have a very large average altitude compared to other districts, especially those in the North-West, compared to those in the East. Nevertheless, we do not have the geographical coordinates of households or villages: we only know in which district they are. We need to rely on a proxy for land gradient. We have one information, which is the main source of water each household is using, within a village. We notice that many of them use spring water, instead of private or public pipe. In particular, there are some villages where all households are using spring water. We expect this to indicate to some extent a probability that this village is located in a high altitude and subsequently, it is a proxy for land elevation and connection cost.[11] Figure 3 gives support to this idea: in the districts with high average elevation, universal spring water usage villages are more frequent than in lower altitude districts. One natural critique that could be raised against the choice of this instrument, is that households using spring water may be just a consequence of the absence of infrastructures allowing for pipe water usage, itself related to the village dynamism and hence the employment rate. Hence we represent in the map the male wage employment rate in the rural areas of the districts, and we do not observe systematic lower employment rates in high altitude districts. Nevertheless, we still control further for district fixed effects, that are a first level of location-specific economic strength, and the few village characteristics we can have access to, which are distance to main road, and distance to administrative cell office. Note that Kigali's district has the highest employment rate, because it's the capital city. As of the instrument's relevance, it turns out that it is a quite strong predictor, all other household characteristics being held constant, of village access to electricity. Detailed evidence of this is provided in table 12 in the appendix.

[11] Rwanda receives an average annual precipitation of 1200 mm. Rainfall ranges from as low as 700 mm in the Eastern Province to about 2000 mm in the high altitude north and west (Republic of Rwanda, Ministry of Natural Resources, 2015).

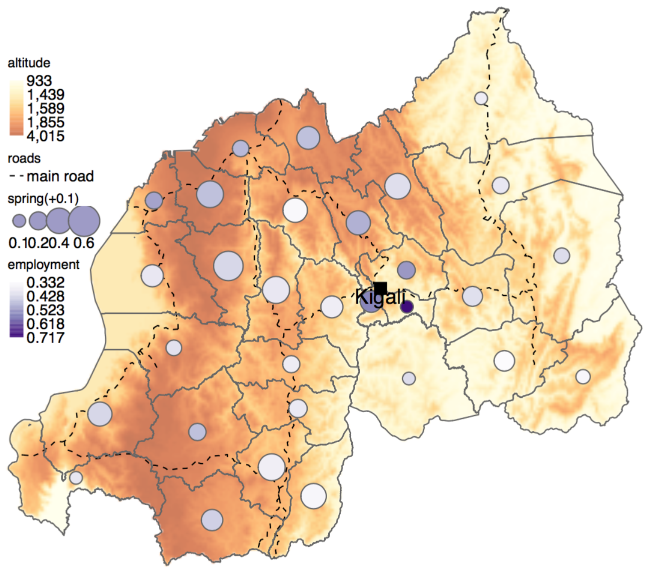

Figure 3: Map of Rwanda districts. *Employment* variable refers to the weighted male wage employment rate in rural areas in a given district, computed with the EICV4 data. *Spring* variable refers to the weighted proportion of villages where all households use spring water, among rural villages, in a given district. The proportions are augmented by 0.1 in order to observe districts where real proportions are zero. Kigali is the capital city of Rwanda. Dashed lines are main roads.

To instrument for household level connection treatment status, we choose to use the village connection status as an instrument. However, we do so when we regress outcomes for which we suspect particular household-level endogeneity, but conditionally on control variables, no village-level endogeneity. More precisely, we rule out first order spillover effects when considering weekly work hours outcomes (intensive labor impact margins) and school repetition. Indeed, we do not expect that, *conditional on working or attending school*, the village connection status is correlated with how many hours are supplied or with school repetition. To enforce the credibility of this assumption we control for village remoteness through the household distance to road, as well as for full set of district fixed effects which are arguably exogenous to household level characteristics.

More formally, let's $Y_{ijk}^{h,r}$ be the outcomes of interest representing the positive quantity of hours worked in various type of labor, or grade repetition. Let Y_{ijk}^{p} be the participation outcomes equal to 1 if the individual participates to various kind of labor or attends school. Then the equations to estimate are the following

$$Y_{ijk}^{h,r} = \beta_0 + \beta_1 X_{ijk} + \beta_2 H_{jk} + \tau E_{jk} + FE_d + \mu_{jk} + \sigma_{ijk}$$

$$Y_{ijk}^{p} = \phi_0 + \phi_1 X_{ijk} + \phi_2 H_{jk} + \tau E_{jk} + \nu V_k + FE_d + \eta_k + \epsilon_{ijk}$$

$$W_{ijk} = \begin{cases} \frac{\hat{p}}{1-\hat{p}} & \text{if} \quad E_{jk} = 0 \\ 1 \quad \text{if} \quad E_{jk} = 1 \end{cases}$$

X_{ijk} is a vector of individual characteristics, H_{jk} is a vector of household characteristics, V_k is a vector of village level characteristics and W_{ijk} is the estimated weight of individual i in the regression. μ_{jk} contains household unobservables and η_k village unobservables. ϵ_{ijk} and σ_{ijk} are the unobservable outcome determinants.

To sum-up, we estimate 4 models. First, we use simple OLS estimations using the provided sampling weights. The OLS estimates provide conditional mean difference between treated household at different levels, and different control groups. In order to insure that all treatment effects are estimated taking a reference control group with the same characteristics we apply weights W_{ijk}. In that case we reweight our sample such that, whatever their treatment status, households have on average the same characteristics as those living in electrified dwellings. This makes our results more comparable across different treatment status. Third, we use IV 2SLS estimation. First, in equations relating intensive margin labor and school quality equation to household connection, E_{jk} is instrumented using V_k. In the second set of equations, the participation equations, V_k is instrumented using a variable at the village level indicating whether or not all households in the village use spring water, that we denote S_k. Fourth, we also consider an IPW version of our IV-2SLS taking as a reference group, in the reweighting strategy, the electrified households.

4.3 Sample selection under endogenous electrification

One issue arising when regressing the weekly hours of work is that we observe this variable only for the sample of women who actually work. We expect that the participation decision is affected by outcomes systematically related to work hours. This can lead to a *selection bias* of the coefficient estimate on hours. We tackle this issue using a heckman sample selection procedure under endogeneity of the electricity dummy, and implement it as it is presented in Woolridge (2002).

In a first stage, we model labor participation decision with a probit, including both the regressors present in the work hours equation, and variables excluded from the work hours decision. Let X_{ij} be the vector of variables included in the work hours equation, except the treatment E_{jk}, but including the instrument S_k. Let Z_{ij} be a vector of excluded variables determining the participation but not the hours. The first stage equation is :

$$P[Y_{ij}^p = 1 | X_{ij}, Z_{ij}] = \Phi(\alpha X_{ij} + \beta Z_{ij})$$

That we estimate using all female spouses sample to compute the inverse mills ratio $\lambda(\cdot) \equiv \frac{\phi(\cdot)}{\Phi(\cdot)}$:

$$\hat{\lambda}_{ij} = \lambda(\hat{\alpha} X_{ij} + \hat{\beta} Z_{ij})$$

for all female spouses. We include this ratio in the work hours equation as specified in previous subsection, omitting Z_{ij}.

In practice, the challenge is to find a good vector Z_{ij}: some female spouse and household characteristics that we expect to impact participation decision but that we can omit from work hours equation with confidence. Inspired by Woolridge (2002) we use three variables : the number of kids in very young age, between 0 and 6 years old, the total number of kids, and the number of adults and old persons in the household. If a women has to take care of young kids, she is less likely to participate. If other adults are there to take care of them, this is less likely to be true. In addition, we do not expect these variables to affect work hours *conditional on working*. We compute $\hat{\lambda}_{ij}$ for all activities for which we estimate the effect on work hours: *wage worker*, *unpaid employment*, *independent business*.

5 Results

5.1 Women and children

Our first set of results concern children. We focus on employment participation in one hand, and on educational outcomes on the other hand, measured by attendance and grade repetition. Results are presented in Table 5. In all subsequent estimations, though they are not presented to lighten the exposition, all regressions include an extended set of control variables presented in the descriptive statistics section. In particular, we control for the household's head socio-demographic characteristics, for household wealth, for district fixed effects, and household distance to the closest road. The OLS estimations shows a negative effect of electrification at household level on labor market participation, which is large but not statistically significant when we instrument for village level electrification. The effect at the village level is rather positive, in IV estimations, but not statistically significantly different from zero. The signs of the coefficients (positive for the household level and negative for village level effect) suggest that spillover effect, without the benefits of household access to electricity, may increase child labor. However at that stage we lack sufficient power to be conclusive. We also investigate the impact on schooling outcomes. We find a negative effect on attendance for the village level treatment effect in IV estimations. However, this effect seems to be specific to the sample characteristics of non-connected households. Indeed, combining our IV estimation with a reweighed approach in order to make our three samples comparable across treatment status, the effect on attendance turns out to be small and no statistically significant. Lately, we found that the effect on the quality of education acquired, as measured by grade repetition is little affected by household level electrification. Our data allows us to distinguish between paid and unpaid employment. Detailed results that distinguish boys and girls are presented in next section.

In Table 6 we turn to women employment outcomes. We first consider women engagement in paid and unpaid employment ; this distinction is central to assess the potential of electrification for empowerment. We distinguish between impacts occurring at the extensive margin (columns 1 to 4), i.e. the participation decision, and those occurring at the intensive margins (columns (5) to (8)) i.e., conditional on participation, how many work hours are supplied. As before we rule out spillover effect at the intensive margin but not at the extensive margin. Therefore at the extensive margin we distinguish between direct effect of household electrification an indirect effect occurring through spillover effect at the village level, on non-connected household. As in the previous table we take into account that the village electrification status may be endogenous to village level unobserved determinants of participation. Therefore we also present results where village electrification is instrumented using intensity of spring water usage which is a proxy for village elevation. Considering first the impact on participation into paid employment, in general, whatever the estimation method, we find that the impact on participation into paid employment is rather small. Indeed, conditioning on a full set of control variables does not substantially change the results we already obtained using simple OLS regressions in Table 3. This results again highlight the fact that on one hand non electrified households are very similar whatever their village treatment status. Remember that village level effect in this specification is measuring any spillover effect that will benefit non-electrified household in connected villages. Therefore we do not find evidence of such an effect. On the other hand, the fact that there is no effect of electrification at the household level suggests that difference across household within connected villages is not an important source of selectivity bias in our context, which supports our assumption of exogenous electrification conditional on the village being connected. When we constrain the whole sample to have the same characteristics as those of electrified households, the village level effect turns to be large and positive in IV estimation, which is however not statistically significant. Additionally, the effect of electricity access at the household level is also low and non statistically different from zero. Here again, controlling for household characteristics does not alter the result relatively to those already obtained in the simple regressions (see Table 3). This suggests there is probably little selectivity into the decision to connect that may be related to unobserved household level determinants to participate into paid employment.

Unlike for participation, impact on women materializes sharply and more precisely at the intensive margin. Given previous results, we assume that spillover effects occurring at the village level are negligible, and consider village access to electricity, conditional on household connection status and participation decision, as exogenous to unobserved determinants of women labor supply. This allows us to use village access as an excluded instrument for household decision to connect. Unlike for participation we found significant, positive and sizable effect of household level connection on hours worked for a paid activity. On average and taking into account endogenous connection choices, women in connected households supply on average 5 additional hours into paid employment, this effect size is in general higher when we instrument for electricity access. This suggests that, contrary to conventional wisdom, there is no evidence of positive selection of household into the decision to connect with respect to unobserved determinants of labor supply. The impact estimated also depends on the specific characteristics of the control group, it is somehow lower when we constraint the impact to be estimated on a control group having similar characteristics as those of electrified household (column 8 vs column 7) using re-weighted IV estimation. However, given the size of the estimated standard errors, the difference is not statistically significant.

What are the sources of this additional time supplied into paid employment? In particular: did these additional hours come from a lower number of hours supplied into unpaid employment activities? Indeed, and unlike for men, unpaid employment is a pervasive phenomena among women in rural Rwanda; where a majority of them is engaged into unpaid employment (see Table 2). Results provided in Panel B of Table 6 shows that there is virtually no effect on unpaid labor, the estimated impact is small, negative and non statistically significantly different from zero. Presumably, the additional time supplied in the labor market for paid employment does not seem to have lowered women labor supply to unpaid employment. Therefore, this should come from other sources of time gain such as a more efficient used of time due to lower constraint on availability of light time.

Women may engage into paid employment as a wage worker, or as an independent worker into various small business activities. In Table 7 we distinguish these two activities since they may reveal different mechanism through which women economic empowerment is affected by access to electricity. We present results on the extensive margin for which we find statistically significant effects[12]. Household access to electricity seem to trigger more labor supply in both independent businesses and wage labor activities. Surprisingly, the effect on wage activities is larger. We expect the effect to be small because of limited spillover effects and therefore limited effects on village level labor demand channel. This effect suggests that there is potentially a demand for labor that does not meet a supply due to household time constraint. What these result suggest is that it is presumably women that are already engaged in independent and wage labor activities that experience an increase in their labor supply. We also not that controlling for household characteristics to match those of treated households is important in particular for labor supply of workers, this underline some important heterogeneity of impact across households. In the next section we dig further into this heterogeneity, focusing on the impact of electricity on economic empowerment across genders and within household couples.

Table 5: Effect of electrification on children

	Labor participation				School attendance			
	OLS	OLS^{IPW}	IV	IV^{IPW}	OLS	OLS^{IPW}	IV	IV^{IPW}
$E^{village}$	-0.0105	-0.0250**	0.0313	0.134	0.0111	0.0183	-0.136*	0.00176
	(0.00649)	(0.0110)	(0.0714)	(0.110)	(0.00815)	(0.0105)	(0.0820)	(0.103)
$E^{household}$	-0.0267***	-0.00134	-0.0554	-0.116	0.0240**	-0.0103	0.124**	0.00167
	(0.0102)	(0.0119)	(0.0498)	(0.0795)	(0.0116)	(0.0118)	(0.0575)	(0.0739)
N	10890	10890	10890	10890	10890	10890	10890	10890
F_stat			109.0	101.8			109.0	101.8
	Grade repetition							
$E^{household}$	-0.0118	0.0325	0.0129	0.0240				
	(0.0167)	(0.0176)	(0.0374)	(0.0327)				
N	7384	7384	7384	7384				
F_stat			1090.9	967.7				

Statistical significance are * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors are reported in parenthesis under coefficient estimates. In all estimations, we include the following control variables: literacy, education, age, age square of household head, child's gender, age and age square, household composition, wealth proxies (consumption deciles dummies, dwelling construction material, land ownership), distance to the closest road, and a full set of district fixed effects. In columns label "OLS" survey sampling weight is used in OLS estimations; in columns label OLS^{IPW} we use the same set of control variable in an IPW regression using the estimated propensity score. See Appendix B.1 for a detail of the variables included in the estimation of the propensity score. The treatment variables $E^{village}$ and $E^{household}$ are a binary variables taking respectively the value of one if the village is connected to the grid, and the value of one if the household dwelling is connected.

[12] As in Table 6 results on the extensive margin are not consistent across estimation methods and in all cases coefficient estimates display large standard errors

5.2 Within household Spouse Differences and differences across child gender

5.2.1 Differences across boys and girls

As in many countries, inequalities across gender are prevalent early on in life and particularly during childhood. In developing countries rural areas, these inequalities also materialize with respect to time use of children inside and outside household and family activities. We saw that electricity has triggered more employment into independent business. These are mostly family held farm, labor intensive activities. In this context children may provide flexible and cheap labor. In Table 8 we present results concerning the effects of electricity access at the village level and household level on participation into paid and unpaid employment separately for girls and boys. For the sake of brevity we only present IV estimations. As before, the village status into treatment is instrumented, and we assume, for a given village electrification status, that household electrification status is uncorrelated with unobserved determinants of children employment. We find contrasting impact across genders and levels of treatment. Among boys, access at the village level is associated with higher employment into paid activity. This positive effect is compensated by a negative effect, of similar magnitude, when the household has access to electricity. Overall, adding the two opposite effects explains that household access to electricity has little effect on boys employment. For girls, the direction of impact on paid employment goes in opposite direction: a higher propensity to work for household level treatment and a lower propensity for the village level treatment. Estimates for girls are nevertheless non significant. Instead, we find for girls greater propensity to engage in unpaid employment once the village is connected to the grid. These contrasting impacts for girls and boys underline different roles of girls and boys within household. It seems that greater employment opportunity triggered by village access to electricity pulls boys into employment, out of family business, while girls are rather pulled into activities, mostly farm, generated within households ; However we acknowledge that these interpretations at that stage remain speculative, and should deserve further investigations. Another explanation could be that boys and girls employment are substitutes and interrelated within households. However, and against this explanation, results of Table 8 do not reveal any negative impact on boys involvement in unpaid work.

5.2.2 Differences between spouses

Does access to electricity increase the relative position of women within household. Our theoretical approach to the issue suggest that access to electricity will particularly benefit women since they are more involved within household in domestic production and unpaid family work. Conversely, electricity may have widespread effect that may also affect men employment opportunities. For instance if men wage potential and employment opportunities raises more than that of women. Women empowerment in relative term may not increase, despite greater involvement in paid employment. To pursue further this issue, let's Y_{sh} denote an outcome of interest for spouse member s of household h. Where s=1 for male and 0 for women. As previously we consider employment outcomes at the extensive and intensive margins: i.e., participation and hours of work, conditional on women participation, into a paid or an unpaid activity. We form the following dependent variable within each household: $Y_{sh} = Y_{1h} - Y_{0h}$ which denote the gap in outcome between spouses within household. An increase in Y_{sh} represents an

Table 6: Effect of electrification on Women

	A: Paid labor							
	Extensive margin				Intensive margin (hours)			
	OLS	OLS^{IPW}	IV	IV^{IPW}	OLS	OLS^{IPW}	IV	IV^{IPW}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
E^{village}	0.0412**	0.0572*	0.0177	0.161				
	(0.0163)	(0.0318)	(0.139)	(0.223)				
E^{household}	0.0227	0.0165	0.0384	-0.0562	4.388***	2.744**	7.465***	5.113**
	(0.0272)	(0.0358)	(0.0972)	(0.158)	(1.201)	(1.302)	(2.368)	(2.319)
IMR					-10.71	-32.20**	-10.61	-31.29**
					(8.043)	(14.34)	(7.945)	(13.96)
N	6075	6075	6075	6075	2734	2734	2734	2734
F_stat			96.25	66.13			357.0	373.0
	B: Unpaid labor							
	Extensive margin				Intensive margin (hours)			
E^{village}	-0.0674***	-0.0570	-0.0542	-0.0910				
	(0.0155)	(0.0295)	(0.126)	(0.227)				
E^{household}	-0.0677**	-0.0726*	-0.0765	-0.0489	0.273	1.810	-1.692	-0.890
	(0.0258)	(0.0342)	(0.0883)	(0.160)	(1.058)	(1.116)	(1.971)	(2.147)
IMR					-21.00***	13.89	-21.58***	13.66
					(6.260)	(13.02)	(6.238)	(12.94)
N	6075	6075	6075	6075	3867	3867	3867	3867
F_stat			96.25	66.13			376.1	346.8

Statistical significance are * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors are reported in parenthesis under coefficient estimates. In all estimations, we include the following control variables: literacy, education, age, age square of household head; age and age square, education and literacy of women, household composition, wealth proxies (consumption deciles dummies, dwelling construction material, land ownership), distance to the closest road, and a full set of district fixed effects. In columns label "OLS" survey sampling weight is used in OLS estimations; in columns label OLS^{IPW} and IV^{IPW} we use the same set of control variable in an IPW OLS and 2SLS regression using the estimated propensity score. In columns (3) and (4) we use 2SLS regression where E^{village} is instrumented with a variable taking the value of one if all household in the village report using spring water as main source of water. In columns (7) and (8) we run a 2SLS regression where E^{household} is instrumented with a variable taking the value of one if the village is connected to the grid. The row label IMR indicate the coefficient associated to the inverse mills ratio estimated from the participation equation. The excluded variable in the participation equation are : the number of kids in very young age, between 0 and 6 years old, the total number of kids, and the number of adults and old persons in the household (see Appendix B.3). See Table 5 for additional notes. The row label F_stat refers to the first stage F-statistics on the excluded instrument.

Table 7: Effect of electrification on women wage labor and independent business labor supply

	(1)	(2)	(3)	(4)
	Independent business		Wage labor	
	IV	IV^{IPW}	IV	IV^{IPW}
$E^{household}$	5.399**	6.728***	21.59***	10.57**
	(2.459)	(2.204)	(5.647)	(4.228)
IMR	-9.156	-9.888	-4.395	-11.02
	(8.391)	(12.38)	(4.430)	(7.714)
N	1829	1829	1202	1202
First stage F_stat	304.8	363.9	82.29	94.38

Statistical significance are * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors are reported in parenthesis under coefficient estimates. See notes of Table 6.

Table 8: Effect of electrification on boys and girls paid and unpaid employment

	Girls				Boys			
	Paid		Unpaid		Paid		Unpaid	
	IV	IV^{IPW}	IV	IV^{IPW}	IV	IV^{IPW}	IV	IV^{IPW}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$E^{village}$	0.0161	-0.0815	-0.00171	0.197*	0.134*	0.128**	-0.0238	0.131
	(0.0634)	(0.176)	(0.0914)	(0.119)	(0.0760)	(0.0532)	(0.0939)	(0.110)
$E^{household}$	-0.0208	0.0420	-0.00605	-0.142	-0.124**	-0.110***	-0.0110	-0.104
	(0.0448)	(0.121)	(0.0650)	(0.0867)	(0.0516)	(0.0394)	(0.0641)	(0.0807)
N	5461	5461	5461	5461	5429	5429	5429	5429
First stage F_stat	61.73	47.34	61.73	47.34	47.27	51.26	47.27	51.26
N	6075	6075	6075	6075	2734	2734	2734	2734
F_stat			96.25	66.13			357.0	373.0

Statistical significance are * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors are reported in parenthesis under coefficient estimates. See notes of Table 5.

improvement in women relative empowerment within household. Considering differences of outcome within households has also another advantage in terms of identification. Indeed, one may assume that each spouses outcome each depends on unobserved common household and village fixed effects. First differencing allow swiping out these two unobserved heterogeneity and therefore greatly relaxes our identification assumptions. Indeed, this specification requires more credible conditional independence assumptions, allowing us to put a greater confidence on standard adjustment methods using either OLS estimations or propensity score reweighted approaches. Specifically, village level electrification, i.e. selective program placement, should not be systematically correlated with unobserved individual differences between spouses that may affect their relative propensity to work. Additionally, it requires that household level decision to connect (conditional on control variables) is not correlated with unobserved individual differences between spouses that may affect their relative employment propensity. Results for spouses are presented in Table 9. Considering paid employment as a whole, participation into paid employment is relatively higher among women in electrified villages. This within household beneficial effect is mostly due to village level treatment, that is to positive spillover effects of electrification across households. In details, women specialize in independent work within household where their employment relative to their spouse increases by almost 9% (column 4), once the village is connected to the grid. Instead, men get more specialize in wage jobs outside family activity (column 6). However looking at the intensive margins, women, tend to work relatively less hours as independent, and relatively more hours as a wage worker. Overall results of Table 9 are suggestive that women economic power within the household improve as a consequence of electricity. However most of improvement occurs within family businesses; yet the few women that are working as wage worker outside the household, are also working more intensively.

Table 9: Effect of electrification on spouses labor participation and supplied hours inequality

	Extensive margin					
	OLS	OLS^{IPW}	OLS	OLS^{IPW}	OLS	OLS^{IPW}
	Δ paid	Δ paid	Δ independent	Δ independent	Δ wage	Δ wage
	(1)	(2)	(3)	(4)	(5)	(6)
$E^{village}$	-0.0386**	-0.0643*	-0.0863***	-0.108***	0.0327*	0.0455
	(0.0194)	(0.0341)	(0.0228)	(0.0401)	(0.0189)	(0.0300)
$E^{household}$	-0.0727**	-0.0339	-0.0644*	-0.0369	0.00124	0.0380
	(0.0324)	(0.0391)	(0.0388)	(0.0479)	(0.0297)	(0.0358)
Observations	6149	6149	6149	6149	6149	6149
	Intensive margin					
	Δ paid	Δ paid				
	(1)	(2)				
$E^{village}$	-0.112	0.509				
	(0.827)	(1.649)				
$E^{household}$	-1.067	-0.404				
	(1.612)	(1.937)				
Observations	6149	6149				

Note. We consider, at the household level, differences between male and female spouse in terms of binary participation to paid labor and supplied hours to paid labor. For each household at the extensive margin Δpaid $\in \{-1, 0, 1\}$, i.e either the women works but not the man, both work (or do not work), or the man works but not the women. At the intensive margin Δpaid is just the difference in hours supplied to any paid labor, independent or wage. We do not condition this variable on participation of man or women, because most of the men do so (see Table 2) so there is little issue of zero-hours probability mass. Each pair of columns represent OLS and reweighted OLS models for a specific dependent variable, where controls are applied as in the previous main results. Robust standad errors in parenthesis below coefficients estimates. Significance levels are * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 Conclusion

Grid coverage is expanding rapidly across Sahara Africa. Is this expansion a vehicle of children and women empowerment? We consider the case of Rwanda where most households are too poor to buy time-saving appliances, and we postulate that extended active daytime with electric-grid based light could affect education, work, home production and leisure, effectively empowering women and children. We highlight that impact of electricity should be assessed at the village level economy since non-connected households may benefit from local spillover, this is important in a context where electricity take-up in grid catchment areas is far from being universal. Though not unequivocal, our results reveal that access to electricity allows connected households to diversify their income sources as women are supplying more hours into paid job. Children and especially boys are also more likely to engage in paid employment. We also found limited evidence of spillover effects. When they are present, children are the most affected: boys being pulled into paid employment while girls are getting more specialized in unpaid employment. Overall empowerment potential of rural electrification is essentially concentrated among the small share of women in rural areas which are more likely to already engaged in paid employment, access to electricity at the household level allows them to extend their weekly working hours substantially. Therefore, to reach its full empowerment potential, access to electricity will need to generate changes in the village economy which will provide sufficient incentives and resources for women to engage in paid employment (intensive margin) among both connected and non-connected households. Such changes could be activated through targeted and complementary policies that will ease the business environment of households' farm and non-farm micro-enterprises (access to credit, markets, training, technology, etc.), in order to increase households' returns from making a productive usage of electricity. The challenge will be to dampen the trade-off among children between working and studying. Indeed, our results show that children labor supply is the most sensitive to local employment spillover effects. However, in light of our results such a threat seems to be limited. Notwithstanding our results, we are aware that lacking a pure natural experiment generating a strong credible research design, several of our conclusions at that stage are still tentative and does provide a definitive causal answer regarding the effect of energy and women empowerment. Most of our conclusions still need to be confirmed with a stronger and more credible research design.[13] Lately, we should stress that beyond identifying and musing causal impact of RE, understanding the channel of transmissions is a priority in order to design public policy that are complementary to RE. This will require a better understanding of the village economies in rural areas, and how it interacts with household level choices to electrify and eventually to make a productive use of electricity. We are confident that this will make a promising research agenda that we look forward to pursue.

References

Philippe Alby, Jean-Jacques Dethier, and Stephane Straub. Firms operating under electricity constraints in developing countries. *World Bank Economic Review*, 2013.

[13] In particular, it is well known that even though our instrument were valid, it will identify local average treatment effect, this effect could be different for treatment effect on the treated.

Oriana Bandiera, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. Labor Markets and Poverty in Village Economies. *The Quarterly Journal of Economics*, 2017.

Gunther Bensch, Jochen Kluve, and Jorg Peters. Impacts of Rural Electrification in Rwanda. *Ruhr Economic Papers*, 2011.

Tanguy Bernard. Impact Analysis of Rural Electrification Projects in Sub-Saharan Africa. *World Bank Research Observer*, 2012.

Matias Busso, John DiNardo, and Justin McCrary. New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators. *The Review of Economics and Statistics*, 2014.

Marco Caliendo and Sabine Kopeinig. Some Practical Guidance for the Implementation of Propensity Score Matching. Iza discussion papers, Institute for the Study of Labor (IZA), 2005.

Taryn Dinkelman. The Effects of Rural Electrification on Employment: New Evidence from South Africa. *American Economic Review*, 2011.

Matthias Doepke and Michèle Tertilt. Does female empowerment promote economic development? Working paper, National Bureau of Economic Research, 2018.

Esther Duflo. Women empowerment and economic development. *Journal of Economic Literature*, 2012.

Louise Grogan and Asha Sadanand. Electrification and the household, 2009.

Guido M Imbens and Jeffrey M Wooldridge. Recent developments in the econometrics of program evaluation. Working Paper 14251, National Bureau of Economic Research, 2008.

Raul Jimenez. Development effects of rural electrification. Policy brief, InterAmerican Development Bank, 2017.

Charles Kirubi, Arne Jacobson, Daniel M. Kammen, and Andrew Mills. Community-Based Electric Micro-Grids Can Contribute to Rural Development: Evidence from Kenya. *World Development*, 2009.

Gunnar Kohlin, Erin O. Sills, Subhrendu K. Pattanayak, and Christopher Wilfong. Energy, gender and development: what are the linkages ? where is the evidence ? Policy research working paper series, The World Bank, 2011.

Kenneth Lee, Edward Miguel, and Catherine Wolfram. Electrification and economic development : A microeconomic perspective. *Oxford Policy Management*, 2017.

Kenneth Lee, Edward Miguel, and Catherine Wolfram. Experimental Evidence on the Demand for and Costs of Rural Electrification. *NBER Working Papers*, 2018.

Luciane Lenz, Anicet Munyehirwe, Jörg Peters, and Maximiliane Sievert. Does large-scale infrastructure investment alleviate poverty? impacts of rwanda's electricity access roll-out program. *World Development*, 2017.

Juan Pablo Rud. Electricity Provision and Industrial Development : Evidence from India. *Journal of Development Economics*, 2012.

Claire Salmon and Jeremy Tanguy. Rural Electrification and Household Labor Supply: Evidence from Nigeria. *World Development*, 2016.

Tim Squires. The impact of access to electricity on education: Evidence from honduras. *Job market paper*, 2015.

Catherine Wolfram, Orie Shelef, and Paul Gertler. How Will Energy Demand Develop in the Developing World? *Journal of Economic Perspectives*, 2012.

J. Woolridge. *Econometric analysis of cross section and panel data*. MIT press, 2002.

A Additional results

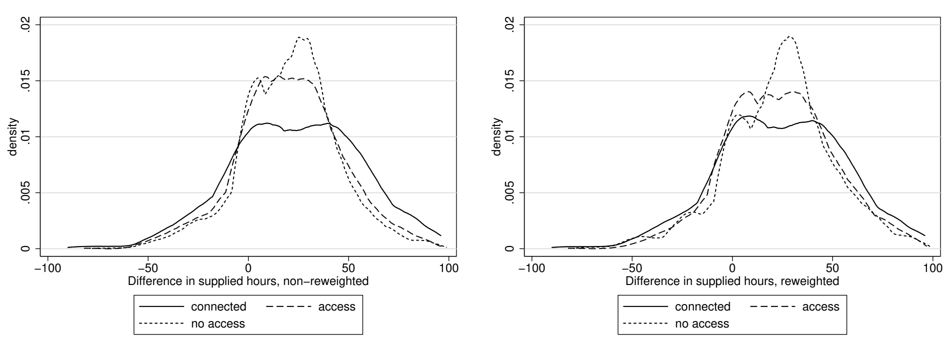

Figure 4: Distribution of the difference in hours supplied to any paid labor between the male and female spouse, before reweighting (LHS) and after reweighting (RHS).

Table 10: Household chores hours

	Children					
	Access to the grid				No access to the grid	
	Connected		Not Connected			
	(1)		(2)		(3)	
	Mean	N	Mean	N	Mean	N
water	4.16	724	4.21	2078	4.32	7109
wood	3.54	276	4.39	1193	4.39	4046
fodder	4.65	414	5.63	1151	6.48	4172
market	1.24	180	1.57	341	1.92	933
cooking	6.19	415	6.10	1094	6.30	3566
N	886		2258		7683	
Women						
water	2.96	278	2.81	862	2.88	3091
wood	2.86	173	3.33	771	3.54	2678
fodder	6.14	281	6.89	775	7.06	3171
market	2.85	476	2.69	1018	2.98	3119
cooking	14.49	569	14.69	1348	15.21	4633
N	583		1382		4721	

	Children				
	(1)	(2)	(3)	(4)	(5)
	water	wood	fodder	market	cooking
$E^{household}$	-1.554***	1.498***	-1.096**	0.128**	-2.099***
	(0.00480)	(0.000180)	(0.431)	(0.0498)	(0.00208)
$E^{village}$	3.510***	0.919***	8.707***	0.105***	0.952***
	(0.293)	(0.118)	(0.254)	(0.0304)	(0.0201)
N	9833	5463	5695	1443	5046
Women					
$E^{household}$	-0.0321	0.689***	-0.0928	-0.678***	-2.929***
	(0.0778)	(0.0540)	(0.0794)	(0.0298)	(0.0205)
$E^{village}$	-1.157***	-0.647***	-1.926***	0.383***	2.095***
	(0.0703)	(0.0155)	(0.0171)	(0.00365)	(0.0149)
N	4200	3598	4202	4583	6507

Note. The left hand side of the table reports average number of hours spent in each chore among children and women, across connection status, as in 1. Sampling weights are used. In the right hand side, each column reports a regression model where the dependent variable is the number of hours spent in a given chore activity. In each model household and available village characteristics are controlled for along with district fixed effects. Robust standard errors are in parenthesis and usual critical probabilities thresholds. Water stands for fetching water. Wood for wood foraging, fodder searching, going to market, and cooking respectively.

B Additional details on estimation procedures

B.1 Propensity scores estimation

We assume a logistic distribution of the household connection decision and we estimate this model on the sample of households in connected villages, who had the possibility to decide whether to connect or not. Using this model we predict the probabilities of connection on the whole sample, including households which did not have access to the grid and therefore no choice to connect.

When one wishes to weight regressions with the propensity scores, one should be careful both with the support condition and extreme probability values that can strongly affect estimates. To ensure that support condition is verified, we trim the estimated probabilities (Busso et al., 2014). We keep only observations for which the estimated probabilities lie in the following interval :

$$[min(\hat{P}^{E=1}), max(\hat{P}^{E=0})]$$

Where, for example, $min(\hat{P}^{E=1})$ indicates minimum of the estimated connection probability among households who decided to connect.

In table 11 we present the household level variables selected to be included in the household connection decision model. We also included district dummies. Finally figure 1 showed in section 3 distributions of predicted probabilities by connection status, before and after reweighting procedure. These figures show that the model catches well the prevalence of household-level selection and that our reweighting procedure balanced propensities well

Table 11: Variables included in propensity score estimation procedure

Household head	Composition	Wealth
Age (+sq) Literate Primary diploma Secondary diploma Always lived in dwelling	Size (+sq) Nb literate (+sq) Nb kids (+sq) Nb adult 50yo + Proportion of female	Water source Floor area (+sq) Nb rooms (+sq) Domestic worker Toilet type Nb health problems Owns land
+ District dummies		

B.2 IV 2SLS first stage

We report in table 12 the results of the IV 2SLS first stage when regressing women paid work binary participation outcome on household and village connection while instrumenting village connection with the binary variable *all spring* indicating if all households use spring water in the village. We present these results to support the relevance of our instrument: we expect that villages where all households use such as source of water are villages located in high altitudes, with a higher cost to electrify the village.

Table 12: First stage IV 2SLS on female spouse paid work participation

	(1) $E^{village}$	
all use spring water	-0.117***	(0.00622)
$E^{household}$	0.678***	(0.00611)
age	-0.000221	(0.000843)
age squared	0.00000574	(0.0000125)
is literate	0.000698	(0.00712)
primary sch. diploma	0.0133	(0.00883)
secondary sch. diploma	-0.0299**	(0.0132)
HH age	0.000164	(0.000341)
HH literate	-0.0000621	(0.00758)
HH secondary dip.	0.0149*	(0.00830)
nb kids 0-6 yo	0.000814	(0.00378)
nb kids	0.000256	(0.00190)
nb + 40 yo (extended)	0.00529	(0.00326)
consumption decile	-0.00244	(0.00229)
walls in cement	0.0340***	(0.00770)
owns land	0.112***	(0.0180)
distance to main road	-0.000000738*	(0.000000396)
N	24166	

We report in column (1) the first stage of the IV 2SLS estimation of the effect of electrification treatments while instrumenting village level electrification with a binary indicating whether all the households in the village where the spouse lives use springs as a main source of water. The instrument's coefficient is strongly significant. Covariates included district fixed effects that are not reported for clarity. Robust standard errors are reported in parenthesis. Significance levels are : * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B.3 Sample selection estimation procedure

Table 13: First-stage sample selection estimation procedure for female spouse *paid* dummy

	(1)
	Paid employment (extensive margin)
nb kids 0-6	-0.0284
nb kids	-0.0270*
nb persons +40 yo (extended)	0.0301
age	0.0865***
age squared	-0.00109***
is literate	-0.0271
owns land	0.399***
primary	0.344***
secondary	-0.263***
HH age	-0.00300
HH literate	-0.0219
HH secondary	0.122**
consumption decile	-0.0570***
walls in cement	0.0577
all spring water	-0.0216
N	6822

We provide an example in column (1) of the first stage of a sample selection correction procedure for labor participation of female spouses. A Probit model is used with a paid work participation dummy as independent variable. Excluded covariates are in the three first lines. District fixed effects are included in the controls, significant but omitted for clarity. The inverse mills ratio is then computed using predicted probabilities and used in last stage regressions. Robust standard errors in parenthesis. Significance levels are * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.