The Effects of Rural Electrification on Employment: New Evidence from South Africa

Taryn Dinkelman*
Princeton University

This version: December 2009 First version: November 2007

Abstract

This paper establishes that new access to public infrastructure affects both home production technologies and market employment in a developing country. Exploiting spatial and temporal variation in electricity project placement during South Africa's mass roll-out of rural household electricity, I estimate district fixed-effects models of fuel-use and employment growth, instrumenting for project placement using land gradient. Within five years, electrified areas substitute sharply towards electricity in home production and IV employment results are large, positive and significant for women. Complementary evidence using a different data set and an alternative identification strategy confirms increases in male and female hours of work and a decline in female wages in areas of expanding infrastructure. Rather than stimulating large-scale rural industrialization, the weight of evidence in this paper points towards rural electrification increasing labor supply and facilitating jobs in new, small-scale market-based services.

*Email: tdinkelm@princeton.edu. Electrification data and institutional background provided by Jenny Barnard, Sheila Brown, Ed Bunge, Jim Stephenson, Steven Tait, Innocent Nxele and Amos Zuma. Conversations with and comments from Lori Beaman, Diether Beuermann, Matias Busso, Anne Case, HwaJung Choi, John DiNardo, Ann Ferris, Raymond Fisman, Ben Keyes, David Lam, Jim Levinsohn, Claudia Martinez-A, Justin McCrary, Nancy Qian, Dori Posel, Samara Potter-Gunter, Mark Rosenzweig, Jeff Smith, Gary Solon, James Tybout, Rebecca Thornton and Martin Wittenberg substantially improved the paper. I also thank seminar participants at the following universities, organizations and conferences for their valuable input: Brown, California-Los Angeles, California-Riverside Dartmouth, Duke, the Harris School, the Harvard School of Public Health, the International Institute of Economic Studies at Stockholm University, the Kennedy School, KwaZulu-Natal, London School of Economics, Michigan, Michigan State, the Microeconomics of Growth Network, Northwestern, North Carolina-Chapel Hill, Penn State, Princeton, Resources for the Future, Syracuse, Washington-St Louis, Western Ontario and Yale. Data collection was funded through a National Science Foundation Doctoral Dissertation Grant and an Eva Mueller New Directions in Demography research award. I also acknowledge support from the Fogarty International Center of the U.S. National Institutes of Health (D43TW000657) and the Population Studies Center at the University of Michigan. All errors are my own.

1 Introduction

Electricity is pervasive in all industrialized countries and largely absent in the developing world. An estimated 1.6 billion people across the globe do not have access to electricity and 75 percent of Africans are without access (Saghir (2005), Sustainable Development Network (2007)). Over the next several decades, this situation is likely to change as commitments to energy infrastructure investments in poor countries are realized. Some of this planned expansion targets industry while other "Bottom of the Pyramid" initiatives focus on making cheap modern energy available to 250 million Africans by 2030.[1] Regardless of the form this expansion takes, we know very little about the effects that new access to modern energy will have on economic growth, health, human capital accumulation and labor market outcomes.

In this paper, I use the experience of rural electrification roll-out in post-apartheid South Africa to analyze the impact of new access to modern energy sources on an outcome of considerable interest: the ability of the poor to use their labor resources for market production. I focus both on estimating the direct impact of electrification on market employment and on understanding the mechanisms through which this new infrastructure affects employment. In many parts of the world, the ability of the poor to work in the market is constrained both by the demands of labor-intensive home production and by lack of market opportunities for work. Household electrification plausibly affects both of these constraints, allowing individuals in rural areas to substitute away from time-intensive methods of food preparation and shift labor time in to the market, as well as allowing them to engage in new market activities.[2] Using data from rural KwaZulu-Natal (KZN), this paper provides some of the first evidence in support of claims that new access to modern energy sources can increase employment in the market. Part of this employment effect appears to be driven by the time being released from home production in to market work, while the rest is likely driven by improvements in market opportunities for small-scale businesses facilitated by new access to infrastructure and by in-migration. None of my results suggest that electrification stimulates employment through large-scale rural industrialization.[3]

The roll-out of grid infrastructure in South Africa provides a good opportunity to evaluate the effects of household electrification on market employment and household fuel use. In 1993, one year

[1]World Bank commitments to energy infrastructure in sub-Saharan Africa rose from $447 million in 2001 to $790 million in 2007. Investments in renewable energy sources around the world amount to $3.3 billion in 2009. See also the World Bank's Lighting Africa initiative www.lightingafrica.org. The Bank has also committed substantial resources to energy development projects in Asian countries.

[2]Although time-use data are scarce across Africa, what evidence exists suggests that substantial labor resources are concentrated in home production. In South Africa in the mid-1990s, over 80% of households collected wood for household needs; over three-quarters of those collecting wood were women; and on average, these individuals spent the equivalent of two working days per week in fuel-wood collection (Budlender et al, 2002). Food preparation in rural areas consumes on average 3 hours per day. Data from South African's 1997 October Household Survey. See also Charmes (2005), Saghir (2005). and "The Energy Challenge for Achieving the Millenium Development Goals" at httE://esa.un.org/un-energy/pdf/UN-ENRG/20paper.pdf accessed 8 August 2006 for discussions of labor resources that could be released when switching from traditional to modern fuels.

[3]See Rud (2009) for an analysis of the effects of a different type of electrification on rural industrialization in India.

before the end of apartheid, over two-thirds of South African households were without electricity and more than 80% relied on wood for home production. Following the new government's commitment to universal electrification, 2 million (23% of all) households across the country were newly connected to the grid by 2001; 470,000 of these households were in KZN. A key feature of this roll-out, which distinguishes it from many other national electrification programmes, was its focus on low-capacity household connections rather than on industrial connections (Gaunt, 2003).

Even with this focus on households, evaluating the effects of electrification is tricky. A large literature on the relationship between infrastructure and economic growth acknowledges that infrastructure spending may be targeted to growth centers, or towards areas that are lagging behind, but politically important. This selection implies that estimates of infrastructure effects obtained by comparing electrified and non-electrified areas may be biased upwards or downwards.[4] Furthermore, teasing out the effects of infrastructure on economic variables (for example, on the demand for or supply of labor) is difficult without adequate controls for economic trends. These difficulties characterize an analysis of South Africa's roll-out: it was socio-politically motivated and occurred during a time of economic restructuring. Electricity projects were not randomly placed, making it difficult to solve the identification problem by comparing employment outcomes across electrified and non-electrified communities.

To address the bias arising from endogenous placement and confounding economic trends, I collect and use information on the technological constraints on roll-out to generate exogenous variation in electricity project allocation to communities. I match this administrative and spatial data with two waves of Census data on outcomes for rural KwaZulu-Natal (KZN), a province containing one-fifth of the population of South Africa.[5] Using this two-wave panel of communities, I compare changes in male and female employment rates in communities that receive an electricity project between the two periods to changes in those that do not. I use community land gradient to instrument for project placement between 1996 and 2001 since gradient directly affected the average cost of electrifying a household and was a primary factor in prioritizing areas for electrification. I argue that in the case of rural KZN, gradient is unlikely to directly affect employment outcomes, conditional on baseline controls and district fixed-effects. [6] To bolster confidence in the research design, I show there is no correlation between land gradient and growth in major sources of female employment. I also conduct a placebo test for areas electrified prior to 1996 and find no evidence that employment grows differentially in flat versus steep areas, post-electrification.

To complement the main analysis, I also present evidence on employment, hours of work, wages

[4]An established macroeconomic literature estimates the effects of public infrastructure on total factor productivity using time-series data. Aschauer (1989) is a classic reference on the relationship between public infrastructure and productivity growth in the US, Canning (1998) provides cross-country evidence and Bogotic and Fedderke (2006) perform an analysis for South Africa. See also World Bank (1994) and Jimenez (1995) for good overviews of this literature.

[5]Electricity projects are likely to have larger impacts in rural areas where reliance on wood is high in the baseline period. Administrative data on these projects is also only available for KZN province.

[6]This identification strategy is similar to Duflo and Pande (2007) who use functions of gradient to instrument for dam allocation within Indian districts.

and earnings using an alternative research design with four waves of household survey data spanning the same time period. In this approach, the unit of analysis is the magisterial district (larger than a community) and I address the same issues of endogenous placement and confounding trends in a different way: by directly controlling for magisterial district (MD) fixed effects and MD-specific trends. Effects of electrification on labor market outcomes are identified from variation in the within-MD growth in electrification rates, after controlling for MD-specific trends in outcomes.

My results indicate large shifts in energy use for home production in electrified areas. The share of households using electric lighting rises by 23 percentage points in electrified areas and the share of households cooking with wood falls by 3.9 percentage points over the five-year period. Even larger shocks to home production technology occur in communities that are electrified by virtue of being flatter: IV estimates are three to seven times larger than ordinary least squares (OLS) estimates.

One of the unusual aspects of the South African case is that electricity projects seem to be targeted towards areas doing more poorly over time: OLS estimates indicate that employment rates are not significantly different for women in electrified areas and are 1.1 percentage points lower for men. I argue that much of this result can be explained by the fact that electrification is differentially likely to occur in places experiencing industrial restructuring: shifts out of male-dominated commercial agriculture and mining sectors were occurring during the 1990s and contributed towards increases in national unemployment rates. Not being able to control for these differential trends across communities in the Census data biases OLS estimates downwards. In contrast, IV results indicate that female employment rises by a significant 9.5 percentage points. Given a baseline female population of 165,637 in electrified areas, this translates into an increase of approximately 15,000 newly employed women or 0.75 percent of the estimated 2 million new jobs created across the country over the period (Casale and Posel, 2004). Results from the MD-fixed effects analysis using the household survey data also point in the direction of increases in employment and to significant increases in hours of work for both men and women.

After establishing that employment increases in electrified areas once key sources of selection are accounted for, I explore the channels through which household electrification had this effect on rural labor markets. I use Census data to show the lack of cross-community employment spill-overs generated by the roll-out. Since communities are small geographical entities, this is some evidence that electrification did not stimulate large increases in labor demand outside of the household via the establishment or expansion of firms. Given that the capacity of household electricity supplied was too small to be of use to large or mid-size firms (South African Department of Minerals and Energy, 2006), this result is not surprising.

The weight of evidence from the fuel use and female employment results points to household electrification operating as a labor-saving technology shock to home production in rural areas, which lead to a corresponding increase in female labor supply. These measured employment effects are not uniform across all women: I show that electrification has larger effects on female employment in middle-poor communities (where new access more easily translates into more use of the new technology) and for women in their thirties and forties, who are less likely to live with young children requiring full-time care.[7] Furthermore, using the alternative research strategy (MD-fixed effects and MD-trend), I show that wages for women fall significantly in magisterial districts where electrification is expanding more rapidly. This fact is difficult to reconcile with electricity causing large increases in labor demand.[8] However, despite all of this evidence, my data do not allow me to conclusively rule out that electricity lowered the cost of producing new, home-based services for the market, thereby presenting individuals with alternative ways to use their labor time (e.g. in self employment). Indeed, the fact that the MD fixed effects analysis shows increasing female and male hours of work in electrifying areas suggests that electrification did not exclusively affect rural labor markets by freeing up time from home production.

A final important channel that I investigate is related to migration. I show that communities receiving electricity projects experience much higher population growth – in OLS and IV specifications – than non-project areas. In these areas, the composition of the population also changes: electrified areas experience increases in the fraction of men and women with a high school certificate. In such a short five-year window, these changes in educational composition of adults are going to be driven by migration behavior. While migration itself is a potential labor market response to electrification, it also potentially confounds the employment results - since the entire increase in employment could be due to in-migration. To rule out this explanation, I use information on who has recently migrated to the community to perform a simple bounding exercise which shows that electrification also raises employment for incumbent women.

This paper contributes to the literature on the effects of physical infrastructure in developing countries in several ways. First, it places a new emphasis on employment effects in an area where the current focus is largely on poverty, health and education outcomes.[9] Studies which do not measure

[7] In some settings, access to cleaner cooking technology may lead to increased ability to work for women who experience improvements in their own health or the health of their children. However, health effects are less likely to drive employment responses in South Africa: the 1998 Demographic and Health Survey indicates that respiratory disease prevalence in rural KZN is substantially lower than prevalence reported in Asian countries where much of the indoor air pollution research is conducted (see www.who.int/indoorair/en/ as well as Rosenzweig, Pitt and Hassan (2006) and Duflo, Greenstone and Hanna (2008)); most respiratory infections are treated within public health clinics and about 30% less time is spent in cooking than in comparable Asian countries (data from 2001 South African Time Use Survey.)

[8] In fact, the main result that female employment rises in electrifying areas follows in the tradition of a large literature on female labor force participation that connects aspects of economic development, labor economics and economic history.Many explanations have been proposed for the stylized fact that as countries develop, women tend to work more in the market (Mammen and Paxson, 2000): changing social norms about female work (Goldin (1994) and Fernandez (2002)), economic growth that is more complementary to women's labor than men's labor (Galor, 1996), new access to technologies of fertility control (Goldin and Katz (2000) and Bailey (2006)) and the changes in cost of child care services (Simonsen (2005) and Gelbach (2002)). Recent simulations by Greenwood et al (2005) show that price reductions in household appliances contributed over 50 percent of the rise in female labor force participation in the US between 1900 and 1980, and Coen-Pirani, Leon and Lugauer (2008) provide microeconomic evidence that the diffusion of dryers and freezers in particular explains about half of the increase in married women's labor force participation in the US between 1960 and 1970.. Although see Bailey and Collins (2006) for some counter-evidence using variation in electrification use rates for an earlier period, 1930 to 1960.

[9] See Cutler and Miller (2005) for the effects of clean water technology in the US; Loshkin and Yemtsov (2005) for effects of infrastructure upgrades in Georgia, Russia; Duflo and Pande (2006) on the effects of Indian dam construction; Cattaneo et al (2007) for the effects of cement floors in Mexico. Existing evidence on how infrastructure affects work and wages is

employment effects could be missing important economic impacts, particularly when the infrastructure has a home production bias, as in the case of water and sanitation services.[10] Second, the results highlight short-run heterogeneity in the effects of infrastructure across types of communities and types of women and shows that important demographic changes may result from providing infrastructure to some regions and not others. Third, the fact that inference from comparing electrified and non-electrified areas at two points in time is misleading highlights the importance of designing an identification strategy robust to omitted variables' bias and selection issues. These are critical issues in all infrastructure studies. The main identification strategy used in this paper, which relies on administrative and spatial data to model project allocation, is likely to be feasible in other networked infrastructure settings. Combining empirical strategies that use different sources of variation to estimate the effects of infrastructure is also a potentially useful exercise, particularly in understanding some of the mechanisms through which infrastructure operates on labor markets.

The paper begins with a brief discussion of how a positive shock to home production technology may affect both labor supply and labor demand. Sections 3, 4 and 5 describe the data, context of South Africa's electrification and empirical strategy. Section 6 presents the main results and robustness checks while section 7 investigates some of the channels through which electrification affects employment. Section 8 concludes.

2 The theoretical impact of household electrification on home and market production

New access to household electrification potentially affects both home and market production, in ways that change the nature of work in the home and in ways that increase market labor. Providing new public infrastructure to a location may also induce employed and unemployed individuals to migrate in to the area. Outlining the form each of these changes may take is important for interpreting the empirical results in the paper.

First, consider home production. Household electrification may be characterized as a shock to the technology of home production that has an ambiguous effect on the supply of labor to the market.[11] This is because the arrival of electricity generates both a substitution and an income effect. Since the new, labor-saving technology makes the household more productive in time-intensive activities in the home – it reduces time spent in fetching wood, speeds up meal

limited, but suggests that effects could be large. Banerjee et al (2007) find that Chinese wages are 37 percent higher in areas transected by railroads while Akee (2006) estimates the effects of road construction on wage employment at 27 percent and on agricultural employment at 38 percent.

[10] Field (2007) highlights a similar point relating to how the establishment of property rights in urban Peru releases time spent on monitoring household assets for more market work.

[11] Becker (1965) and Gronau (1986) provide the canonical model for working through the effects of a home production technology shock.

preparation and enables more efficient food storage – families may decide to substitute more time towards these home-based activities.[12] However, the same technology shock also increases the total effective (time) endowment of the household (it increases the length of the effective day by allowing households to shift home production to the night hours) and in turn increases the demand for all normal goods. This endowment effect leads households to increase the supply of labor to the market, allowing them to earn income to purchase market-intensive commodities.[13]

While the net impact of substitution and endowment effects of electrification on market labor supply is theoretically ambiguous, the substitution effect is likely to be relatively smaller since the demand for home-produced goods (e.g. meals) is bounded above. Hence, we should expect the arrival of household electricity to both alter the time spent in and the nature of home production as well as increase labor supplied to the market. Responses to the technology shock will also differ within and across households. If the primary channel through which electrification affects the household is through changes in home production activities, we should expect larger market labor responses for women than for men, since women are traditionally the agents most heavily involved in home production.If there is heterogeneity across households in initial home production technologies (some of which can be measured), or in the degree of substitutability of home for market commodities (for example, meals versus child-care) then differences in the labor supply effects of electrification across these different types of households will be observed.

Note that home production activities are important and time consuming in my study area. Figure 1 shows the fraction of rural African households in KZN that report different sources of fuel for cooking and lighting in the 1996 and 2001 Census, separately in communities that do and do not gain access to electricity during this period. Almost 80% of all households rely on wood for cooking and candles for lighting in the mid-1990s. In electrified areas, the fraction of households using electricity for cooking increases almost three-fold in five years while the fraction of households using electricity for lighting increases by more than three. In the empirical work below, we investigate whether getting access to household electrification in rural KZN changes methods of home production (we have no data on time spent in these activities), whether electrification increases employment in the market and whether these effects differ by gender and household characteristics.

In addition to effects on the supply of market labor, consider the impact that electrification may have on work opportunities in these rural areas. New infrastructure may allow firms to open up in places where they would previously have been unable to operate and may also encourage the development of market activities inside the household. In the first instance, if household electrification is accompanied by firm electrification, market jobs would be created outside of the household. In the second instance, household electrification alone could create new jobs. This is

[12] This is an argument for why the ‘industrial revolution in the home’ contributed to “More work for mother” (Cowan, 1985).

[13] The more income-elastic the demand for market-intensive goods is, the stronger this endowment effect will be in pushing households to supply more labor to the market. There is good evidence to show that as people get wealthier, expenditure on non-food items increases relatively more than expenditures on food items and that the demand for food and calories is income inelastic (Houthakker (1957), Deaton and Subramanian (1996)).

because electricity makes a range of new activities possible: households may be able to store food prepared for larger groups more easily, operate small appliances that generate valuable market services (e.g. hairdryers for hairdressers, cell phone charging stations) and establish small-scale cottage industries (e.g. production of textiles or crafts for local markets). In this way, household electrification could unlease previously unrealized demand for labor and an increase in market work.

To isolate which of these channels is most important for how rural electrification affects market employment, it would be ideal to show what happens to (1) home production activities (2) market employment (3) the prevalence of firms in rural areas (4) the prevalence of home-based services and cottage industries in rural areas and (5) market wages in areas that gain new access to electricity. Data limitations restrict my analysis to examining (1), (2) and (5). However, I will argue that there is very little evidence to suggest that the specific form of electrification in KZN enabled medium or large firms to develop (no evidence that (3) changed). Any labor demand effects of electrification are therefore likely to be operating through the increased range of market services and small-scale productive activities that can be done from home. In effect, the reduced-form market employment results I find capture a combination of increased labor supply to the market (from the home production channel) and increased labor demand (from the ability to produce new goods and services for the market from inside the home). I provide some additional evidence that the net effect of these changes results in a fall in female wages and no significant differences in male wages.

Finally, increases in employment in response to electrification may be driven by in-migration of certain types of individuals. While in-migration could be an important response to electrification (and I provide some evidence that it is, in the South African case), if individuals moving in to the area already have jobs elsewhere, we would mistakenly attribute the increase in employment to the new household electrification, when the main effect of the roll-out is instead to change the composition of the community. To isolate whether this new access to infrastructure directly increases employment, net of migration, I restrict later parts of my analysis to non-movers.

3 Electrification roll-out in South Africa

3.1 Details of the program

Under apartheid, many African households were denied access to basic services. This was true particularly in homeland areas, which were pockets of land designated for African settlement and which functioned as labor reserves for the white economy.[14] In 1994, all homelands were legally reintegrated into South Africa (Christopher, 2001) and the South African government was again responsible for basic service provision in these areas. By 1990, most economic units and white settlements were electrified and the political concerns of the 1980s had led to extensive electrification

[14]Throughout, I retain the use of apartheid-era racial classifications: African for black South Africans, and white and Indian.

of commercial white farms in rural areas (Gaunt, 2003). In contrast, high-voltage lines carrying power from generation plants to white farms and towns often transected homelands that were themselves without power. At the time of the first democratic elections in 1994, over two-thirds of African households had no access to electricity. The National Electrification Programme (NEP) made the elimination of this backlog a development priority.[15]

As part of the NEP, Eskom– South Africa's national electricity utility– committed to electrify 300,000 households annually from 1995 onwards. These targets were regarded as "firm and non-negotiable" (Eskom, 1996) and connections were fully subsidized by the utility (Gaunt, 2003).[16] Since Eskom was a parastatal and a monopolist in electricity generation during this period, internal support for the roll-out was partly strategic. Eskom was interested in signalling to the government that introducing competition to the industry was not necessary to provide full access to previously disadvantaged communities.[17] As a result, Eskom met their connections targets in most years. Between 1993 and 2003, over 10 billion Rands (about USD1.4 billion) was spent on household electrification and over 470,000 households (28% of all KZN households) were electrified in KZN province alone.

Once an area had been targeted for electrification, each household was fitted with the basic connection package, consisting of an electric circuit board, a pre-payment meter, three plug points and one light bulb. Households received a default supply of 2.5 amperes or could upgrade to a 20 ampere supply for a fee of about ZAR40 (USD6.00), which most of Eskom's 3 million customers chose to do (Gaunt, 2003).[18] Default supply was sufficient for television, radio, two lights and several small kitchen appliances; an upgraded supply could power more appliances simultaneously including a refrigerator and a small water heater (South African Department of Minerals and Energy, 2004).[19]

This subsidized roll-out changed the option to use electricity. Industry experts agreed that "Electric lighting was synonymous with the roll-out"[20], and that the NEP did reach poor households. However, households were still required to pay for using the service by purchasing electricity credits loaded on to pre-paid cards. In 1999, household electricity cost $0.039 per kilowatt hour (kWh).[21]

[15]This section draws on written sources (Gaunt, 2003; Eskom, 1996) and interviews with Eskom engineers and planners (Ed Bunge, Eskom Electrification Engineer, Amos Zuma, prior head of Electrification in Pietermaritzburg, Innocent Nxele, prior head of Electrification in Margate) and energy experts (Gisela Prasad, Energy Research Development Council at the University of Cape Town, Trevor Gaunt in the Department of Engineering at the University of Cape Town) conducted in Durban, Cape Town and Johannesburg between May 2006 and May 2007.

[16]In early years, connection fees were charged to consumers but rarely collected.

[17]Personal communication with Trevor Gaunt, head of Department of Electrical Engineering at the University of Cape Town, May 31 2006

[18]An ampere is a unit of electric current. Larger electrical appliances require a higher amperage.

[19]Household survey data collected in the same area as my study sample indicates that newly connected households report large uptake of appliances that improve the efficiency of home production (electric tea kettles, refrigerators and electric lighting). Author's own calculations from the KwaZulu-Natal Income Dynamics Study, a small panel study of households from 1993 and 1998.

[20]Interview with Gisela Prasad, University of Cape Town Energy Research Centre, May 2007

[21]The corresponding residential retail price of electricity per kWh in the USA was $0.083 in 1996 and $0.073 in 2001 (US Department of Energy, 2007)

Estimates of load demand from Eskom reports suggest that most rural households used between 35 and 60 kWh per month, translating into energy expenses of between $1.37 and $2.34 per month (Gaunt, 2003), or 1.8 percent of median monthly household income in rural KZN in 1995. Because of this positive marginal cost, the poorest households are likely to have been the least responsive to the new technology in the short-run. This is something that I test for in the empirical work.

3.2 Selection of project areas

Almost by definition, networked infrastructure of any kind requires that even identical consumers be connected in some order. In the context of the NEP, local political pressures and connections costs each played an important role in prioritizing communities for electrification.

Gaunt (2003: 91) comments that although objective criteria were identified for ranking communities, political pressures were part of the "not-easily-identifiable but good reasons for selecting particular target groups". In KZN, the African National Congress and Inkatha Freedom Party were fierce competitors for provincial governance in 1994 and for local governance in 1995 and 1996. This political rivalry arguably influenced local public goods allocations.[22] As there is no way to measure these political factors in my data at baseline or how they may have affected allocation of projects, I treat them as omitted variables.

Annual Eskom reports and interviews with planning engineers additionally point to the central role of costs in allocating projects to places. The dual pressures of connections targets and internal financing meant that Eskom had strong incentives to prioritize areas with lowest average cost per household connection.[23] These cost factors are central to the identification strategy in this paper. The bulk of electrification cost is in laying distribution lines out from electricity sub-stations to households. Three main factors reduce the cost of these distribution lines: proximity to existing sub-stations and power lines; higher density settlements; and land gradient and terrain. The less of an incline the land has, the fewer hills and valleys and the softer the soil, the cheaper it is to lay power lines and erect transmission poles (Eskom, 1996; West et al, 1997).

I assemble measures of each of these cost factors in my data. Distance from the grid and household density are important control variables as both are cost variables and are likely to be correlated with economic opportunities that could directly affect changes in employment. In contrast, land gradient is much less likely to directly affect employment growth, conditional on other spatial variables and district fixed-effects. Land gradient forms the basis of my instrumental variables strategy that addresses the multiple biases arising from selection on unobservables and confounding

[22]See Christopher (2001), Johnston (1997) and Piper (1999) for overviews of the political landscape in KZN. Khan et al (2006) describe how variation in levels of education and ability among tribal authorities affected their ability to lobby government for effective service delivery.

[23]Barnard (2006) describes factors affecting network extension to rural communities in KZN: "In the case of an electrical network, ideally the best route would run along the least slope, avoid forests, wetlands and other ecologically sensitive areas, be routed near to roads and avoid households, while running near densely populated areas in order to easily supply them with electricity."

trends. Section 5 provides further motivation for using gradient as an instrumental variable.

4 Data

Five sources of data are used in the main empirical work: community aggregate data from two publicly available Census surveys, two data sets which I collected on Eskom infrastructure and administrative data and one geographic data set which I constructed using spatial mapping software (ArcGIS). This software was also used to link the five data sets to each other. Details of the data and matching exercise are provided in the data Appendix 1. For some parts of my analysis, I also use the 10% micro Census data for 1996 and 2001 and individual data from four cross sections of household survey data: the 1995, 1997 and 1999 South African October Household Survey (OHS) and the 2001 September Labor Force Survey (LFS). None of the individual data in the Census or household surveys contain enough geography to identify communities receiving an electricity project, which is necessary for implementing the main IV strategy. However, individual-level data are useful for assessing the quality of the aggregate Census data and for providing complementary evidence on employment, wages and earnings using a different empirical strategy.

The primary unit of analysis in most of the paper is the census community. A community is small; roughly equivalent in size to a US Census tract. The median number of households per community is 203 in 1996, and 253 in 2001. Ninety-nine percent of communities have no more than 900 households. Key variables captured in the aggregate community-level Census data include the fraction of households with electricity in each year, the fraction of African adults by age group and labor market state and the fraction of households living below a poverty line. I also create a household density measure using the land area of the community. Results are not weighted, as all variables are derived from the full population Census.[24]

My sample of communities is restricted to rural ex-homeland (tribal) areas in KwaZulu-Natal (KZN). KZN is home to one-fifth of the population of South Africa, about 9.5 million people in 2001. The period from 1996 to 2001 is a relevant window for examining the effects of rural electrification since the appropriate technology for supplying smaller power loads to rural areas had been developed by the mid-1990s. Since rural households are more likely to rely on time-consuming traditional fuels than urban households, we should see larger effects of electrification in this group. Census micro data from 1996 confirms that 63.4 percent of rural households used wood for cooking, whereas only 2.7 percent of urban African households did so.[25] In addition, there are potentially fewer economic confounders in rural areas than in urban areas in the first years after the end of apartheid. Access to other development services is one source of confounding in rural settings; migration is another clear

[24]Statistics South Africa adjusts for under-count after enumeration (Personal Communication with Piet Alberts, Senior Statistician in the Census department of Statistics South Africa, May 2007).

[25]Expansion of service in rural areas is also more straightforward to analyze than urban areas, since Eskom is the sole distributor of power for rural areas. In urban areas, distribution rights are shared with local municipalities, making it more difficult to model roll-out.

challenge to identification. I discuss my approach to these issues in the next section.

Using digital elevation data, I construct measures of land gradient for each community. Gradient is one of the major factors affecting the average cost per household connection, and the distribution of this variable is illustrated in Figure 2 along with the boundaries of the sample of 1,816 rural communities. The geographical fragmentation that characterized the former KwaZulu is evident; the apartheid government forcefully resettled Africans to areas deemed inhospitable for white settlement, wherever those happened to be (Christopher, 2001). Gradient varies widely across the region and is very steep (dark shaded areas) in some areas. The Food and Agriculture Organisation categorizes the average gradient of this area as "strongly sloping" (FAO, 1998). Note that there are 10 districts that span this area; districts are far larger entities than communities and very much like local labor markets (in household survey data, very few people report working in outside of their district).

Census geography provides measures of the distance from each community (its centroid) to the nearest main road and town in 1996. These variables capture some information about access to local economies and job opportunities. I also collected technical data from Eskom planning engineers on the location of the electricity distribution network in 1996 to create another important cost variables: the distance from each community to the closest 1996 electrical substation.

To assign electrification status to each community, I collected administrative data from Eskom on the location and number of new household connections made between 1990 and 2007. The strength of defining electrification status using project data is that I can identify when a community gets new access to infrastructure instead of relying on on time variation in the use of electricity at the household level, which may be correlated with changes in wealth.

An area is electrified if the community had its first electricity project between 1996 and 2001 (inclusive) and non-electrified if it never received an electricity project or only had a project post-2001. Areas with pre-1996 projects are excluded from the main analysis; there are 406 of these out of the total 2,398 tribal areas in the sample (17 percent). I use these communities to conduct a placebo test in support of the exclusion restriction.

Several important features of project placement are evident in Figure 3, which shows the distribution of (dark shaded) electrified and non-electrified areas. Electrified areas are not all positioned close to the 1996 grid infrastructure, and many areas adjacent to the grid are non-electrified areas. Being close to the original grid is neither necessary nor sufficient for electrification between 1996 and 2001, although the first stage shows that proximity does raise the probability of getting a project. Proximity to a town is also not necessary for electrification. Finally, electrified areas are distributed across the province rather than clustered in one area. This allows me to include district fixed-effects to absorb aggregate differences in employment growth rates across local labor markets. In the next section, I describe how pairing the variation in land gradient with the variation in project status identifies the effects of rural electrification.

5 Empirical strategies

Let y_{jdt} be outcome y for community j and district d in time period $t = [0, 1]$. y_{jdt} measures (for example) the fraction of households using different fuels for cooking, or the fraction of men or women employed.[26] T_{jdt} is an indicator variable for whether a community has received an electricity project by time period t. If electrification T_{jdt} was randomly assigned across communities, we could estimate the average treatment effect of electrification (α_2) by ordinary least squares:

$$y_{jdt} = \alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \rho_d + \lambda_d t + \epsilon_{jdt} \tag{1}$$

where μ_j is a community fixed effect, $\delta_j t$ is a community trend, ρ_d is a district fixed effect, $\lambda_d t$ is a district trend and ϵ_{jdt} is remaining idiosyncratic error. To eliminate μ_j and ρ_d, re-write equation (1) in first differences so that ΔT_{jdt} is 1 if the community has an electricity project between t and $t + 1$, otherwise 0:

$$\Delta y_{jdt} = (y_{jdt+1} - y_{jdt}) = \alpha_1 + \alpha_2 \Delta T_{jdt} + \lambda_d + (\delta_j + \Delta\epsilon_{jdt}) \tag{2}$$

In my dataset, I have measures of Δy_{jdt}, ΔT_{jdt} and λ_d, but I have no way to account for δ_j with only two years of Census data. Estimating equation (2) by OLS will not provide the correct answer to the question: "what is the causal effect of electrification on employment?" if $\delta_j + \Delta\epsilon_{jdt}$ is correlated with ΔT_{jdt}. Positive selection on community trend (δ_j) would bias $\hat{\alpha}_{2,OLS}$ upwards, if electricity projects are allocated to communities growing faster for unobservable reasons. Negative selection on the community trend (δ_j) is also possible if projects are targeted to more disadvantaged areas.[27] Since electrification was driven by a socio-political compact between Eskom and the newly-elected government, political concerns for disadvantaged communities could well have directed negative selection on trend. In addition, during this period, South Africa's economy is restructuring and jobs are being lost in certain sectors and areas (Banerjee et al (2006)). If the areas that are losing jobs are also the ones targeted for projects, $\hat{\alpha}_{2,OLS}$ would be biased downwards.[28]

I take two approaches to dealing with these sources of bias. I include a vector of baseline covariates (X_{jd0}) to control for some factors affecting a community's growth path (δ_j), where the baseline year is 1996. Covariates include household density, fraction of households living in poverty, distance to the 1996 grid, distance to the nearest road and town, fraction of adults that are white or Indian (as a proxy for local employers) and measures of adult educational attainment in the area. Since the Census captures only a crude measure of income– in intervals and only at the household level– I also include two standard proxies for the extent of community poverty: the share of

[26] This definition of labor market participation is typically used in the literature on female employment (see Costa (2000), Goldin (1994) and Mammen and Paxson (2000)).

[27] Banerjee and Somanathan (2007) show that access to public goods in India increases more among more politically mobilized disadvantaged groups.

[28] Measurement error in ΔT_{jdt} presents a third practical challenge for estimating equation (2). See the discussion of this issue in Appendix 2.

female-headed households and the female/male sex ratio (Standing et al,1996).

Even with these controls, confounding trends in community-level employment and unmeasured political factors are still of concern. To overcome these issues, I instrument for program placement using average community land gradient (Z_j). The system of equations to be estimated is:

$$\Delta y_{jdt} = (y_{jdt+1} - y_{jdt}) = \alpha_1 + \alpha_2 \Delta T_{jdt} + X_{jd0}\beta + \lambda_d + (\delta_j + \Delta\epsilon_{jdt}) \tag{3}$$

$$\Delta T_{jdt} = \pi_0 + \pi_1 Z_j + X_{jd0}\pi_2 + \gamma_d + \tau_{jdt} \tag{4}$$

where $(\delta_j + \Delta\epsilon_{jdt})$ and τ_{jdt} are unobserved. The identification assumption is that conditional on baseline community characteristics, proximity to local economic centers and grid infrastructure, land gradient of the community should not affect changes in employment outcomes independently of being assigned an electrification project.

An obvious concern with using land gradient as an instrument in a rural setting is that it may affect agricultural outcomes or characteristics of individuals settling in steep and flat areas.[29] In rural KZN, the direct impact of gradient on agricultural productivity and agricultural employment growth is limited, since most people are not farming. Less than 10 percent of employed individuals are involved in agriculture (see the tables and discussion in Appendix 2).[30] In addition, validity of the instrument is threatened only if non-random sorting of individuals across flat and steep areas resulted in differential employment growth, independent of the effects of new electrification. Mobility within homeland areas during this time period is somewhat limited by a lack of property titling and the role that tribal authorities (rather than the market) play in allocating land.[31] I investigate in-migration to rural areas as a possible response to electrification and find substantial migration in to flatter areas. However, I also use information on former place of residence to estimate (3) and (4) for the subset of non-movers. The main results are robust to this re-specification of the outcome

[29]Gradient is sometimes used as a control variable in estimating agricultural production functions (Udry, 1996). More recently, time-invariant topographical variables have been used to generate random variation in infrastructure allocation (Duflo and Pande, 2006) and intensity of agricultural crop type (Qian, 2006).

[30]See Simkins (1981), Standing et al (1996) and Aliber (2001) for a description of historical and current agricultural conditions in ex-homeland areas. Vink and Schirmer (2002) describe how overcrowding and agricultural decline worsened in the homeland areas in the 1970s and that by the mid-1980s, farming accounted for only about 10% of total household earnings in these areas. Ardington and Lund (1996) examine various national, sub-national and local datasets that try to measure employment in rural areas of KwaZulu-Natal in the mid-1990s. They paint a nuanced picture of the role of agriculture in supporting the ex-homeland economies and describe the importance of migrant remittances and pension income in creating livelihoods for rural dwellers. Different datasets indicate that only about one third of households in rural KZN do any subsistence agriculture and that this activity contributes an even smaller fraction – between 7% and 16% – towards total household income. The largest sources of income in these areas are wages, remittances and transfers (Ardington and Lund, 1996: 39). They write that "a significant percentage of the income of rural households is sourced outside the household and indeed outside rural areas" (Ardington and Lunch, 1996: 48), and that households typically have access to multiple sources of income. They also note that "it is clear that land is nowhere the 'main source' of income for the majority of rural households" (Ardington and Lund, 1996: 55).

[31]Personal communication, Department of Land Affairs, Pietermaritzburg, June 2006. Household survey data from the 1990s indicates that about 60% of households that do farm, have land allocated by tribal authority. Non-random initial settlement across flat and steep areas is less likely in these areas given the forced nature of these settlements under apartheid spatial planning laws (Christopher, 2001).

variable.

To further bolster confidence in the research design, I test for whether gradient is correlated with outcomes for areas that were electrified before 1996, and whether it is correlated with changes in other services including access to water and to flush toilets. I also test for whether gradient is correlated with changes in two important sources of employment. In each case where we might be concerned that gradient could have a direct effect on outcomes, the data cannot reject zero correlation.

Conditional on instrument validity, $\alpha_{2,IV}$ captures the local average treatment effect (LATE) of electricity projects on employment growth at a community level. It is typical to think about LATE's in terms of marginal effects for individuals who are affected by the instrument. In this paper, individuals aggregate to communities and so community composition drives these marginal effects. If individuals living in flatter areas can better afford electricity once it arrives, then a larger than average treatment effect may be measured for these areas. Or, if individuals living in flatter communities have fewer home production demands, they may also respond more to the arrival of the new technology. In addition, employment returns to electrification may differ by gradient itself, leading to larger estimated employment effects for marginal than for average communities. For example, flatter areas always have lower commuting costs, so individuals in flatter areas always face a higher net wage. Since these individuals are initially closer to the employment participation margin, they will always be more likely to respond when electricity arrives.[32] These reasons imply that we might expect IV estimates to be larger than average treatment effects for a more general population. I explore several of these avenues for heterogeneity in responses in the last part of the paper.

In addition to the main analysis using an IV strategy to deal with the likely correlation between community trends (δ_j) and project status ΔT_{jdt}, I present complementary evidence using an alternative identification strategy that controls directly for differential trends. I pool information from four cross-sections of South Africa household survey data to estimate the impact of electrification on male and female employment, hours of work, wages and earnings and account for trend differences across electrified and non-electrified areas by including a full set of area fixed effects and area-specific trends. Because the household survey data contain a richer set of labor market outcomes, they can be informative about whether electrification is operating on these labor markets mainly through labor demand or supply channels. The major drawback to using these data is that an individual can only be situated in the magisterial district (MD) in which they reside, which cannot be linked to an Eskom project area. These magisterial districts (MD) are larger than Census communities (there are 38 of them in my sample), but still smaller than districts j. In place of using the Eskom administrative data to define "new access" to the grid, I use the fraction of household

[32]That gradient is correlated with transportation costs (specifically through access to roads) is a potential threat to validity. Changing economic activities in distant markets may be more easily accessible for flatter communities, hence making gradient itself a 'treatment'. See Nunn and Puga (2007) for more on the direct economic effects of terrain ruggedness on transportation costs. To test whether it is largely access to roads that drives the employment response in KZN, I restrict to areas without any main roads. Although the coefficient on electrification falls slightly in the female employment regressions, there is no significant difference in results (not shown here, available from author upon request).

using electric lighting as an indicator of access to electrification infrastructure. This is a reasonable indicator for expanding access to the grid, as almost all households that got access to the grid were able to use electric lighting.[33]

I restrict the sample in each of these household surveys to African men and women aged 20 to 59, living in rural (not only tribal) areas of KZN and create measures of employment (an indicator variable), usual weekly hours of work, earnings per month and wages (combining earnings and reported usual hours of work) for each individual. There are at least 1,500 observations for each year of data and sex. I regress each of these outcomes on age, age-squared and years of education, obtain the residuals from these regressions and average the residuals within year (t), magisterial district (m) and sex (s) to create up to 304 observations on outcomes (4 years*38 m observations for males and for females). I also create the fraction of households with electric lighting for each MD-year ($ELEC_{mt}$). Then, I estimate an OLS regression of these residuals ($\bar{\hat{\epsilon}}_{mt}$) on $ELEC_{mt}$ and a time trend (t) as well as a fixed effects regression that includes a full set of MD fixed effects (λ_m) and MD-specific trends (δ_{mt}) as below:

$$\bar{\hat{\epsilon}}_{mt} = \alpha_0 + \alpha_1 * ELEC_{mt} + \alpha_2 * t + \lambda_m + \delta_m * t + \nu_{mt} \tag{5}$$

It is not possible to instrument for $ELEC_{mt}$ in this framework, as gradient has no predictive power in explaining electrification rates in magisterial districts (i.e. there is no first stage). In the OLS regression, α_1 is identified using variation in electric lighting within and across MDs. In the FE regression, α_1 is identified using variation in electric lighting within the MD over time, after accounting for λ_m and for $\delta_m * t$. The ability to include MD-specific trend terms helps to control for the fact that areas where electrification is rolling out more quickly are likely to be trending differently to those where electrification is rolling out more sluggishly. If electrification is being targeted towards areas that are doing worse (or better) over time, these MD FE and MD-specific trends should account for much of this selection. Although these regressions contain only a small number of observations (38 MDs in each wave of data), they provide useful additional evidence of the effects of electrification on employment on the extensive and intensive margins and on earnings and wages.

6 Main Results

6.1 Comparing electrified and non-electrified areas at baseline

Table 1 presents baseline (1996) characteristics for all 1,816 communities, and separately by Eskom project status of the community. 25% of communities received an Eskom project between 1996 and 2001. In column (1), we see that communities in the sample area are poor: 61 percent of households

[33]Industry experts agreed that "Electric lighting was synonymous with the roll-out". Interview with Gisela Prasad, University of Cape Town Energy Research Centre, May 2007

live on less than 6,000ZAR per year (approximately USD840 at a 2006 USD/ZAR exchange rate).[34] On average, over half of households within a community are female-headed and the female/male adult sex ratio is well over 1. These variables underscore the historical function of the homelands as migrant-labor communities.

The stark differences across communities with and without an Eskom project are highlighted in column (4). Compared to non-electrified areas, electrified communities are significantly less poor, have lower adult sex ratios and higher fractions of high school-educated men and women and are 2.8-2.9 kilometers closer to the nearest road and town. Given that low average cost areas were prioritized for projects, it is not surprising that electrified areas have significantly higher household densities, are 4.1 kilometers closer to the nearest substation, and have a 1.2-degree flatter average gradient than areas without an Eskom project. If electricity projects had been randomly assigned to communities, we should expect most of these observable characteristics to be balanced across project and non-project areas. Instead, a joint test that all of these coefficients are zero can be rejected at the 1% level.

In the last two columns, values of each covariate are compared across steep and flat areas, first without any controls and second with all other controls (that is, in the final column I regress each covariate individually on the electrification dummy, controlling for all other covariates and ten district fixed effects). The results in column (6) show that gradient balances more of the community-level variables at baseline, conditional on all other controls. There is now no significant difference in poverty rates, the fraction of female-headed households, any of the distance variables or the fraction of females with high school. There are remaining, although small, differences in the adult sex ratio (0.004), household density (0.95 households per square kilometer) and men with high school (0.003).[35] A joint test cannot reject the null of all coefficients equal are to zero, at the 1% level (although it can be rejected at the 5% level).

Male and female employment rates and population totals in areas with and without an Eskom project are shown in Table 2. The main outcome variable is the employment to population rate of African women and men aged 15 to 59 inclusive. Questions about employment are similar across Census waves, although 2001 employment definition is somewhat broader than the 1996 variable, describing individuals who work for even one hour per week as employed. Since the main outcome variable is the change in employment rate, these differences will only be problematic if reported part-time work differentially contributes to new employment in flatter areas. Column (2) of Table 2 indicates that over the period, employment rates fall by 4 percentage points for men in these areas. Female employment rates remain steady on average across communities but low, at 7 percent. Employment is uniformly higher in electrified than in non-electrified communities in the baseline

[34] This is the cut-off for households in the two lowest income brackets reported in the Census.

[35] Consider the joint hypothesis that all coefficients are equal to zero, where we focus on the first ten variables from poverty rate to distance from the grid. Then, the relevant p-value for rejection of this joint null at the 1% level of significance is $p < 0.01/10 = 0.001$: if at least one p-value is less than 0.001, the null is rejected. In column (4), the null is decisively rejected; while in column (6), this null is not rejected.

period. Comparing changes in employment rates in Eskom project areas to the same change in non-project areas (column 6), the unadjusted estimate for women is not significantly different from zero while for men it is a statistically significant -1.7 percentage points.

Both in- and out-migration to rural areas is occurring during the period (Leibbrandt et al, 2003). Migration could contaminate any comparison of outcomes across electrified and non-electrified areas or it could represent a response to electrification.[36] Some of these changes are evident in the lower part of Table 2. Population growth is faster in Eskom project than in non-project areas, even though electrified areas begin with higher populations. Electrified areas grow at about 6 percent per year while non-electrified areas grow at about 3 percent.[37] Keeping in mind that communities are small, a 3 percent change in population over five years is an increase of 30 people in the median community.

Two striking points emerge from the unconditional difference-in-differences comparisons in Table 2: employment rates are very low for men and women, and are falling and falling faster for men in electrified areas between 1996 and 2001. That South Africa has low levels of employment is not a new insight. High unemployment rates have been documented by Klasen and Woolard (1999), Bhorat (1999) and Nattrass (2000) among others, in multiple years of household survey data. The fact that communities in Table 2 have even lower rates of employment than usual is partly attributable to the Census not containing detailed questions on work, and partly the result of including only rural, tribal (ex-homeland) areas of KwaZulu-Natal in the sample. As described by Ardington and Lund (2006: 12), the "four independent homelands and six self-governing territories established for the various 'ethnic groups' consigned millions of people to rural areas with few employment opportunities". These ex-homeland areas were ill-suited for agriculture and not intended to develop vibrant labor markets. In these areas, most formal wage work is still in the civil service (teachers) and in domestic work (both jobs that favor the employment of women); many jobs are marginal, with workers working under 20 hours per week (Ardington and Lund, 2006) and large fractions of households rely on income from welfare grants (old age pensions) and migrant workers.[38] Although the Census likely under-counts employment, in Appendix 2 I compare the community Census data to the individual level Census data and household survey data from the same years to show that much of the explanation for low levels of employment in my sample is due to a focus on tribal areas.

The large reductions in male employment in Eskom project relative to non-project areas should not be interpreted as the causal effect of electrification. Rather, the employment responses for men and women are confounded by broad changes in the South African labor market during the 1990s. Figure 4 shows trends in male and female employment over time using data from four waves of household survey data from 1995, 1997, 1999 and 2001. These employment rates are higher than the

[36]The overall population growth rate of 20 percentage points over the five-year period is approximately equivalent to a 3.7 percent growth rate per year.

[37]For areas that receive an electricity project, the increase in the number of women is 135/426; for men, the increase is 114/312. Averaging this increase over the five years between the Census translates into about 6% growth per year for Eskom project areas. A similar calculation can be done for non-electrified areas.

[38]Standing et al (1996). The 1996 micro Census data show that 32 percent of African households in rural KZN contain either a pensioner or a migrant worker, and 20 percent of households receive remittances from outside the household.

Census rates (they do include all rural parts of KZN, not just the tribal areas) but are still extremely low. Employment for men falls significantly between 1995 and 2001 and falls to a much lesser extent for women. Figure 5 shows wage trends using the same data: over the period, male wages are roughly constant while female wages fall and are lower in 2001 than in 1995. Dissecting the overall changes in employment, Banerjee et al (2006) document large shifts in the composition of jobs away from commercial agricultural and mining sectors, and towards service and retail sectors. These changes were largely a continuation of trend during the 1990s and impacted heavily on jobs in male-dominated sectors. The types of new jobs created during this time were predominantly low skill and in the informal sector: industries that favor female workers (Casale and Posel, 2004). Banerjee et al (2006) present evidence that the number of jobs for self-employed workers and household workers increased by over 200 percent and 44 percent respectively between 1995 and 2000.

While a common challenge in evaluating the effects of an infrastructure expansion on economic activity is how to identify whether infrastructure is being built in a simultaneously expanding economy or whether the infrastructure is causing the economic expansion, this challenge is slightly different in the South African case. Eskom was more likely to electrify places that were experiencing longer-term declines in employment and economic activity, largely because the expansion of the grid infrastructure was spatially constrained by the initial placement of the network and the network that existed at the end of apartheid was designed to service commercial farms and white towns in the province. Many of the factors that determined whether a community got early access to electricity were the same factors that increased the community's exposure to the industrial restructuring of the 1990s. Political reasons for targeting public investments towards previously disadvantaged areas would further reinforce this correlation. Both the IV strategy and the MD-fixed effects and MD-specific trends strategy proposed above are attempts to deal with these confounding factors in different ways.

6.2 Employment and wage effects of electrification: results

6.2.1 OLS and IV analysis

First-stage estimates for the allocation of an electricity project to a community are presented in Table 3.[39] The outcome variable is the indicator for whether a community received an electricity project between 1996 and 2001. A two standard deviation increase in gradient (about 10 degrees) reduces the probability of electrification by about 8 percentage points. Across columns, the size of the coefficient does not change substantially with the addition of more controls and the precision of the estimate improves, particularly after including district fixed-effects in column (3). When restricting to areas where no households had electricity in 1996 (column (5)), the gradient coefficient is slightly larger.[40]

[39]Results from a logit model of electrification are very similar to these linear probability model results.

[40]First stage estimates using alternative definitions of project status (years since electrified, fraction of households connected) convey the same message.

The other two cost variables have coefficients of the expected signs: a three-quarter standard deviation increase in distance from the grid (about 10 kilometers) reduces the probability of electrification by 1 percentage point, although this is not significant when other controls are added. A one standard deviation increase in household density (30 households) per square kilometer increases the probability of electrification by about 1.3 percentage point and this is robust and strongly significant across specifications.

These results provide mixed evidence on whether newly electrified areas are positively selected on wealth. While areas with more female-headed households (i.e. poorer areas) are significantly less likely to receive an electricity project, areas with more white and Indian adults (i.e. richer areas) are also less likely to be electrified during these years. The community poverty rate and sex ratio variables also have large positive coefficients in all specifications, suggesting that projects may be targeted to poorer areas. This lack of strong evidence for project placement in richer areas is consistent with the overarching socio-political motivation for the roll-out.

If electrification affects patterns of employment through the channel of reduced time in home production, households must switch out of traditional fuels when their communities are connected to the grid. Table 4 illustrates that this is indeed occurring. Each coefficient reported in the table is from a separate OLS or IV regression, where the outcome variable is the change in the fraction of households using electricity for lighting or cooking or using wood for cooking. Columns (2) and (3) do not contain any additional controls while columns (3) and (6) report results from regressions containing all relevant control variables. Both OLS and IV regressions illustrate substantial shifts towards using electricity for home production and IV results are substantially larger than OLS estimates. Average rates of electric lighting rise by 23 percentage points more in communities with an electricity project than in communities without in the OLS comparison of column (3). In the same column, reliance on wood for cooking falls by 3.9 percentage points and cooking with electricity rises by 5.6 percentage points. Column (6) indicates that in areas chosen to be electrified because of their flatter gradient (i.e. lower cost), use of electric lighting increases by a substantial and significant 65 percentage points, wood use falls by 28 percentage points and cooking with electricity rises by 23 percentage points.

To check that gradient is not simply picking up easier access to all types of services that could make home production easier, rows (4) and (5) of Table 4 present results for two additional outcome variables: the change in fraction of households with access to piped water in the home or within 200 meters of the house, and the change in fraction of households with a flush toilet at home. There is no evidence that electrified regions experience differential changes in these basic services. In fact, the IV results for water services in column (5) and (6) are in the opposite direction to what we would expect if gradient was simply a noisy measure of wealth: slightly flatter areas have larger reductions in access to water sources close by, although these estimates are not significant once all controls are added. The change in access to flush toilets is not systematically associated with electrification status or with land gradient. Table 4 demonstrates two points: gradient does not appear to capture access to development projects in electrified areas more generally (there is no reduced form relationship between increases in these services and gradient), and instrumented employment responses could be large, since the effects of electrification on household fuel use are much larger in IV than in OLS comparisons.

Employment effects for men and women presented in Tables 5 and 6 are consistent with this latter point.[41] In each column, the dependent variable is the change in sex-specific employment rate between 1996 and 2001. The coefficient on electrification in column (1) reflects the falling employment rates from Table 1: there is no significant change in female employment across project and non-project areas while male employment falls by 1.7%. Adding controls and district fixed-effects in columns (2) and (3) increases the coefficient on electrification slightly, with the female employment effect still not significantly different from zero and male employment becoming less negative and less statistically significant. The positive, significant coefficients on community poverty rate, sex ratio and female-headed households in both tables indicate that female and male employment is growing faster in poorer places.

IV estimates of electrification are substantially larger than OLS estimates and significantly positive for women in columns (6) to (9). Since gradient is correlated with some of the control variables (as Table 1 indicated) and the F-statistic on the excluded variable in the first stage is larger once other controls are included to absorb residual variation, my preferred estimates are the results in columns (8) and (9). In these columns, female employment increases by 9 percentage points in local labor markets that were electrified due to flatter gradient.[42] The Anderson-Rubin (AR) test for the significance of electrification effects for female employment strongly rejects zero and the 5% confidence interval is wider than the standard 5% confidence interval, between 5 and 35 percentage points. Male employment increases by a substantially smaller 3.2 percentage points, and this is not significantly different from zero under either the standard test or the AR test (column (9), Table 6). In fact, there is no reduced form for male employment (column (5)). The difference in the male and female employment effects estimated in Table 5 and 6 are, however, not significantly different from zero, so it is difficult to conclude that electrification differentially affected female relative to male employment.[43]

Another aspect of the results in Tables 5 and 6 that bears mentioning is the sensitivity of the

[41] The table shows robust standard errors, clustered at the main place level, one level of aggregation up from the community level and one level below the district. I separately implemented a correction for spatial correlation in the error terms in both the first stage and the IV regressions using methods developed by Conley (1999 and STATA code provided at $httE$: $//faculty.chicagogsb.edu/timothy.conley/research/$). While the standard errors increase slightly under this correction, all coefficients displayed in Tables 5 and 6 retain significance (or remain insignificant, respectively).

[42] To address concerns about over-optimistic inference with a possibly weak instrument, heteroscedasticity-robust Anderson-Rubin (AR) confidence intervals are computed for the second stage parameter estimate. These corrected confidence intervals have correct coverage properties in the presence of weak instruments while standard Wald tests do not. For more on these tests, see Moreira and Cruz (2005), Mikusheva and Poi (2006) and Chernosukov and Hansen (2007).

[43] I implemented this test by differencing the male and female outcome variables within community and then performing the same set of OLS and IV regressions on this new variable. This test respects the correlated structure of errors across male and female regressions. Results not shown, available on request.

female (but not male) employment results to inclusion of district fixed effects. In part, this reflects large differences in gradient across districts: about one quarter of the variation in gradient is accounted for by cross-district variation. Since some districts are very flat and other districts are very hilly, results excluding district FE are driven by cross-district comparisons in female employment growth. Including district fixed effects allows us to identify the effect of electrification by comparing slightly steeper to slightly flatter areas within the same local labor market. Since the female labor market is much more likely to be spatially segregated across districts (females are less likely than males to migrate far from home, for work), this provides an additional reason to prefer the estimates that include district fixed effects.

The IV results suggest that in a non-electrified community with the median number of adult women in 1996 (N=264), a 9 percentage-point increase in female employment raises the number of women working (where working is broadly defined) by 23 women, from 18 to 41. If we assume this 9 percentage point increase applies to the entire group of electrified communities (rather than marginal communities only), this translates into an increase of approximately 15,000 newly employed women out of the baseline female population of 165,637. This is 0.75 percent of the estimated 2 million new jobs created across the country over the period (Casale and Posel, 2004).

6.2.2 Threats to validity

If employment rates in steep and flat areas evolve differently, the gradient IV would be invalid. However, checking for differential trend is difficult without more years of data. This is where having the administrative data on electricity projects is helpful for conducting a placebo test. These data identify which areas are electrified before 1996– a set of areas that were excluded from the main analysis. For these areas, there should be no reduced-form relationship between gradient and employment growth between 1996 and 2001, since they have already received an electricity project. If there is, we would be concerned that gradient has a direct effect on employment growth. To test this, I select the sample of areas electrified prior to 1996 ($N = 406$) and estimate an OLS regression of employment growth between 1996 and 2001 on the full set of controls, and gradient. Column (1) of Table 7 contains the results of this placebo test. The coefficient on gradient is small (-0.001) and insignificant, providing no evidence of any such reduced-form relationship female employment (the same is true for males, results not shown). This boosts confidence in the research design.

A second potential threat to the validity of the IV strategy arises in the form of positive labor demand shocks that happen in flatter communities at the same time that electricity projects are being rolled-out. For example, businesses may expand in flatter ex-homeland areas after the end of apartheid for reasons unrelated to electrification. I do not have any information on the presence of firms in these areas. Instead, I test whether there are increases in the major sources of demand for female labor in flatter areas. Appendix Table A1 shows that professional and elementary occupations account for the majority of female employment in these areas. Data from the 10% micro Census sample (not shown) indicate that 75 percent of African women in rural KZN working as professionals or associate professionals are teachers and that domestic workers make up the majority of elementary occupations. New schools and new households are therefore the primary sources of new demand for teachers and elementary occupation workers, so labor demand shocks in these two industries are the most likely candidates for confounding IV estimates of electrification effects for women.

Using two waves of the South African Schools Register of Needs (1995, 2000) that capture the location of schools, I construct a variable measuring the change in the number of schools in each community over time. Over the five-year period, the number of schools across the rural KwaZulu-Natal area increases by 19 percent, from 1,770 to 2,801. This undoubtedly increases the demand for teachers across the province: however, column (2) of Table 7 indicates that this increase is uncorrelated with community gradient.

As a second indirect check that female employment is not being driven by a demand shocks that happen to occur in flatter areas, I proxy for "new employment opportunity" using the change in the fraction of adult population that is Indian or white. These are the individuals most likely to hire household workers (Dinkelman and Ranchhod, 2008). The number of Indian and white adults is not changing differentially across areas of different gradient, as column (3) of Table 7 shows. While other researchers have documented the growth in low-skilled and informal sector jobs in the economy during the 1990s (Banerjee et al, 2008; Casale and Posel, 2004), there is no evidence that these job openings are occurring differentially in flatter parts of KZN.

6.2.3 A complementary analysis using MD fixed effects and MD-specific trends

Since Census data provide no wage or earnings information, it is informative to turn to a complementary analysis using the four waves of household survey data from the October Household Surveys and Labor Force Survey. Individual-level data on employment, hours of work, earnings and wages is aggregated to the magisterial district level (MD) after removing age and education effects. First, I compare employment rates and usual weekly hours of work (Table 8, upper panel) and log wages and log earnings (Table 8, lower panel) across MDs that have different fractions of households with electric lighting, controlling for a common trend. The results of these regressions are shown in the OLS columns of Table 8. Then, I include MD fixed effects and MD-specific trends to control for the fact that electricity may be increasing differently in MDs with different trends. These results are presented in the FE columns of Table 8.

Consider first the estimates for employment: in areas where electrification increased, employment rises substantially under the OLS specification, for men and women. The average increase in electrification over the period (0.15) translates into a 1.5 percentage point increase in employment for men and a 2 percentage point increase for women, although the male-female differences are not statistically significant. Coefficients are similar under OLS and FE specifications, however, because of the small sample, none of the electrification coefficients are precisely estimated once all fixed effects and trend terms are included.[44] For men and women, usual hours of work per week increase

[44] I experimented with included a set of MD-year dummy variables instead of a linear trend for each MD. Results are robust

significantly in electrifying areas compared to non-electrifying areas (columns (5) and (7)) and this increase in work on the intensive margin is even larger in the fixed effects specification. Weekly work hours are between 3.15 (for women) and 3.45 (for men) hours higher in the MD in periods when electrification rates are higher, compared to work hours in the same MD in periods when electrification rates are lower. The male-female differences are not statistically significant. Note that the types of hours increase observed in these data are consistent with the type of work being informal and/or in self-employment rather than being in full-time formal sector positions.

Turning to the effects of electrification on wages and earnings: the female results suggest that wages are falling in areas where electricity is rolling out - and more so, in the fixed effects specification. For men, there are no significant differences in wages across electrifying and non-electrifying areas. Combining the increase in female hours of work and what looks like an increase in female employment on the extensive margin with the decline in wages, it is not surprising that there are no significant differences in female earnings across electrifying and non-electrifying areas (column (5)) or within an MD that sees growing electrification over time (column (6)). In contrast, male earnings do rise when electrification rates are higher. This also makes sense, given that the other outcomes show an increase in male hours of work but no decline in average wages.

Putting the household survey results together with the results from the IV analysis suggests the following interpretation of the effects of electrification on rural labor markets: employment on the extensive margin increases for women (IV results) and possibly for men (household survey data) although the male effects are difficult to estimate precisely. On the intensive margin, the best household survey evidence we have suggests that electrification raises the number of weekly hours of work for women and for men, and by about the same quantity. Although home production may be part of the channel through which electricity "frees up" women's time for the labor market, it is less likely to operate for men. More likely, electricity facilitates many new activities for both men and women in these areas. However, there is still no strong evidence to suggest that electrification stimulates large net increases in labor demand that translate into higher wages for men and women. Rather, female wages are falling in places where electrification is rolling out and male wages do not appear to be changing at all. In the next section, I use Census data to present two final pieces of evidence against electrification having a major impact on labor demand.

7 Channels

7.1 Electrification and labor demand

An additional test of the demand channel uses the fact that these communities are small and so any electricity project that generates new firms and hence new demand for labor should have spill-over effects in to neighboring areas. If firms create jobs for people living in neighboring areas, positive

to this alternative specification of MD-specific trend.

spill-overs in these unelectrified areas would dampen any effects of household electrification. If people move out of neighboring non-electrified areas towards electrified areas to get one of the new jobs, a negative spill-over would amplify electrification effects. In both cases, the effect is the sum of an incumbents' effect and a spill-over effect. In both cases, OLS and IV coefficients should be substantively different when adjacent non-electrified areas most susceptible to these spill-overs are excluded from the analysis.

To test this, I re-estimate OLS and IV regressions after excluding non-electrified areas within a one- and five-kilometer radius of an electrified area.[45] Table 9 presents results for each restriction. OLS coefficients are never significantly different from zero, while IV coefficients are large, positive and close to the main IV estimate: neither 0.076 nor 0.069 could be rejected in the full sample. Using this test, there is no evidence of large spill-overs across communities. Combined with the fact that the roll-out was driven by household targets and that capacity was too small to stimulate even mid-size manufacturing or service enterprizes (South African Department of Minerals and Energy, 2004), and along with the earnings results from the previous section, the lack of evidence for spill-overs supports the idea that electrification increased employment through some combination of a labor supply channel and increased opportunities for making market work inside the home.

7.2 Heterogeneous effects of electrification by income

IV estimates identify effects for communities which are cheaper to electrify by virtue of having a flatter gradient. In these communities, female employment may be more responsive to electrification than in an average newly electrified community. Recall that the expansion of infrastructure did not entail free electricity. So, one way in which marginal communities could differ from average communities is that they could contain more households able to switch home production technologies when the new service arrives.

Since the Census provides only a crude measure of poverty (household income is reported in intervals that are not consistent over time) I combine the three poverty indicators into a poverty index and consider the characteristics of communities in each quintile of this index. To create the index, I follow Card (1995) and Kling (2001): for the sample of communities in the steepest half of the gradient distribution, I use a logit model to estimate the probability of receiving an electricity project using baseline poverty rate, the baseline female/male sex ratio and the baseline share of female-headed households. Using coefficients from this regression, a value for every community in the sample is predicted. Each community is then assigned to a quintile of the predicted poverty index, where quintile cut-points are defined on the estimation sample only.

The graph in Figure 6 shows the fraction of the predicted poverty quintile that is electrified, for communities in the flattest and steepest halves of the gradient distribution. Both lines slope upwards, indicating that areas with higher predicted values of the poverty index (i.e. richer) are

[45] This is similar to what Black et al (2005) do in estimating the employment effects of coal booms and busts affecting local labor markets differentially.

more likely to actually receive an electricity project. The gap between the two lines shows that flatter areas are systematically more likely to be electrified than steeper areas. The middle-poorest and second-richest quintiles are most likely to have the probability of a project manipulated by the instrument which can be seen in the larger gap between the lines occurring at these quintiles. This larger gap is also evident in column (3) of Table 10 which shows the difference in electrification probability for flat versus steep areas, within each quintile. In column (4) of that table, I compute the contribution of each quintile to the final IV estimate by calculating a weight: the middle quintile and the second richest quintile together contribute over 65 percent to the IV result.[46]

Middle quintiles in particular may have larger employment effects because they contain households that experience larger changes in home production technology when electricity arrives compared to richer quintiles and they are more able to effectively use the new technology than the poorest quintiles. Table 11 shows that middle-poor areas are initially less likely to be using electricity than richer areas and more reliant on wood for cooking (columns 1 to 3). Columns (4), (5) and (6) of Table 11 present within-quintile reduced-form coefficients from regressions of the change in fuel use on a gradient dummy (1 is flat, 0 is steep). These columns indicate large increases in the use of electricity and large decreases in reliance on wood for cooking in flatter areas for middle-poor, second-richest and richest areas.[47] Finally, column (7) of Table 11 indicates that the female employment result is indeed driven by women living in middle- and second-richest quintile communities: the effects for these communities are large, positive and significant and are weighted most heavily in the final IV results.[48]

7.3 Heterogeneous effects related to other constraints on women's time

Women who have additional home-production responsibilities are less likely to be able to respond to new access to electricity, even though productivity at home may be substantially enhanced by the use of electricity. For example, child-care responsibilities raise the value of a woman's time at home and in the absence of pre-school care, this value only falls when children start school. Officially, school-starting age is between ages 6 and 7 in South Africa, but enrollment only reaches 90% by around age 9 (results from 2001 10% Census micro data, not shown). Children also create work at home though, and so the more children in the house that require child-care, the more time can potentially be saved with access to a more efficient power source.

[46]The computation of these weights is explained in the table notes. The IV coefficient is a weighted sum of effects for different groups that are differently affected by the IV (Kling, 2001). Each group may experience a different electrification effect and the weights determine which group's effect contributes the most to the total measured effect in the IV regressions.

[47]This is related to the point by Greenwood et al (2005) who argue that poorer households are the last to adopt durable goods for home production.

[48]The coefficients in this table are akin to reduced-form coefficients from a regression of the outcome variable on a binary version of the instrument and all controls. Dividing each coefficient by the corresponding coefficient in column (3) of Table 10 will give the IV coefficient.

Census micro data from 1996 give some indication of which women are more likely to live with a child younger than age 9. Figure 7 is a lowess-smoothed graph of the fraction of women of each age living with at least one child aged 9 or under. The graph is drawn for African women between ages 15 and 59 living in rural areas of KZN and shows a clear distribution of youngest children to households with both younger and older women.[49] After age 30 and up to about age 50, the probability of a woman living with a child who requires constant care falls substantially.

To investigate whether the employment effects of household electrification are largest for this latter group of women, I redefine the outcome variable to be $y_{ajdt} = \frac{E_{ajdt}}{P_{jdt}}$, where E_{ajdt} is the number of employed women in age group a for each of nine five-year cohorts and P_{jdt} is the total adult female population in each community in each year. This definition decomposes the employment result into effects for each age cohort: the estimated coefficients sum to the main electrification coefficient in the final column of Table 5. Table 12 presents OLS and IV coefficients (and standard errors) on the electrification dummy for separate regressions.[50] IV results are large and positive for each age group, but significant only for women in their thirties and late forties. Employment grows by 3 percentage points for women between the ages of 30 and 34, by 1.7 percentage points for the 35 to 39 year old group and by a smaller but still significant 1.4 percentage points for women in their late forties. Together, these age groups account for 65 percent of the total female employment result.

7.4 The importance of migration and demographic change in electrified areas

Out-migration from rural areas is occurring during this period of roll-out.[51] Cross et al (1998) also document rural-to-rural migration in KZN in the 1990s and show that part of this migration is towards areas with better infrastructure and amenities. Each of these flows could alter the composition of the population in electrified and non-electrified communities and contribute to employment effects in OLS and IV specifications. If workers with more education are migrating in to electrified areas, this could also account for some of the employment increases we see.

We can get some sense of the prevalence of migration flows at an aggregate level using the 10% sample of the 2001 Census micro data. These data provide information on the district that an individual was living in as of 1996 and the district in which they report living in 2001. Matching reports of out-migrants from each district to the total number of people living in each district in 1996, there is evidence of substantial out-migration from the rural KZN districts in our sample. Fifteen percent of men and 10 percent of women report out-migration during the five-year period.[52]

[49]The allocation of young children to households with older women is a common pattern in South Africa, where pension-aged women care for grandchildren in skip-generation households (Case and Deaton, 1998).

[50]Results for men are not shown as the electrification coefficient was never significant for any cohort.

[51]Leibbrandt et al (2002) find that men with intermediate levels of education tend to leave rural areas, leaving both the least and the most educated men behind.

[52]Out-migrants from KZN rural areas are defined as individuals in other parts of the country who report that they were resident in a sample sub-district in 1996. Sub-districts are larger than communities and the lowest available level of geography

These out-migration rates are, however, not significantly different by gradient.[53]

Table 13 presents differences in population growth rates in Eskom project and non-project areas, which tell us something about net migration at the community level. Note that with small communities, numerically small increases in population can translate into large percentage changes. The first two columns of the table show that electrified areas have significantly higher population growth than non-electrified areas, both in the OLS and IV results. Over the five-year period, the population in areas with an electricity project grows 17 percent more than in non-electrified areas, and this growth is 380 percent under the IV specification. The population is growing faster in marginal communities that are allocated to Eskom projects by virtue of their flatter gradient.

Next, I examine some of the compositional changes in population across project and non-project areas by regressing the change in the fraction of men and women with a high school education on the full set of controls, with and without instrumenting for project status. Columns (3)-(6) present these results. In the OLS comparison, there is no change in the fraction of women with a high school education, even though the population in electrified areas is increasing while for men, the fraction of those with high school is falling in places that are getting access to the grid. The IV results for education certainly give us some pause: in column (4) and column (6), the coefficient on Eskom project status is very similar to the main result for employment for women and men. The most likely explanation for such large increases in the fraction of (female) adults with a high school education in a five-year period is that high-skilled people are migrating towards communities that are being electrified by virtue of flatter gradient. If it is the case that high-skilled people are simply moving to flat areas differentially, then all the gradient instrument may be picking up is the increase in people with better job prospects or with existing jobs.

Ideally, it would be possible to isolate employment growth for in-migrants and incumbents. Although it is not possible to identify the migrant status of employed individuals in the Census, I can identify which individuals are in-migrants versus incumbents.[54] I use this information to implement a type of bounding test for whether in-migrants are driving the entire employment response. First, I assume that all individuals who report themselves as recent in-migrants (in-migrants in the last 5 years) are employed. Then, I redefine the dependent variable by excluding this total number of recent in-migrants from the numerator of the employment rate in 1996 and in 2001. The new employment rate variable is therefore the most conservative measure of changes in employment rates for incumbents only. Re-estimating the main results of the paper and comparing these results to those in Tables 5 and 6 inform us about the maximum contribution of in-migration (regardless of skill-level) to the total employment result.[55]

---

in the 10 percent sample. Results available from the author on request.

[53]Ideally, we could test directly whether gradient predicted out-migration from communities in my sample rather than using the aggregated district-level data. Unfortunately, the community data do not contain information on prior place of residence.

[54]This is because the Census community data do not allow all possible cross tabulations of variables.

[55]The in-migration data are far from perfect. Individuals are asked "Were you living in this place 5 years previously?" While this leaves room for a wide interpretation of 'this place', the exercise is still useful as a significant difference in results

Table 13 provides employment results separately for men and women in columns (7) to (10). For women, the OLS and IV results are remarkably similar across the full definition of employment and the migrant-excluded definition. Using a Hausman test to compare the electrification coefficient for female employment defined in the original way and in this more conservative manner, I cannot reject the hypothesis that these coefficients are the same. This is the case for both OLS and the IV specifications. Differential in-migration of employed women, high-skilled or not, cannot account for the entire female employment effect. In contrast, for men, redefining the employment variable in this way raises the IV point estimate a great deal (although this is still not significant). Male employment is 8.2 percentage points higher in electrified communities; although again, since there is no reduced form relationship between gradient and changes in the fraction of educated men, it is difficult to know what to make of these second stage estimates.

Within the limitations of the aggregate Census data then, there is some evidence in columns (1) to (3) that in-migration of individuals towards newly electrified areas may be an additional response to electricity projects. Given reported preferences for household services, this effect is not surprising: in a recent household survey conducted in a different rural part of the country (also an ex-homeland area) individuals ranked electricity in the home as the second most important service, after water.[56] However, the main employment results are robust to a focus on incumbents.

8 Conclusions

This paper uses a period of household electrification in South Africa to do two things: measure the direct effects of public infrastructure on employment in rural labor markets and uncover some of the mechanisms through which these effects operate. I combine hand-collected administrative and spatial data on electricity project roll-out with aggregate Census data to estimate large increases in the use of electric lighting and cooking, and reductions in wood-fueled cooking over a five-year period. Results from the main analysis indicate that female employment rises by 9.5 percentage points in electrified areas while there are no significant effects for male employment; complementary evidence from household-level surveys points towards increases in employment on the intensive margin for both women and men. The female employment response in the Census is driven by middle-poor and second-richest communities that initially rely heavily on wood for cooking and are able to respond more when the new service becomes accessible. Effects are also larger for women in their thirties and forties, who are much less likely to be living with young children requiring constant care. Controlling for magisterial district fixed effects and trends, female wages fall in areas that are experiencing larger increases in electrification while there are no statistically significant differences in male wages in the same areas. While not altogether satisfying in terms of precision, the MD-level analysis uses a different identification strategy and a different source of variation in electrification

would indicate a substantial in-migration response to electrification.

[56]Study conduced by Fort Hare Institute of Social and Economic Research (2007).

roll-out to show similar employment responses to the IV strategy and provides additional evidence that electrification did not appear to stimulate large increases in the demand for labor. While electrification of households did change the technology of home production and likely had an effect on female labor supply (as evidenced by falling female wages in electrifying areas), my results also suggest that electricity probably changed the types of market activities in which people could engage. The fact that similar employment effects for men and women cannot be rejected in either the Census data or the household survey data suggests that electrification does not exclusively operate on rural labor markets through the mechanisms of releasing time from home production.

These results represent some of the first pieces of evidence on the impact of infrastructure for rural electrification on labor markets in a developing country.[57] Regardless of the mechanism, electrification enabled South Africans living in rural areas to increase their participation in modern labor markets. More generally, the paper highlights the importance of measuring employment effects in infrastructure evaluations. I also emphasize the importance of interpreting the measured effects of electrification within the context of existing economic conditions - in the case of South Africa, economic adjustments after the end of apartheid. This is an emphasis that is present in much of the historical literature on changes in women's labor force participation: in that literature, women's responses to changing constraints on their ability to work are most often interpreted within the context of a restructuring economy.[58]

Using new data and instrumental variables methods, this paper also provides an example of how we might study other networked infrastructure roll-outs that are inherently difficult to randomize. Collecting project and spatial data from implementing agencies is often feasible, and may generate more actual variation than legal changes would, in institutionally weak environments. Combining empirical approaches and data sources is also potentially useful for dealing with the multiple sources of bias that make it challenging to identifying the effects of infrastructure and for unpacking some of the mechanisms through which infrastructure affects labor markets.

Finally, one of the last results in the paper highlight how migration presents both challenges and opportunities for research into the effects of infrastructure roll-out. Although migration is a potential confounder of labor market effects, it was possible to use the South African data to show that the effects of electrification raised employment of incumbents, separately from any migrant responses. This was a useful exercise for the specific research question of this paper. However, the significant increases in population in electrifying areas certainly raise additional questions about the role that infrastructure-building might play in transforming rural communities into more urban entities. Addressing such questions successfully is likely to require more of a general equilibrium approach than I have presented in this work.

[57] Grogan and Sadanand (2008) investigate the impact of electrification on a range of outcomes in Guatemala..

[58] For example, effects of new fertility-control technologies or falling appliance prices in the US have been analyzed over periods characterized by World Wars, changing social norms and alterations in the structure of jobs available for women. See for example, Bailey (2006), Goldin and Katz (2000) and Greenwood et al (2005).

References

Akee, Richard, "Road to Palau," 2006. IZA Working Paper No. 2452.

Aschauer, David Alan, "Is public expenditure productive?," *Journal of Monetary Economics*, 1989, *23*, 177–200.

Bailey, Martha, "More power to the pill: The impact of contraceptive freedom on women's labor supply," *Quarterly Journal of Economics*, February 2006, (121), 289–320.

__ **and William Collins**, "The economic and demographic effects of household electrification,1920 to 1960," March 2006.
http : //hubcap.clemson.edu/ sauerr/seminarpapers/BaileyCollinsClemson.pdf.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "Railroad to success: The effect of infrastructure on economic growth," 2007. Mimeo, Department of Economics, MIT.

__ **, Sebastian Galiani, James Levinsohn, and Ingrid Woolard**, "Why has unemployment risen in the new South Africa?," 2007. NBER Working Paper No. 13167.

Barnard, Jenny, "An Investigation into Using GIS in Electrification and Network Planning in Rural Kwazulu-Natal," Technical Report, GISDevelopment.net
*http : //www.gisdevelopment.net/application/utility/power/maf*0630.*htm* 2006.

Becker, Gary, "A Theory of the Allocation of Time," *Economic Journal*, 1965, *75* (1), 493–517.

Black, Dan, Terra McKinnish, and Seth Sanders, "The economic impact of the coal boom and bust," *The Economic Journal*, April 2005, (115), 449–476.

Bound, John and Gary Solon, "Double Trouble: on the value of twins-based estimation of the returns to schooling," *Economics of Education Review*, 1999, *18*, 169–182.

Budlender, Debbie, Ntebaleng Chobokoane, and Yandiswa Mpetsheni, "A survey of time use: How South Africans spend their time," Technical Report, Statistics South Africa
*http : //new.hst.org.za/pubs/index.php/*442/ 2001.

Canning, David, "A Database of World Infrastructure Stocks," 1998. World Bank Policy Research Working Paper No. 1929.

Card, David, "Using geographic variation in college proximity to estimate the returns to schooling," in L.N. Christofides et al, ed., *Aspects of labour market behaviour: Essays in honor of John Vanderkamp*, Toronto: University of Toronto Press, 1995, pp. 201–221.

Case, Anne and Angus Deaton, "Large cash tranfers to the elderly in South Africa," *Economic Journal*, September 1998, *108* (450), 1330–1361.

Cattaneo, Matias, Sebastian Galiani, Paul Gertler, Sebastian Martinez, and Rocio Titiunik, "Housing, health and happiness," 2007. World Bank Policy Research Working Paper No. 4214.

Charmes, Jacques, "Chapter 3," in Mark Blackden and Quentin Wodon, eds., *Gender, time use and poverty in sub-Saharan Africa*, Washington, D.C.: The World Bank, 2005.

Chernozhukov, Victor and Christian Hansen, "The reduced form: A simple approach to weak inference," 2007. University of Chicago, Graduate School of Business *http://faculty.chicagogsb.edu/christian.hansen/research/chweakivmar07.pdf*.

Christopher, A.J., *The Geography of Apartheid*, Johannesburg: Witwatersrand University Press, 2001.

Coen-Pirani, Daniel, Alexis Leon, and Steven Lugauer, "The effect of household appliances on female labor force participation: Evidence from micro data," March 2008. Carnegie Mellon University *http://ideas.repec.org/p/cmu/gsiawp/407427706.html*.

Conley, Timothy, "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, September 1999, *92* (1), 1–45.

Costa, Dora, "From mill town to boardroom: The rise of women's paid labor," *Journal of Economic Perspectives*, 2000, *14* (4), 101122.

Cowan, Ruth Schwartz, *More work for mother: The ironies of household technology from the open hearth to the microwave*, Basic Books, 1985.

Cross, Catherine, Tobias Mngadi, and Themba Mbhele, "Infrastructure, poverty and development in KwaZulu-Natal," *Development Southern Africa*, Summer 1998, *15* (4).

Cutler, Grant and David Miller, "The role of public health improvements in health advances: The twentieth-century United States," *Demography*, February 2005, *42* (1), 1–22.

Deaton, Angus and Shankar Subramanian, "The demand for food and calories," *Journal of Political Economy*, 1996, *104* (1).

Dinkelman, Taryn and Vimal Ranchhod, "Compliance with minimum wage laws when penalties and enforcement are weak: Domestic workers in South Africa," 2008. Mimeo, University of Michigan *http://www.personal.umich.edu/ tdinkelm*.

Duflo, Esther and Rohini Pande, "Dams," *Quarterly Journal of Economics*, May 2007, *122* (2), 601–646.

___ , Michael Greenstone, and Rema Hanna, "Indoor Air Pollution, Health and Economic Well-being," February 2008. Mimeo, MIT *http://econ − www.mit.edu/faculty/eduflo/papers*.

EnergyNet Limited, "Rural Energy 21: Bringing Power to Africa," 2004. *http://www.energynet.co.uk/Infosite/images/RE21report2004.pdf*.

Eskom, "Eskom Annual Report," Technical Report, Eskom 1996.

___ , "Eskom Annual Report," Technical Report, Eskom 1997.

___ , "Eskom Annual Report," Technical Report, Eskom 1998.

___ , "Eskom Annual Report," Technical Report, Eskom 1999.

Fedderke, Johannes and Zeljko Bogetic, "Infrastructure and growth in South Africa: Direct and indirect productivity impacts of nineteen infrastructure measures," 2006. World Bank Working Paper No. 3989.

Fernandez, Raquel, Alessandra Fogli, and Claudia Olivetti, "Marrying your mom: Preference transmission and women's labor and education choices," September 2002. NBER Working Paper No. 9234.

Field, Erica, "Entitled to work: Urban property rights and labor supply in Peru," *Quarterly*

Journal of Economics, November 2007, *122* (4), 1561–1602.

Food and Agriculture Organisation, "Topological classifications," Technical Report, FAO 1998.

Fort Hare Institute of Social and Economic Research, "Rapid assessment of service delivery and socio-economic survey in the Eastern Cape," Technical Report, University of Fort Hare, Development Research Africa, Take Note Trading 2006.

Galor, Oded and David N. Weil, "The gender gap, fertility, and growth," *American Economic Review*, June 1996, *86* (3), 374–387.

Gaunt, Trevor, "Electrification technology and processes to meet economic and social objectives in Southern Africa." PhD dissertation, University of Cape Town 2003.

Gelbach, Jonah, "Public schooling for young children and maternal labor supply," *American Economic Review*, March 2002, *92* (1).

Goldin, Claudia, "The U-Shaped female labor force function in economic development and economic history," April 1996. NBER Working Paper No. 4707.

__ **and Laurence Katz**, "Career and marriage in the age of the pill," *American Economic Review Papers and Proceedings of the One Hundred Twelfth Annual Meeting of the American Economic Association*, 2000, (XC), 461–465.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu, "Engines of Liberation," *Review of Economic Studies*, 2005, *72*, 109–133.

Grogan, Louise and A. Sadanand, "Electrification and the household," 2008. University of Guelph *http://www.economics.uoguelph.ca/lgrogan/electrifhh.pdf*.

Gronau, Reuben, "Home production: A survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of Labor Economics*, Vol. I, Elsevier Science Publishers, 1986.

Houthakker, H.S., "An international comparison of household expenditure patterns, commemorating the centenary of Engel's law," *Econometrica*, 1957, *25* (4), 532–551.

Jimenez, Emmanuel, "Human and physical infrastructure: public investment and pricing policies in developing countries," in Jere Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics*, Vol. 3B, Elsevier Science Publishers, 1995, chapter 43, pp. 2774–2836.

Johnston, A.M. and R.W. Johnston, "The local elections in KwaZulu-Natal: 26 June 1996," *African Affairs*, 1997, *96* (384), 377–398.

Kane, Thomas, Cecilia Elena Rouse, and Douglas Staiger, "Estimating returns to schooling when schooling is misreported," 1998. NBER Working Paper No. 7235.

Khan, Sultan, Benoit Lootvoet, and Shahid Vawda, "Transcending traditional forms of governance: Prospects for co-operative governance and service delivery in Durban's tribal authority areas," *Transformation*, 2006, *62*.

Kling, Jeffrey, "Intepreting instrumental variables estimates of the returns to schooling," *Journal of Business and Economic Statistics*, July 2001.

Leibbrandt, Murray, Laura Poswell, Pranushka Naidoo, Matthew Welch, and Ingrid Woolard, "Measuring recent changes in South African inequality and poverty using 1996 and

2001 Census data," 2003. Center for Social Science Research Working Paper 05-94.

Loshkin, Michael and Ruslan Yemtsov, "Has rural infrastructure rehabilitation in Georgia helped the poor?," *World Bank Economic Review*, August 2005, *19* (2), 311–333.

Mammen, Kristin and Christina Paxson, "Women's work and economic development," *Journal of Economic Perspectives*, 2000, *14* (4), 141–164.

Mikusheva, Anna and Brian Poi, "Tests and confidence sets with correct size in the simultaneous equations model with potentially weak instruments," *The Stata Journal*, 2006, (1), 1–11.

Moreira, Marcelo and Luiz M. Cruz, "On the validity of econometric techniques with weak instruments: Inference on returns to education using compulsory school attendance laws," *Journal of Human Resources*, 2005, *40* (2), 393–410.

Nunn, Nathan and Diego Puga, "Ruggedness: The blessing of bad geography in Africa," March 2007. Harvard University *http://diegopuga.org/papers/rugged.pdf*.

Piper, Laurence, "Democracy for a bargain: The 1999 election in KwaZulu-Natal," *Politikon*, 1999, *26* (2), 145–154.

Posel, Dori and Daniela Casale, "'Two million net new jobs': A reconsideration of the rise in employment in South Africa, 1995-2003," *South African Journal of Economics*, 2004, *72* (5).

__ **, James A. Fairburn, and Frances Lund**, "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa," *Economic Modelling*, 2006, *23*, 836–853.

Qian, Nancy, "Missing women and the price of tea in China: The effect of relative female income on sex imbalance," 2006. Brown University *www.chass.utoronto.ca/ brandt/ECO2703/papers/Qian*$_2$*006*$_M$*issing*$_m$*imeo.pdf*.

Rosenzweig, Mark, Mark Pitt, and Nazmul Hassan, "Sharing the burden of disease: Gender, the household division of labor and the health effects of indoor air pollution in Bangladesh and India," March 2005. CID Working Paper No. 119 *http://www.cid.harvard.edu/cidwp/119.htm*.

Rud, Juan Pablo, "Electricity provision and industrial development: Evidence from India," 2009. Memo, LSE.

Saghir, Jamal, "Energy and Poverty: Myths, Links and Poverty Issues," Technical Report, The World Bank *http://esa.un.org/un − energy/pdf/UN − ENRG/20paper.pdf* 2005. Energy Working Notes No 4.

Simkins, Charles, "Agricultural production in the African reserves of South Africa, 1918-1969," *Journal of Southern African Studies*, April 1981, *7* (2), 256–283.

Simonsen, Marianne, "Availability and price of high quality day care and female employment," May 2005. Aarhus University Economics Paper No. 2005-08.

South African Department of Minerals and Energy, "Electricity Supply Guide," Technical Report, Department of Minerals and Energy 2004.

Standing, Guy, John Sender, and Jeremy Weeks, *Restructuring the Labour Market: The*

South African Challenge, Geneva: International Labour Office, 1996.

Statistics South Africa, *Census Mapping Manual 2000.*

The World Bank, *World Development Report 1994: Infrastructure for Development*, The World Bank, 1994.

__ , "Clean energy for development investment framework: The World Bank Group action plan," Technical Report, The World Bank 2007.

Udry, Christopher, "Gender, agricultural production, and the theory of the household," *Journal of Political Economy*, October 1996, *104* (5), 1010–46.

United States Department of Energy, "Average retail price of electricity to ultimate customers by end-use sector," Technical Report, Department of Energy
http://www.eia.doe.gov/cneaf/electricity/epa/epat7p4.html 2007.

van der Berg, Servaas, Rulof Burger, Murray Leibbrandt, and Cecil Mlatsheni, "Migration and the changing rural-urban interface in South Africa: What can we learn from census and survey data?," 2003. South African Labour and Development Research Unit
http://www.csae.ox.ac.uk/conferences/2002-UPaGiSSA/papers/Leibbrandt-csae2002.pdf.

West, Nicola, Barry Dwolatzky, and Alan Meyer, "Terrain-based routing of distribution cables," *IEEE Computer Applications in Power*, January 1997.

Appendix 1: Data

Census data

Census community data 1996 and 2001: Available from Statistics South Africa (at www.statssa.gov.za). Proprietary software enables extraction of community totals for various combinations of variables at enumeration area (in 1996) or sub-place (2001) level, including: counts of employment, population, and levels of educational attainment by sex and age group; counts of households, female-headed households, and households living beneath a poverty line; counts of households using different sources of fuel for lighting. In response to a special request, Statistics South Africa also provided counts of households using different fuel-sources for cooking at the enumeration area (1996) and sub-place level (2001).

Employment variables in the Census: As in most Census data, the measures of employment are broad. In 1996, all adults are asked: 'Does the person work?' Activities listed as work included: formal work for a salary or wage, informal work such as making things for sale or selling things or rendering a service, work on a farm or the land, whether for a wage or as part of the household's farming activities. In 2001, adults were asked: 'Did the person do any work for pay, profit or family gain for one hour or more?' Possible responses were: yes (formal, registered, non-farming), yes (informal, unregistered, non-farming), yes (farming) and no (did not have work).

Census panel of communities: The 2001 Census geography is ordered hierarchically as follows, from largest to smallest unit:

- District: this represents a local labor market area in KwaZulu-Natal and contains between 30,000 and 50,000 households.
- Main place or sub-district: these correspond to groupings of towns and surrounding areas.
- Community or sub-place: this is the lowest unit of observation in the 2001 Census data. Average community size is small: between 200 and 250 households on average.

Boundaries for communities from the 2001 Census define the main unit of analysis. Since boundaries have shifted over time (Christopher, 2001), the 1996 (smaller) areas are aggregated up to the (larger) 2001 boundaries.[59] The matched identifiers from this panel of areas are used to extract Census aggregate data in 1996 and 2001. For each 1996 EA, the proportion of the EA polygon area that falls inside each 2001 community is calculated. This proportion is used as a weight to assign a proportion of the 1996 EA data to the 2001 community. The key assumption in this process is that people are uniformly distributed over 1996 EA's.

[59]Statistics South Africa notes that EA boundaries should never cut across existing administrative boundaries, and all "social boundaries should be respected" (StatsSA, 2000). In most cases, re-demarcation involved the following real changes to 1996 EA's: "splits" that occurred when obstacles or boundaries divided the EA naturally, and "merges" that occurred between EA's that were small or that were legally, socially or naturally a geographical entity. Changes were made only when "absolutely necessary" (StatsSA, 2000: 21, 26).

Census Micro data 1996 and 2001 - 10% sample: Available at: www.statssa.gov.za. This is a 10% sample of the population Census conducted in 1996 and 2001. Observations are at the individual level and can only be assigned to district boundaries for confidentiality reasons.

Household Surveys: 1995, 1997, 1999, 2001

Available at: www.statssa.gov.za. These four waves of household survey data (October Household Surveys for the 1990s and the September Labor Force Survey in 2001), resembling the World Bank LSMS surveys, are collected by Statistics South Africa and provide a nationally representative sample of individuals. I use the sample of African male and female adults (15-59) living in rural KwaZulu-Natal who report information about on employment (or lack thereof) as well as hours of work and total monthly earnings. The lowest level of geography that can be identified in these household surveys is the magisterial district, of which there are 42 in rural KZN.

Schools Register of Needs

These data are provided by the South African Department of Education for school in 1995 and 2000. GPS co-ordinates for each school are used to assign schools to Census community areas. Each community is assigned the total number of schools in each year as well as the change in the number of schools over the five-year period.

Geographic data

Land gradient: The source for these data is the 90-meter Shuttle Radar Topography Mission (SRTM) Global Digital Elevation Model available at www.landcover.org. Digital elevation model data was used to construct measures of average land gradient for each Census community using GIS software (ArcMap 9.1). Gradient is measured in degrees from 0 (perfectly flat) to 90 degrees (perfectly vertical).

Other measures of proximity: Spatial data on Eskom's 1996 grid network (high and medium voltage lines and substations) was provided by Steven Tait. These data were used to calculate straight line distances between Census centroids and the nearest electricity substation.

Census 1996 spatial data were used to generate straight line distances from each community centroid to the nearest road and town.

Electricity project data

Data on Eskom projects in KwaZulu-Natal were provided by Sheila Brown. The project list details the number of pre-paid electricity connections per Eskom area in each year from 1990 to 2007. The year of electrification is defined as the year in which a community experienced a spike in household connections (concentrated project activity). Areas are referenced by name and village code. Eskom's planning units do not line up accurately with Census regions. To match project data to Census regions, the project data were first mapped to a physical location (using a spatial database of transformer codes that corresponded to project codes) and these locations were then merged back to Census spatial boundaries.

Appendix 2: Measurement error

Measurement error in the Census data: Employment

The Census data undoubtedly measure employment with some error. While the employment questions are broad, the Census does not probe for employment information as the household surveys do. This section discusses the extent of this measurement error by comparing the Census data to individual level household survey data.

In Appendix Table 1, I present population totals and employment rates for six different surveys: columns (1) and (4) present household-level data from the 1996 October Household Survey and the 2001 September Labor Force Surveys. These are the closest surveys we have to the relevant Census years and I use the weights in these surveys (constructed using the relevant Census as a benchmark) to create population totals and employment rates. In columns (2) and (4), I use the micro data from the 10% sample of the Census in 1996 and 2001 to create the same statistics using the Census weights; and in columns (3) and (6) I present the statistics taken from the 100% Census community databases in 1996 and 2001. Note that the unit of observation is the individual in columns (1), (2), (4) and (5) and the community in columns (3) and (6). Another important difference is that the individual level data in the Census and household surveys can only be restricted to African adults living in rural KZN while the Census community data can be disaggregated further to include adult Africans living in tribal areas of KZN. Tribal areas refer to the former homelands.

The individual Census and individual household survey data provide population totals and employment rates that are not substantially different from each other in most cases. Employment rates and population totals for women and men in 1996 are fairly close to each other in the household survey and the individual Census data; however, in 2001, there is a larger difference between the household surveys and the individual Census data, with the household surveys picking up a higher rate of employment for both men and women. The information in this table suggests that the Census in 2001 may have under-counted employment more than in 1996, for both men and women.

In addition to these differences over time in how closely the individual level data correspond, there are larger differences between the community data and the individual data. In every year, for men and for women, the Census community data present substantially lower population totals for men and women as well as lower employment rates: between one half and one third of the employment measured in the individual data. A large part of the explanation for this is that the community data are restricted to tribal areas, which do not include all rural areas in the province. Hence individuals who live in rural communities with better average labor market outcomes than individuals in the tribal, rural areas of the province are excluded from population counts and employment statistics based on the community level data.

Appendix Table 2 shows the correlations between the individual and community Census data for different years and for men and women at the magisterial district level. The first four columns show that the community census employment data predict only part of the individual census employment data and that the fraction explained for women is higher than it is for men. The final two columns show the correlation between the change in employment rates measured at the individual level and the change in employment rates measured at the community level. Again, more of the change in female employment at the individual level is predicted by the change in female employment in the community data than for men.

There are a few important points to note from the table:

- The Census community data that is restricted to tribal areas undercounts employment, relative to all rural areas.
- The 2001 Census data (both individual and community data) measure lower levels of employment compared to the household survey data. This is probably due to the way the Census asked about employment in 2001: "Did you work for at least 1 hour last week?" compared with the 1996 question, "Did you work for a formal wage/salary, in informal work, or on a farm last week?" The 2001 question may not have been interpreted to include informal sector work or farm work by respondents in 2001, so the main types of employment that are undercounted in 2001 are probably these types of jobs. As long as the prevalence of these jobs is uncorrelated with gradient, then undercounting of employment in the 2001 data should not be problematic.
- Changes in employment in the Census community data more strongly predict changes in employment in the individual level data for women compared to men. This suggests that the community level data may be missing more of the employment story for men, than for women, in these areas.
- Even though the Census community data undercounts employment, the strong message from the individual level data is that there are very low levels of employment in these rural areas: under 30% of men are employed and under 22% of women are employed. These employment rates fall even further when we restrict to tribal areas of the province in using the Census community data. The low levels of employment in these areas are not an artifact of mismeasurement.
- Finally, focusing on the occupation distribution for men and women, the individual Census data count fewer men employed in agriculture and slightly more women in agriculture in 1996 than the household surveys do. Agricultural employment in both of the individual sources is higher than in the Census community data - bearing in mind that the community data count people living only in tribal areas - and yet is still very low. Regardless of which data set is considered, there is an insubstantial fraction of men and women working in agriculture in the rural areas of KZN.

Measurement error in the electrification project variable

Since Eskom region boundaries do not line up with Census boundaries, I assign values of T_{jdt} in the following way: for any community that lies even partially inside an Eskom project area, all information from that project is assigned to that community. This means some communities are assigned full electrification status when only a fraction of households in the area are electrified. In addition, non-NEP electrification continued during this period in areas where households were willing to pay for their connections.

Measurement error in the binary project status variable could contribute to the difference between OLS and IV coefficients. OLS will underestimate the effect of electricity on outcomes when there is a negative covariance between δ_j and ΔT_{jdt} (which I have argued is likely) and when ΔT_{jdt} is measured with error. However, the valid IV that is uncorrelated with $\delta_j + \Delta\epsilon_{jt}$ will tend to be correlated with any non-classical measurement error in the binary variable ΔT_{jdt}. In this situation, even if the instrument deals with the omitted variables bias, the measurement error in ΔT_{jdt} could lead to an upwards-biased IV estimator.[60]

To get a sense of how much of the difference in OLS and IV results is due to measurement error, I restrict to samples where I expect ΔT_{jdt} to be measured with less error. The first two columns of Appendix Table 3 reproduce the main result for females in the full sample while columns (3) to (6) present results for successive sample limitations. To identify communities where projects had greater coverage, I exclude electrified areas with less than a 10 percent change in coverage of electric lighting, and areas where the connection rate between 1996 and 2001 was under 80 percent of households. All communities that did not have an electricity project during the period are included in all columns. Under the first restriction in columns (3) and (4), the OLS coefficient rises substantially and the IV coefficient is the same as the main result at 13 percentage points. The movement in the OLS coefficient suggests that there is some measurement error is present in the electrification variable. Columns (5) and (6) impose the second restriction. Again, the OLS estimate is large and positive and the IV result is now slightly higher than the main result (at 0.155), although neither is statistically significant due to the smaller sample size.

Although effects estimated under the OLS specification for these sub-samples are between 1 and 1.2 percentage points higher than the OLS result for the full sample, they are still well smaller than the IV results. This provides some evidence that measurement error in the electrification dummy alone is unable to account for the entire gap between OLS and IV estimates.

[60] This result is conditional on the measurement error in electrification status not being too extreme (Kane et al, 1998). See Bound and Solon (1999) and Kane, Rouse and Staiger (1998) for a discussion of what the IV estimator is consistent for in the presence of non-classical measurement error.

Appendix Table 1: Comparing measures of employment in the Census and October Household/Labor Force Surveys

	1996			2001		
	OHS 1996	10% Census	100% Census	LFS 2001	10% Census 2001	100% Census
Level of data	Individual, RURAL	Individual, RURAL	Community, TRIBAL	Individual, RURAL	Individual, RURAL	Community, TRIBAL
Panel A: Women						
Population totals	1,231,422	1,299,475	1,290,869	1,285,357	1,479,848	1,632,826
Total employment/population	0.148	0.130	0.072	0.217	0.130	0.076
Occupational distribution						
Managers	0.001	0.001	0.001	0.002	0.001	0.001
Professionals	0.002	0.013	0.015	0.003	0.003	0.016
Associates	0.022	0.003	0.002	0.020	0.014	0.005
Clerks	0.007	0.003	0.003	0.011	0.008	0.005
Services	0.014	0.005	0.003	0.025	0.007	0.003
Agriculture	0.005	0.010	0.000	0.008	0.008	0.000
Crafters	0.013	0.007	0.004	0.011	0.008	0.004
Machine Operators	0.003	0.004	0.002	0.013	0.005	0.002
Elementary Occupations	0.068	0.069	0.031	0.124	0.062	0.040
Missing occupations data	0.012	0.017	0.010	0.000	0.015	0.000
Panel B: Men						
Population totals	921,862	993,888	1,079,777	1,012,397	1,181,795	1,323,726
Total employment/population	0.282	0.250	0.149	0.263	0.216	0.116
Occupational distribution						
Managers	0.003	0.002	0.002	0.010	0.004	0.003
Professionals	0.004	0.009	0.012	0.000	0.004	0.013
Associates	0.020	0.003	0.003	0.013	0.010	0.006
Clerks	0.007	0.004	0.015	0.006	0.009	0.014
Services	0.045	0.021	0.007	0.026	0.021	0.005
Agriculture	0.011	0.020	0.000	0.022	0.016	0.000
Crafters	0.025	0.041	0.028	0.049	0.027	0.017
Machine Operators	0.052	0.035	0.021	0.060	0.037	0.020
Elementary Occupations	0.091	0.066	0.029	0.078	0.069	0.039
Missing occupations data	0.024	0.048	0.031	0.000	0.021	0.000

OHS is the October Household Survey; LFS is the Labor Force Survey, 10% Census are microdata; 100% Census are community aggregate data. OHS 1996 does not have a separate category for subsistence agriculture workers, so agriculture includes skilled and subsistence. Proportions from the OHS/LFS/10% Census data are weighted using population weights. Census Community data are weighted by the number of people in each community. All proportions are calculated over the sample of rural Africans living in KZN, aged 15-59 inclusive

Appendix Table 2: Correlation between community and individual Census employment data - OLS

	Female		Male			
	Census Individual, 2001	Census Individual, 1996	Census Individual, 2001	Census Individual, 1996	Δ Female employment, Individual	Δ Male employment, Individual
	(1)	(2)	(4)	(5)	(6)	(8)
Females, Community, 2001	0.62654***					
	(0.076)					
Females, Community, 1996		0.50758***				
		(0.077)				
Males, Community, 2001			0.28403***			
			(0.055)			
Males, Community, 1996				0.33985***		
				(0.044)		
Δ Female employment, Community					0.38448***	
					(0.088)	
Δ Male employment, Community						0.14004***
						(0.047)
N	42	42	42	42	42	41
R^2	0.55	0.44	0.39	0.33	0.17	0.01

*** p<0.01, ** p<0.05, * p<0.1, robust standard errors in parentheses. Census, Individual refers to employment rates computed using the 1996 or 2001 Census microdata. Census, Community refers to employment rates computed using the 1996 or 2001 Census community data. Δ variables refer to change in the employment rate over time. The unit of observation is the magisterial district.

Appendix Table 3: Contribution of measurement error in electrification status to female employment result

Outcome is Δ_t in female employment	Full sample		Restricted to areas with > 10% change in electricity coverage		Restricted to areas with over 80% coverage by 2001	
	OLS (1)	IV (2)	OLS (5)	IV (6)	OLS (7)	IV (8)
Eskom Project	-0.001	0.095*	0.009	0.095	0.011	0.082
	(0.005)	(0.055)	(0.007)	(0.060)	(0.009)	(0.087)
N	1,816	1,816	1,461	1,461	1,273	1,273

Dependent variable is change in employment rate of African females aged 15-59. Each coefficient (standard error) is from a separate regression that controls for all covariates. Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level. Electrified is 1 if community had first Eskom project between 1996 and 2001, else 0. Columns (1) and (2) replicate the coefficient on Eskom Project from Table 5. Columns (3) and (4) restrict the sample to all non-electrified areas and electrified areas with a 10% or larger change in fraction of households using electric lighting. Columns (5) and (6) restrict the sample to all non-electrified areas and electrified areas in which Eskom connected at least 80% of households between 1996 and 2001. All control variables (except electrification status and change in access to other services) are measured in 1996: distance to the grid, household density, community poverty rate, adult sex ratio (F/M), fraction of female-headed households, share of Indian/white adults, distance to nearest road, distance to nearest town, share of adult men and women with at least high school. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression.

Table 1: Baseline (1996) community covariates by Electrification Project status and by gradient

Covariates	All	Eskom Project	No Eskom Project	Difference (3) - (2)	Difference by gradient: Without controls	Difference by gradient: With controls
	(1)	(2)	(3)	(4)	(5)	(6)
Poverty rate	0.61	0.59	0.61	-0.024**	0.00	0.002
	(0.19)	(0.17)	(0.20)	(0.01)	(0.00)	(0.001)
Fracton female-headed households	0.55	0.55	0.55	0.00	0.005***	0.001
	(0.13)	(0.12)	(0.13)	(0.01)	(0.00)	(0.001)
Adult sex ratio (f/m)	1.48	1.41	1.49	-0.080***	0.011***	0.004**
	(0.28)	(0.25)	(0.29)	(0.02)	(0.00)	(0.002)
Proportion Indian and White adults*10	0.00	0.00	0.00	0.00	0.00	0.000
	(0.01)	(0.00)	(0.01)	(0.00)	(0.000)	(0.000)
Kilometers to road	37.95	35.62	38.54	-2.917**	-0.201*	-0.156
	(24.57)	(24.18)	(24.64)	(1.44)	(0.12)	(0.184)
Kilometers to town	38.57	36.34	39.13	-2.790***	0.278***	0.180
	(18.12)	(15.34)	(18.72)	(1.06)	(0.09)	(0.130)
Fraction men with high school	0.06	0.08	0.06	0.016***	-0.002***	-0.003**
	(0.05)	(0.05)	(0.05)	(0.00)	(0.000)	(0.000)
Fraction women with high school	0.07	0.08	0.06	0.020***	-0.002***	0.000
	(0.05)	(0.05)	(0.05)	(0.00)	(0.000)	(0.000)
Households per km^2	22.05	32.56	19.41	13.152***	-0.523***	-0.945***
	(30.48)	(49.31)	(22.75)	(1.76)	(0.15)	(0.301)
Kilometers from the grid	19.06	15.75	19.89	-4.139***	-0.235***	0.029
	(13.32)	(10.20)	(13.88)	(0.77)	(0.06)	(0.125)
Land gradient - mean	10.10	9.12	10.35	-1.232***		
	(4.89)	(4.21)	(5.02)	(0.29)		
Land gradient - std. dev.	5.35	4.92	5.46	-0.540***	0.435***	0.661***
	(2.42)	(2.25)	(2.45)	(0.14)	(0.006)	(0.025)
Land gradient - range	28.50	26.52	28.99	-2.476***	2.010***	2.494***
	(11.85)	(11.30)	(11.93)	(0.69)	(0.032)	(0.103)
Min p-value [for joint test]				0.000	0.002	0.002
N communities	1,816	365	1,451	1,816	1,816	1,816

Sample contains the set of tribal KwaZulu-Natal (KZN) communities not electrified before 1996; Eskom project is 1 if first Eskom project occurred between 1996 and 2001, else 0. Communities with fewer than 100 adults in either year are excluded. Community-level means (s.d.) in columns (1)-(3). Differences in columns (4)-(6) significant at p<0.01***, p<0.05** or p<0.1* level. Columns (5) and (6) show coefficients from regressions of each covariate on gradient, controlling for all other covariates and district fixed effects (in column (6)). Community-level variables are measured in 1996. Poverty rate is fraction of households earning below ZAR6,000 per year. Sex ratio is numbe African females (aged 15-59) over number of African males (aged 15-59). Female-headed households expressed as fraction of all households. Number of Indian and white adults expressed as a fraction of all adults. Distances to nearest road, town, sub-station are straight-line kilometer distances from community centroid to nearest object. African men and women with at least completed high school as a share of all African men or women. Household density is per square kilometer. Land gradient statistics created in ARCMAP at the community level. With these ten covariates (including household poverty to household density), the Bonferroni joint test of significance requires p<0.05/10=0.005 to reject the null of all coefficients zero at a 5% level of significance; or p<0.01/10=0.001 to reject at the 1% level.

Table 2: Average community-level outcomes in 1996 (before) and 2001 (after)

	Year	Mean	Min	Max	Eskom project	No Eskom project	Difference (5)-(4)
		(1)	(2)	(3)	(4)	(5)	(6)
Employment rates							
Female Employment Rate	1996	0.07	0.00	0.87	0.09	0.06	0.020***
		(0.08)			(0.07)	(0.08)	(0.004)
	2001	0.07	0.00	0.69	0.08	0.06	0.017***
		(0.07)			(0.07)	(0.07)	(0.004)
	Δ_t	0.000			-0.003	0.001	-0.004
		(0.002)			(0.005)	(0.00)	(0.004)
Male Employment Rate	1996	0.14	0.00	0.98	0.16	0.13	0.031***
		(0.11)			(0.11)	(0.11)	(0.007)
	2001	0.10	0.00	0.72	0.11	0.10	0.014**
		(0.09)			(0.09)	(0.09)	(0.005)
	Δ_τ	-0.04***			-0.050***	-0.033***	-0.017***
		(0.00)			(0.01)	(0.00)	(0.006)
Population totals							
N Adult Females	1996	356.07	48	4,553	453.80	367.53	86.28***
		(347.84)			(377.53)	(341.29)	(20.43)
	2001	446.04	42	3,392	575.26	413.54	161.72***
		(407.31)			(498.74)	(374.11)	(23.55)
	Δ_t	89.97***			121.46***	46.01***	75.45***
		(12.61)			(32.74)	(13.29)	(17.82)
N Adult Males	1996	273.56	32	3,136	332.87	258.64	74.23***
		(264.87)			(290.63)	(255.93)	(15.42)
	2001	328.59	30	2,770	439.37	300.72	138.65***
		(319.36)			(407.11)	(286.70)	(18.42)
	Δ_t	55.03***			106.50***	42.08***	64.42***
		(9.74)			(26.18)	(10.09)	(14.05)
N		1,816			365	1,451	

Columns (1), (4) and (5) contain variable means and standard deviations (s.d.). Mean differences and standard errors are shown for the difference (5)-(4) in column (6) and for Δt rows. Eskom project is =1 if first Eskom project occurred between 1996 and 2001, else 0. All variables constructed for Africans adults (ages 15-59) only. Sample excludes areas with fewer than 100 adults in either year.

Table 3: First stage assignment to Eskom Project: OLS results

	Outcome is Eskom Project = 1				
	(1)	(2)	(3)	(4)	(5)
Gradient*10	-0.083**	-0.075**	-0.078***	-0.077***	-0.100***
	(0.040)	(0.034)	(0.027)	(0.027)	(0.036)
Kilometers to grid*10		-0.040*	-0.012	-0.011	0.025
		(0.021)	(0.023)	(0.023)	(0.025)
Household density*10		0.017***	0.012**	0.013**	0.012
		(0.004)	(0.006)	(0.006)	(0.012)
Poverty rate		0.023	0.019	0.017	0.054
		(0.069)	(0.070)	(0.069)	(0.076)
Adult sex ratio (f/m)		0.393***	0.165	0.155	0.082
		(0.120)	(0.107)	(0.107)	(0.114)
Female-headed hh's		-0.173***	-0.130***	-0.121***	-0.072
		(0.052)	(0.042)	(0.042)	(0.056)
Fraction of Indian and white adults*10		-1.236***	-1.116**	-1.105**	53.057
		(0.401)	(0.459)	(0.452)	(61.000)
Kilometers to road*10		0.003	-0.010	-0.010	-0.003
		(0.009)	(0.010)	(0.010)	(0.009)
Kilometers to town*10		0.016	0.008	0.008	-0.023
		(0.015)	(0.015)	(0.016)	(0.017)
Men with high school		-0.269	-0.185	-0.152	-0.276
		(0.500)	(0.411)	(0.417)	(0.535)
Women with high school		1.046**	0.965**	0.984**	1.368**
		(0.475)	(0.413)	(0.409)	(0.583)
Change in water access				0.012	0.111
				(0.048)	(0.073)
Change in toilet access				0.155	0.535**
				(0.104)	(0.269)
District FE	N	N	Y	Y	Y
Sample	All	All	All	All	No electricity in 1996
Mean of Y variable	0.201	0.201	0.201	0.201	0.110
N	1,816	1,816	1,816	1,816	418
R^2	0.010	0.074	0.177	0.178	0.199
F-stat: gradient	4.201	4.870	8.336	8.257	7.835
Prob>F:	0.041	0.028	0.004	0.004	0.006

Sample is restricted to communities with at least 100 adults in both 1996 and 2001. Robust standard errors clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level. Dependent variable uses indicator for Eskom Project status (1 if the area had a project in between 1996 and 2001). Ten district fixed-effects included in columns (3) to (5). Land gradient in degrees, all distances measured in kilometers. Sample in column (5) is restricted to set of areas where no households had electricity in 1996. All control variables measured in 1996, except change in access to water and flush toilet measured between 1996 and 2001. Change in water captures change in fraction of households with access to water in the house or less than 200 meters away; change in toilets captures the change in fraction of households with a flush toilet.

Table 4 Effects of electricity projects on household energy sources and other services

		OLS			IV	
Outcome is Δ_t	Mean of dependent variable	No controls	Controls	Reduced form coefficient on gradient*10	No controls	Controls
	(1)	(2)	(3)	(4)	(5)	(6)
(1) Lighting with electricity	0.800	0.251***	0.239***	-0.008***	0.577***	0.658***
		(0.032)	(0.031)	(0.002)	(0.188)	(0.144)
(2) Cooking with wood	-0.036	-0.045***	-0.039***	0.002**	-0.266	-0.275*
		(0.012)	(0.012)	(0.001)	(0.179)	(0.147)
(3) Cooking with electricity	0.037	0.068***	0.056***	-0.002***	0.250**	0.228**
		(0.009)	(0.009)	(0.001)	(0.107)	(0.101)
(4) Water nearby	0.007	-0.029	0.005	0.003	-0.483*	-0.372
		(0.029)	(0.024)	(0.002)	(0.249)	(0.248)
(4) Flush toilet	0.030	0.003	0.008	-0.001	0.018	0.067
		(0.006)	(0.005)	(0.000)	(0.069)	(0.068)

Each cell contains the coefficient on Eskom Project indicator (robust standard errors clustered at sub-district level) from OLS regressions of dependent variable on electrification dummy and all explanatory variables. Significant at p<0.01***, p<0.05** or p<0.1* level. Column (1) contains the mean value of each dependent variable, the change in the fraction of households using electricity for lighting or cooking, using wood for cooking, or having access to nearby water or adequate toilet facilities. Electrified is 1 if the first Eskom project occurred between 1996 and 2001, otherwise 0. Excluded instrument is mean community gradient. Other control variables included in each regression: distance to grid, household density, community poverty rate, adult sex ratio (F/M), share of female-headed households, share of Indian/white adults, distance to nearest road, distance to nearest town, share of adult men and women with at least high school, change in fraction of households with water close by, change in proportion of households with flush toilets and ten district fixed-effects. Change in water (toilet) access excluded from controls in rows (4) and (5). Each regression contains N=1,816 except for change in fraction of households using wood; I set 9 observations to missing (rather than to zero) for 2001 observations.

Table 5: Effects of electrification on female employment

Outcome is Δ_t female employment rate	OLS				Reduced form	IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Eskom Project	-0.004	-0.001	0.000	-0.001		0.025	0.074	0.090*	0.095*
	(0.005)	(0.005)	(0.005)	(0.005)		(0.045)	(0.060)	(0.055)	(0.055)
Kilometers to grid *10		0.002	0.000	0.001	0.001		0.005	0.001	0.002
		(0.002)	(0.002)	(0.002)	(0.002)		(0.003)	(0.003)	(0.003)
Household density *10		0.000	0.000	0.000	0.000		-0.002	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Poverty rate		0.029***	0.033***	0.031***	0.032***		0.027**	0.032**	0.031**
		(0.011)	(0.010)	(0.010)	(0.010)		(0.012)	(0.013)	(0.013)
Fraction female-headed hhs		0.042**	0.051***	0.047**	0.048**		0.014	0.036	0.033
		(0.019)	(0.019)	(0.020)	(0.020)		(0.031)	(0.026)	(0.026)
Sex ratio (Female/Male)		0.019**	0.017**	0.020***	0.021***		0.033**	0.029**	0.032***
		(0.009)	(0.008)	(0.007)	(0.007)		(0.014)	(0.012)	(0.012)
Prop. Indian & white adults		-0.562	-0.528	-0.525	-0.526		-0.466	-0.429	-0.421
		(0.448)	(0.446)	(0.450)	(0.449)		(0.451)	(0.436)	(0.437)
Kilometers to road*10		0.001	0.000	0.000	0.000		0.000	0.001	0.001
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.002)	(0.002)
Kilometers to town*10		-0.003*	-0.001	-0.002	-0.002		-0.004*	-0.002	-0.002
		(0.002)	(0.002)	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)
Fraction men with h/s		0.051	0.054	0.072	0.065		0.065	0.064	0.080
		(0.064)	(0.066)	(0.065)	(0.066)		(0.079)	(0.078)	(0.079)
Fraction women with h/s		-0.067	-0.080	-0.072	-0.072		-0.147	-0.166	-0.165
		(0.095)	(0.097)	(0.092)	(0.092)		(0.124)	(0.113)	(0.111)
Change in water access				0.022***	0.022***				0.021***
				(0.006)	(0.006)				(0.008)
Change in toilet access				0.062	0.060				0.046
				(0.045)	(0.044)				(0.046)
Gradient*10					-0.007**				
					(0.003)				
District FE	N	N	Y	Y	Y	N	N	Y	Y
N	1,816	1,816	1,816	1,816	1,816	1,816	1,816	1,816	1,816
R^2	0.001	0.061	0.072	0.085	0.087				
Standard 95% C.I.	[-0.01;0.01]	[-0.01;0.01]	[-0.01;0.01]	[-0.01;0.01]		[-0.06;0.11]	[-0.04;0.19]	[-0.02;0.2]	[-0.01;0.2]
AR 95% C.I.								[0.05;0.3]	[0.05;0.3]

Outcome variable is change in employment rate of African females aged 15-59. Robust standard errors in parentheses, clustered at sub-district level. Electrified is 1 if community had the first Eskom project between 1996 and 2001, otherwise 0.Significant at p<0.01***, p<0.05** or p<0.1* level. Excluded instrument is mean community land gradient. All control variables (except electrification status indicator and change in access to water and toilet services) are measured in 1996. Change in access to water and toilets is measured between 1996 and 2001. Ten district fixed effects included in columns (3),(4),(5),(8) and (9). For the IV results, standard confidence intervals are provided as well as confidence intervals from the Anderson-Rubin test. The AR test is robust to weak instruments and is implemented to be robust to heteroscedasticity.

Table 6: Effects of electrification on male employment

Outcome is Δ_t male employment rate	OLS				Reduced form	IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Eskom Project	-0.017**	-0.015***	-0.009	-0.010*		-0.063	0.069	0.033	0.035
	(0.007)	(0.006)	(0.006)	(0.006)		(0.073)	(0.082)	(0.064)	(0.066)
Kilometers to grid *10		0.005*	0.000	0.001	0.001		0.008*	0.001	0.001
		(0.003)	(0.003)	(0.003)	(0.003)		(0.005)	(0.004)	(0.003)
Household density *10		0.001	0.002	0.002	0.002		0.000	0.001	0.002
		(0.002)	(0.002)	(0.002)	(0.002)		(0.003)	(0.002)	(0.002)
Poverty rate		0.062***	0.064***	0.063***	0.063***		0.059***	0.064***	0.062***
		(0.020)	(0.018)	(0.018)	(0.018)		(0.022)	(0.019)	(0.019)
Fraction female-headed hhs		0.217***	0.233***	0.227***	0.225***		0.187***	0.227***	0.220***
		(0.029)	(0.030)	(0.030)	(0.030)		(0.042)	(0.034)	(0.034)
Sex ratio (Female/Male)		0.018*	0.012	0.017	0.019*		0.034*	0.018	0.023
		(0.011)	(0.011)	(0.011)	(0.011)		(0.019)	(0.015)	(0.015)
Prop. Indian & white adults		0.032	0.075	0.080	0.090		0.139	0.121	0.130
		(0.549)	(0.547)	(0.549)	(0.545)		(0.555)	(0.543)	(0.544)
Kilometers to road*10		0.001	-0.001	-0.001	-0.001		0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)		(0.002)	(0.002)	(0.002)
Kilometers to town*10		-0.005**	0.001	0.000	0.000		-0.006**	0.001	0.000
		(0.002)	(0.002)	(0.002)	(0.002)		(0.003)	(0.003)	(0.002)
Fraction men with h/s		-0.058	-0.069	-0.043	-0.044		-0.042	-0.064	-0.039
		(0.091)	(0.088)	(0.088)	(0.088)		(0.108)	(0.092)	(0.091)
Fraction women with h/s		0.059	0.046	0.059	0.050		-0.031	0.005	0.015
		(0.114)	(0.111)	(0.106)	(0.106)		(0.151)	(0.124)	(0.123)
Change in water access				0.024***	0.024***				0.023***
				(0.009)	(0.009)				(0.009)
Change in toilet access				0.105**	0.102**				0.097**
				(0.046)	(0.046)				(0.049)
Gradient*10					-0.003				
					(0.005)				
District FE	N	N	Y	Y	Y	N	N	Y	Y
N	1,816	1,817	1,817	1,817	1,817	1,817	1,817	1,817	1,817
R^2	0.005	0.157	0.181	0.194	0.192				
Standard C.I.	[-0.03;0]	[-0.03;0]	[-0.02;0]	[-0.02;0]		[-0.21;0.08]	[-0.09;0.23]	[-0.09;0.16]	[-0.09;0.16]
AR confidence interval								[-0.05;0.25]	[-0.05;0.25]

Outcome variable is change in employment rate of African females aged 15-59. Robust standard errors in parentheses, clustered at sub-district level. Electrified is 1 if community had the first Eskom project between 1996 and 2001, otherwise 0.Significant at p<0.01***, p<0.05** or p<0.1* level. Excluded instrument is mean community land gradient. All control variables (except electrification status indicator and change in access to water and toilet services) are measured in 1996. Change in access to water and toilets is measured between 1996 and 2001. Ten district fixed effects included in columns (3),(4),(5),(8) and (9). For the IV results, standard confidence intervals are provided as well as confidence intervals from the Anderson-Rubin test. The AR test is robust to weak instruments and is implemented to be robust to heteroscedasticity.

Table 7: Placebo experiment and reduced form for employers of women: OLS results

	Placebo experiment: Reduced form for Δ_t in female employment in pre-period	Growth in major sources of female employment	
		Δ_t Schools	Δ_t Indian & White adults
	(1)	(2)	(3)
Gradient*10	-0.001	0.007	0.000
	(0.001)	(0.028)	(0.000)
Sample	Areas electrified before 1996 or not before 2002	Full sample	Full sample
N	373	1,816	1,816
R^2	0.106	0.057	0.039

Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level. In column (1), sample includes only areas that had projects prior to 1996 (already electrified). In column (2), outcome variable is the change in the number of schools in a community between 1996 and 2001. In column (3) it is the change in the fraction of Indian and white adults in the community between 1996 and 2001. Each regression controls for all covariates. All control variables (except electrification status and change in access to other services) are measured in 1996: distance to the grid, household density, community poverty rate, adult sex ratio (F/M), fraction of female-headed households, share of Indian/white adults, distance to nearest road, distance to nearest town, share of adult men and women with at least high school. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression. In column (3), the level of Indian/white adults in the community is excluded from regression. Schools data are taken from the 1995 and 2000 Schools Register of Needs. See data appendix for details.

Table 8: Employment, hours of work, wages and earnings for African workers in rural KZN 1995-2001: OLS and FE results from household survey data

	Females		Males		Females		Males	
	OLS	FE	OLS	FE	OLS	FE	OLS	FE
	A: Employment				B: Usual hours of work per week			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MD electrification rate	0.137***	0.118	0.108*	0.143	10.655***	21.194**	7.570**	23.070**
	(0.052)	(0.102)	(0.057)	(0.122)	(3.367)	(7.246)	(1.215)	(3.493)
Trend (1995-2001)	-0.013	0.050**	-0.048***	-0.075	7.443***	6.199*	5.177***	-0.609
	(0.011)	(0.024)	(0.014)	(0.074)	(0.827)	(2.823)	(0.331)	(0.586)
Constant	-0.079***	-0.117***	0.160***	0.360**	3.612**	-2.293	15.120***	19.638***
	(0.019)	(0.044)	(0.029)	(0.171)	(1.821)	(6.760)	(0.601)	(1.289)
N	152	152	152	152	151	151	151	151
R^2	0.08	0.63	0.09	0.75	0.44	0.56	0.25	0.55
	C: Log weekly wages				D: Log monthly earnings			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MD electrification rate	-0.046	-1.154**	0.120	0.228	0.003	-0.356	0.408***	1.168***
	(0.190)	(0.546)	(0.165)	(0.391)	(0.170)	(0.533)	(0.148)	(0.343)
Trend (1995-2001)	-0.085**	0.116	-0.037	0.075	-0.090**	-0.088	-0.053	-0.096*
	(0.038)	(0.097)	(0.036)	(0.058)	(0.036)	(0.169)	(0.032)	(0.052)
Constant	-0.097	0.186	0.149*	-0.017	-0.154*	0.037	0.108	-0.191*
	(0.100)	(0.238)	(0.089)	(0.116)	(0.085)	(0.427)	(0.072)	(0.102)
N	146	146	148	148	146	146	148	148
R^2	0.03	0.60	0.01	0.51	0.03	0.62	0.05	0.57

Robust standard errors, clustered at the magisterial district level. Significant at p<0.01***, p<0.05** or p<0.1* level. Data are from October Household Surveys 1995, 1997 and 1999 and the September Labor Force Survey 2001. The observation is a magisterial district-year. These results are from regressions of the MD-year level residuals on a trend (0,1,2,3) term and the fraction of households that have electric lighting in the MD-year. Sample for initial regression of outcome variable on age and education consists of African men and women ages 20-59 living in rural KZN. All FE regressions contain MD level fixed effects and MD-specific trends. The mean level of MD electrification is 0.3 and the average change over time is 0.15. Regressions for log wages and log earnings exclude MDs in which no-one reports positive earnings. Results for these outcomes are qualitatively similar if I use the level of wages/earnings instead of log wages/log earnings.

Table 9: Testing for spillovers in female employment by excluding adjacent control areas: OLS and IV results

Δ_t Female employment	Coefficient on Eskom Project indicator OLS (1)	IV (2)	N (3)
Full sample	-0.001 (0.005)	0.095* (0.055)	1,816
Sample excludes:			
Non-project areas within 1 kilometer of an project area	-0.004 (0.006)	0.076 (0.057)	1,205
Non-project areas within 5 kilometers of an project area	-0.003 (0.008)	0.069 (0.077)	840

Dependent variable is change in employment rate of African women aged 15-59. Each coefficient (standard error) is from a separate regression. Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level. Electrification is 1 if community had first Eskom project between 1996 and 2001, otherwise 0. Successive sample restrictions exclude non-electrified communities which fall partly/wholly inside an [X] kilometer radius of an area treated prior to 2001. All control variables (except electrification status and change in access to other services) are measured in 1996: distance to the grid, household density, community poverty rate, adult sex ratio (F/M), fraction of female-headed households, share of Indian/white adults, distance to nearest road, distance to nearest town, share of adult men and women with at least high school education. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression.

Table 10: Contribution of each poverty quintile to IV estimate of treatment effect

Quintiles of predicted poverty index	Fraction of Sample in Quintile (1)	Variance of Gradient by Quintile (λ_q) (2)	E(Δelec|z=1, q,x) - E(Δelec|z=0, (3)	IV weight (ω_q) (4)
Poorest quintile	0.159	0.203	0.015 (0.040)	0.036
Second poorest	0.180	0.209	0.045 (0.036)	0.129
Third poorest	0.213	0.212	0.099 (0.038)	0.340
Second richest	0.215	0.203	0.100 (0.038)	0.333
Richest quintile	0.234	0.200	0.045 (0.046)	0.162

Predicted poverty quintile is assigned as follows: for communities in the steepest half of the gradient distribution, I project the indicator of electricity project on to community poverty rate, the fraction of female-headed households and the female/male sex ratio. Predicted values are created for every community using these regression coefficients. Communities are assigned to quintiles, where quintile cut-points are defined by the regression sub-sample. λq is the estimated conditional variance of the gradient dummy (1=flat, 0=steep) within each quintile (q): predicted E(P[Z|x,q][1-P(Z|x,q)|q]). Predicted Δelec|q=predicted E(E((elec|z=1,x,q) - predicted E(elec|z=0,x,q)|q)) is the estimated difference in probability of getting an electricity project across top and bottom halves of gradient distribution within each quintile, controlling for covariates. Each estimated coefficient in column (3) is on the interaction of the gradient dummy (1=flat, 0=steep) with each predicted quintile dummy. ωq is the weight that each quintile contributes to the IV estimate. As described in Kling (2001), it is computed across the columns (1)-(3): for each q, ωq= ([(1)q*(2)q*(3)q])/(Sum of q [(1)q*(2)q*(3)q])

Table 11: Household energy use by poverty quintile: At baseline and over time, 1996 to 2001

	Fuel Use in Home Production: Fraction using [X] in 1996			Δ_t in Fuel Use for Home Production: Within-quintile difference by gradient			Δ_t in employment by gradient	
Quintile of predicted poverty index	Electric Lighting	Electric Cooking	Wood Cooking	Electric Lighting	Electric Cooking	Wood Cooking	Females	Males
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Poorest quintile	0.019	0.011	0.896	0.001	0.002	0.000	0.001	-0.005
	(0.079)	(0.048)	(0.148)	(0.018)	(0.006)	(0.019)	(0.005)	(0.007)
Second	0.043	0.023	0.854	0.002	0.004	-0.012	0.00806*	0.001
	(0.139)	(0.084)	(0.191)	(0.017)	(0.008)	(0.014)	(0.005)	(0.007)
Third	0.072	0.034	0.807	0.036	0.0152*	-0.0193*	0.00792*	0.002
	(0.174)	(0.096)	(0.215)	(0.023)	(0.009)	(0.011)	(0.005)	(0.006)
Fourth	0.123	0.055	0.724	0.027	0.0214**	-0.0337**	0.0101*	0.002
	(0.230)	(0.119)	(0.262)	(0.021)	(0.010)	(0.015)	(0.006)	(0.008)
Richest quintile	0.176	0.090	0.636	0.042	0.0258**	-0.025	-0.002	-0.012
	(0.274)	(0.158)	(0.300)	(0.030)	(0.011)	(0.018)	(0.008)	(0.010)

Significant at p<0.01***, p<0.05** or p<0.1* level. Columns (1)-(3) present the quintile means of outcome variables in 1996, columns (4)-(8) present coefficients from regression of interactions of gradient dummy and predicted poverty quintile.

Table 12: Age-specific effects of electrification on female employment

Δ_t female employment	OLS	IV
Ages 15-19	0.000	0.000
	(0.000)	(0.005)
Ages 20-24	0.000	0.009
	(0.001)	(0.013)
Ages 25-29	-0.001	0.015
	(0.001)	(0.012)
Ages 30-34	-0.001	0.030*
	(0.001)	(0.012)
Ages 35-39	0.000	0.017
	(0.001)	(0.013)
Ages 40-44	0.006***	0.001
	(0.002)	(0.001)
Ages 45-49	0.007	0.014*
	(0.012)	(0.008)
Ages 50-54	-0.001	-0.001
	(0.001)	(0.007)
Ages 55-59	0.001	0.004
	(0.001)	(0.006)

Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level. Dependent variable is change in fraction of employed African females in age group [X] over all females, where [X] is one of nine 5-year age groups. Each cell presents a regression coefficient on the "Electrified community" indicator, from separate regressions. All controls (except "Electrified community" and change in access to other services) are measured in 1996: distance to the grid, household density, community poverty rate, adult sex ratio (F/M), fraction of female-headed households, share of Indian/white adults, distance to nearest road, distance to nearest town, share of adult men and women with at least high school. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression. Excluded instrument is average land gradient. N=1,816 in each regression.

Table 13: Effects of electrification on population growth, employment of incumbants and skill composition of labor force

	Δ_t Log Population		Δ_t Fraction Females with Matric		Δ_t Fraction Males with Matric		Δ_t Female Employment: Excluding In-Migrants		Δ_t Male Employment: Excluding In-Migrants	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Eskom Project	0.171***	3.897***	-0.001	0.095*	-0.010*	0.035	0.000	0.116*	-0.008	0.086
	(0.045)	(1.427)	(0.005)	(0.055)	(0.006)	(0.066)	(0.005)	(0.069)	(0.005)	(0.069)
N	1,816	1,816	1,816	1,816	1,817	1,817	1,816	1,816	1,816	1,816
R^2	0.066		0.09				0.036		0.148	

Dependent variable in columns (1)-(2) is change in log African population. In columns (3)-(6), it is the change in fraction of women or men that have a completed high school education. In columns (7)-(10) it is the change in the employment rate of African females or males, where the numerator has been adjusted downwards to exclude those African adults who report they have moved to the area in the five years before the relevant Census year. Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level. All control variables (except electrification status and change in access to other services) are measured in 1996: distance to the grid, household density, community poverty rate, adult sex ratio (F/M), fraction of female-headed households, share of Indian/white adults, distance to nearest road, distance to nearest town, share of adult men and women with at least high school. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression. The mean number of adult female (male) in-migrants in 1996 is 15.02 (10.07), in 2001 is 11.56 (8.60). Regressions in columns (3)-(6) do not include controls for baseline fraction of women or men with completed high school.

Figure 1A: Change in fuel use over time for areas with Eskom projects 1996-2001

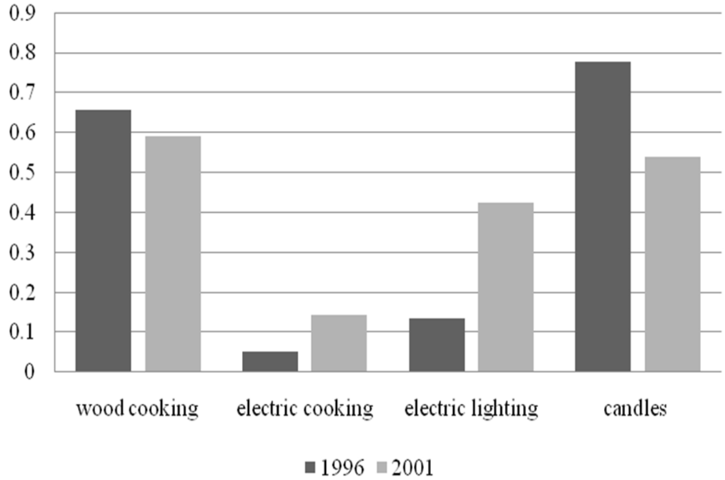
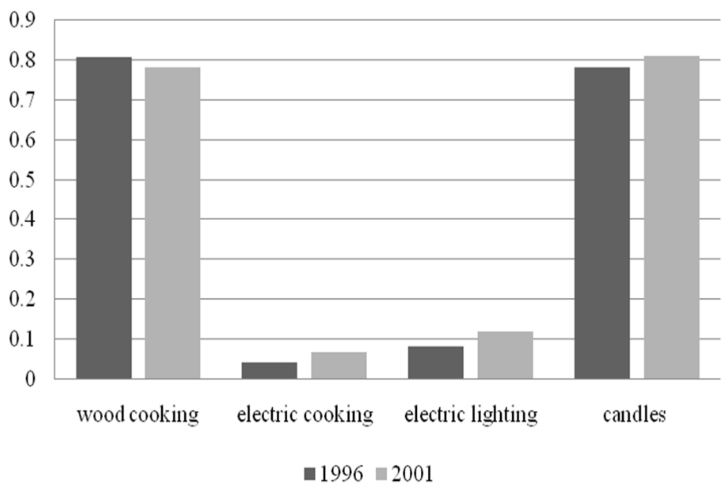
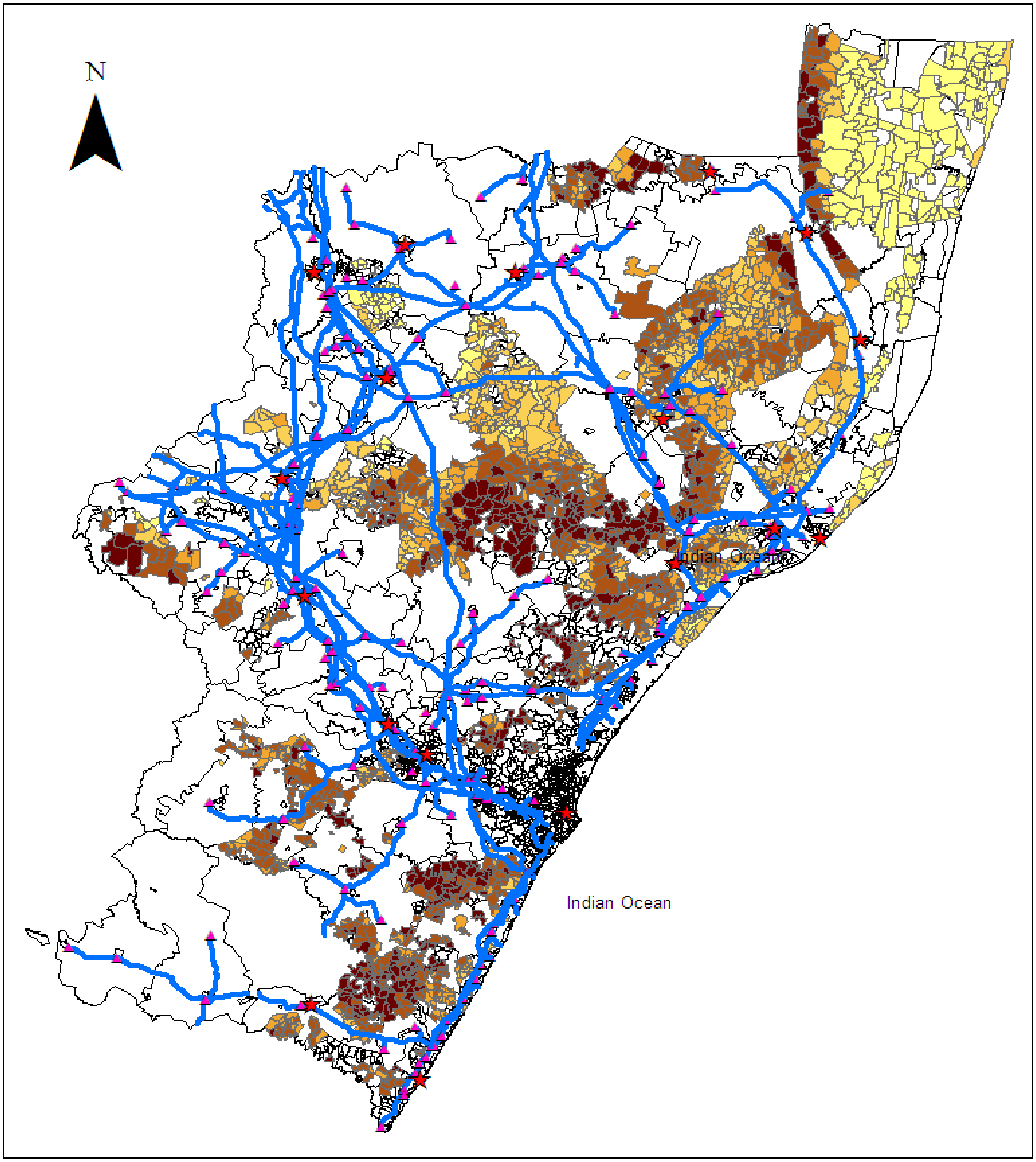


Figure 1B: Change in fuel use over time for areas without Eskom projects 1996-2001



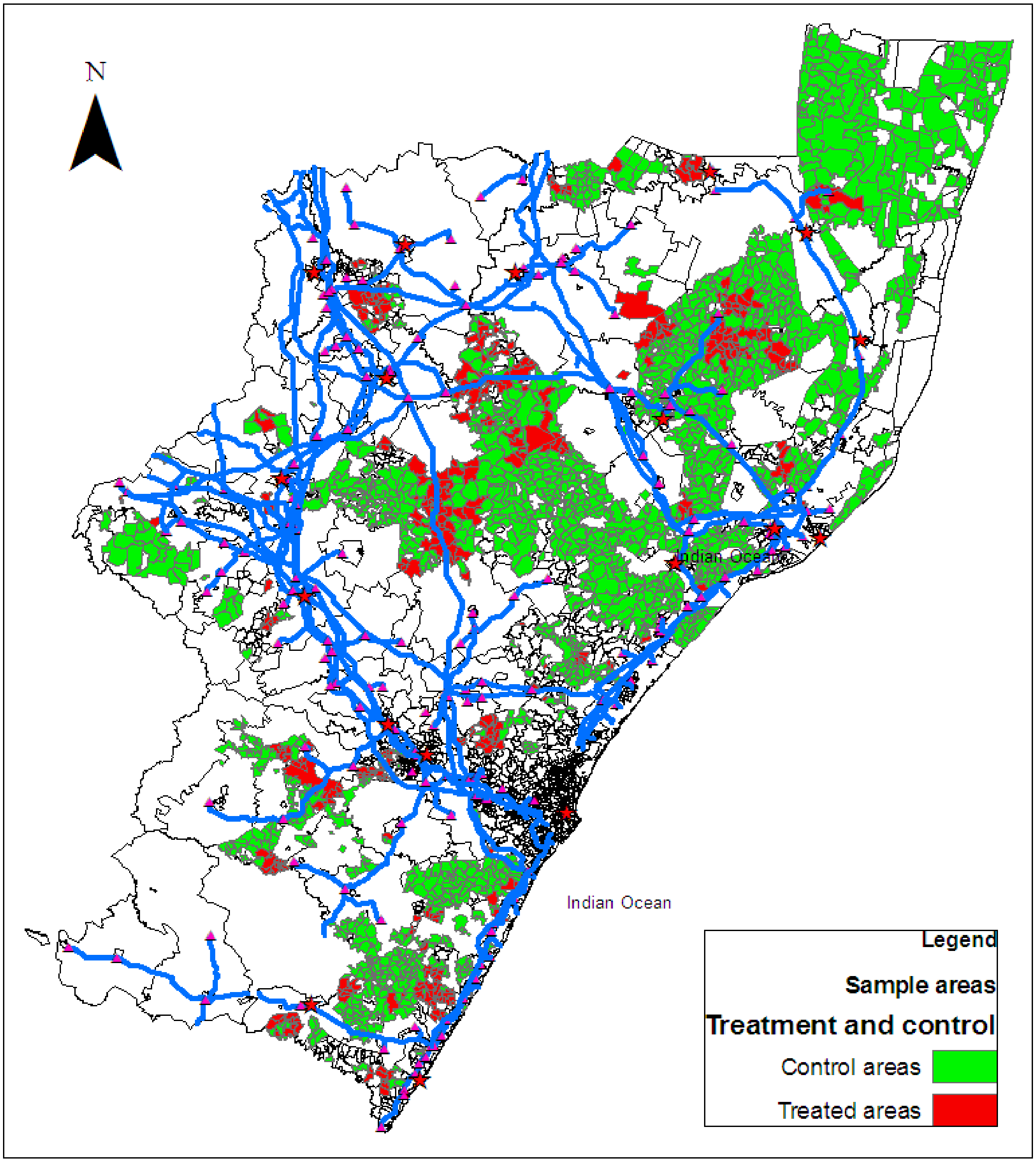
Each graph shows the average fraction of households reporting main source of fuel for cooking and lighting as wood, electricity or candles. The unit of observation is the census community. Sample includes all African households in communities that are part of the main analysis sample, in tribal KZN.

Figure 2: Spatial distribution of gradient in sample areas: KwaZulu-Natal, South Africa



All shaded areas are included in sample. Steeper areas are shaded dark, flatter areas are shaded light (see electronic version for color). Thick lines depict electricity grid lines in 1996, triangles are electricity substations in 1996 and stars represent towns. N=1,816.

Figure 3: Electrified and Non-electrified areas: KwaZulu-Natal, South Africa

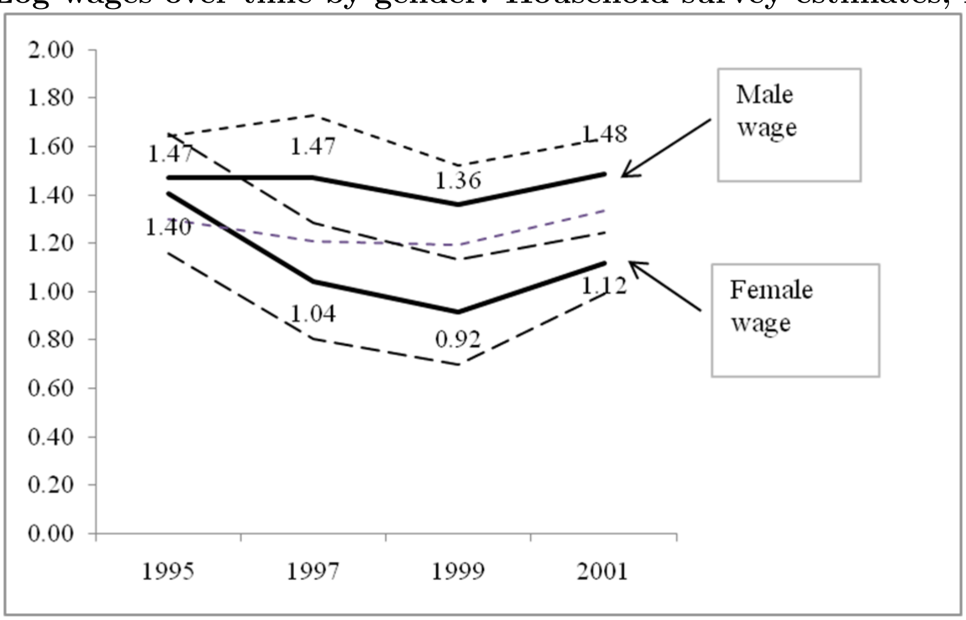


Shaded areas are in the sample: dark-shaded areas receive an Eskom project between 1996 and 2001, lighter shaded areas are electrified after 2001 or not at all (see electronic version for color). Thick lines represent electricity grid lines in 1996, triangles are electricity substations in 1996 and stars represent towns. N=1,816, N_T = 365, N_C = 1,451.

Figure 4: Employment rates over time by gender: Household survey estimates, rural KZN

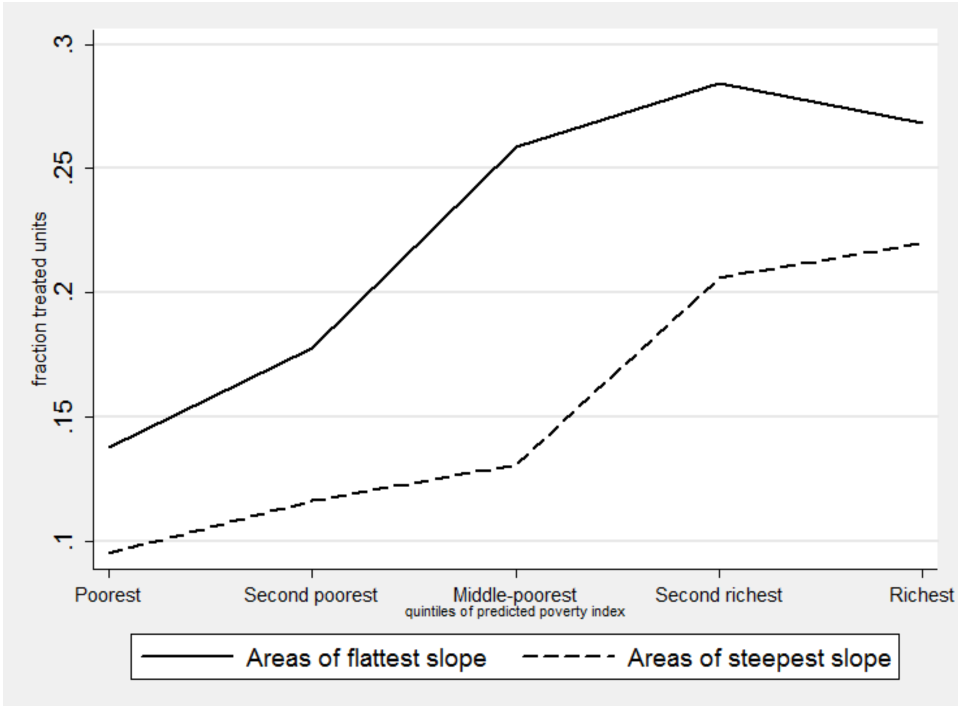
Data are from October Household Surveys 1995, 1997, 1999 and the September Labor Force Survey 2001. Sample includes African men and women aged 15 to 59 living in rural KZN (not just tribal areas). Dashed lines are 95% confidence intervals. The unit of observation is the individual.

Figure 5: Log wages over time by gender: Household survey estimates, rural KZN



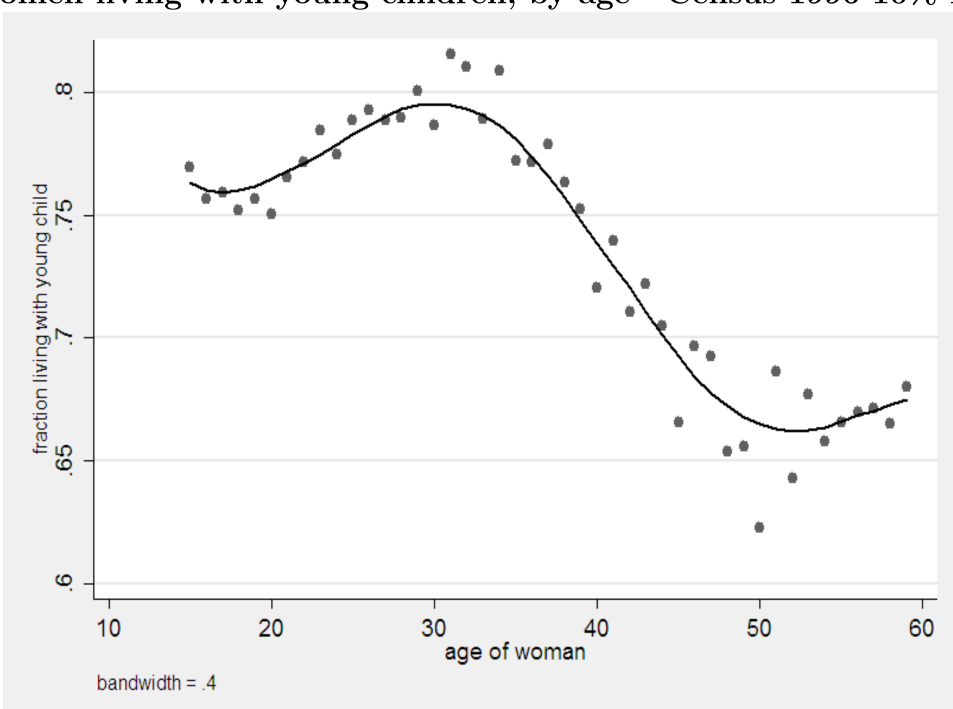
Data are from October Household Surveys 1995, 1997, 1999 and the September Labor Force Survey 2001. Sample includes African men and women aged 15 to 59 living in rural KZN (not just tribal areas) who are employed. Log wage is constructed by combining reported monthly earnings with usual weekly hours of work. Dashed lines are 95% confidence intervals. The unit of observation is the individual.

Figure 6: Effect of gradient on electrification probability, by predicted poverty quintile



Lines show fraction of each predicted poverty quintile that is electrified, by top (steep) and bottom (flat) halves of the gradient distribution. See notes for Table 10 for a description of how poverty index is created. The gap between the two lines indicates at which part of the poverty index the gradient manipulates the probability of electrification the most.

Figure 7: Women living with young children, by age - Census 1996 10% micro sample



Lowess-smoothed graph of the fraction of women of each age living with at least one child under the age of 9. Data are from the 1996 South African Census 10% micro data and include African women aged 15-59 living in rural KwaZulu-Natal. N=116,381 collapsed to 45 age-specific data points.