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

Taryn Dinkelman*
University of Michigan

Job Market Paper

January 11, 2008

Abstract

This paper investigates the employment effects of a mass roll-out of household electrification in rural South Africa. For the 470,000 households that were connected in this area, the new infrastructure represented an improvement in home production technology. I collected administrative and geographic data on this expansion and matched it to two waves of aggregate Census data to test for effects on employment and on fuel use at home. I exploit variation in electricity project placement and timing to estimate district fixed-effects models, and instrument for project placement using land gradient that directly affects the cost of grid expansion. My findings show that cooking with wood falls sharply in treated areas over a five-year period, and lighting and cooking with electricity increase substantially. IV employment results indicate asymmetric responses by gender: female employment rates increase by 13.5 percentage points in treated areas, while there are no significant male effects. Middle-poor communities respond most to the new option to use electricity, and employment effects are large for women in their thirties and forties who are less constrained by child-care responsibilities. This new evidence on how home production infrastructure can affect the extensive margin of work for women contributes to a growing literature on the effects of public infrastructure in developing countries. The results suggest that studies that do not examine employment outcomes may miss important economic effects of these infrastructure investments.

*PhD candidate, Department of Economics. Email tdinkelm@umich.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, HwaJung Choi, John DiNardo, Ann Ferris, Ben Keyes, David Lam, Jim Levinsohn, Claudia Martinez-A, Justin McCrary, Dori Posel, Samara Potter-Gunter, Jeff Smith, Gary Solon, James Tybout, Rebecca Thornton and Martin Wittenberg substantially improved the paper. I also thank seminar participants at the Universities of Michigan, Michigan State Western Ontario and KwaZulu-Natal for their valuable input. Data collection was funded through a National Science Foundation Doctoral Dissertation Grant and an Eva Mueller New Directions in Demography research award. All errors are my own.

1 Introduction

Electricity is pervasive in all industrialized countries and largely absent in developing ones. 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, access to electricity and other modern energy sources is set to expand in many of these poor countries (EnergyNet Limited (2004), Deutsche Presse-Agentur (2006), Sustainable Development Network (2007)). Some of this expansion targets industrial needs. For example, World Bank commitments to energy infrastructure in sub-Saharan Africa rose from $447 million in 2001 to $790 million in 2007. Other "Bottom of the Pyramid" initiatives which are focused on households aim to make cheap, modern energy available to 250 million Africans by 2030.[1] Lack of data and challenging identification issues have constrained what we know about the effects this infrastructure could have. It is an open research question as to how economic growth, health, human capital accumulation and labor market outcomes will be impacted by these investments.

In this paper, I measure the impact of access to modern energy sources on an outcome of considerable interest: the ability of the poor to use their labor resources for market production. To do this, I exploit data from South Africa's post-apartheid electrification roll-out to measure the effects of household electrification on fuel-use, and on rural employment in a part of KwaZulu-Natal (KZN) province. I ask whether women as primary home producers respond more to this new infrastructure and if so, which women experience the largest impact.

Time-use data, where it exists, indicates that fuel-wood collection and food preparation with traditional fuels consumes significant time resources. In South Africa, over 80% of households collected wood for household needs in the mid-1990s; 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).[2] These figures are typical of other developing countries (Charmes, 2005) and provide good reasons to expect that household electrification could change the nature of work for poor women (Saghir, 2005; United Nations, 2005). One of the contributions of this

[1] See the World Bank's Lighting Africa website (www.lightingafrica.org)

[2] Food preparation in rural areas consumes on average 3 hours per day. Data from South African's 1997 October Household Survey and the South African Labor Force Survey (September 2001) suggest that more people are involved in and more time is spent on wood than on water collection.

paper is to provide some of the first evidence in support of these claims. Using new administrative and spatial data matched with two waves of Census data covering rural areas of KZN, this paper shows that female employment does indeed rise in response to new access to electricity in the home.

The roll-out of electricity-grid infrastructure in South Africa provides a good opportunity to evaluate the effects of household electrification on employment. In 1993, one year before the end of apartheid, over two-thirds of South African households were without electricity. Following the new government's commitment to universal electrification, 2 million (out of 8.4 million) 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 was its focus on low-capacity household connections rather than industrial connections (Gaunt, 2003).

A large literature on the relationship between infrastructure and economic growth acknowledges that infrastructure spending may be targeted at growth centers, or towards areas that are lagging behind, but politically important. In addition, teasing out the effects of infrastructure on economic variables is difficult without adequate controls for general trends in economic conditions.[3] These difficulties characterize South Africa's roll-out: it was socio-politically motivated and occurred during a time of economic restructuring. Non-random placement of electricity projects is likely in this context, and the identification problem is unlikely to be solved by comparing outcomes in treated and non-treated areas.

To address the indeterminate bias arising from endogenous placement and confounding trends, I collect and use information on the technological constraints on roll-out to generate exogenous variation in electricity project allocation. The data are drawn from rural KZN, a province containing one-fifth of the population of South Africa.[4] I compare changes in employment rates in areas that have had electricity projects (treated areas) to those that have not (control areas). District fixed-effects and baseline controls absorb some differences in local labor market conditions over time. To deal with selection and confounders, I instrument for project placement between 1996 and 2001

[3]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) analyze South African data. The World Bank (1994) reviews some of the infrastructure evaluation literature and Jimenez (1995) discusses difficulties in establishing causality in this literature.

[4]This sample is chosen for two reasons: electricity projects are likely to have larger impacts in rural areas where reliance on wood is high in the baseline period, and administrative data is only available in KZN province.

using community land gradient. Gradient directly affects the average cost per connection – a primary factor in prioritizing areas for electrification – and is unlikely to directly affect employment outcomes, conditional on controls and district fixed-effects. This identification strategy is similar to Duflo and Pande (2007) who use functions of gradient to instrument for dam allocation within Indian districts.[5]

In areas treated with electricity, I estimate large changes in energy use for home production. The share of households using electric lighting rises by 23 percentage points in treated areas and the share of households cooking with wood falls by 4.2 percentage points. Even greater shocks to home production technology occur in marginal communities affected by gradient: instrumental variables estimates are three to seven times larger than OLS estimates.

OLS results for employment strongly suggest that projects are targeted towards areas which are doing more poorly over time: employment rates are 0.1 percentage points higher for women in treated areas and 1.1 percentage points lower for men.[6] In contrast, instrumental variable results indicate that female employment rises by a significant 13.5 percentage points. This point estimate has a 95 percent confidence interval from 5 to 40 percentage points, indicating a non-trivial positive impact on women's employment. Effects for men are not significantly different from zero. The female results are notable, since over the same period national unemployment rates are rising. A placebo test for areas treated prior to 1996 provides no evidence that gradient is directly associated with employment prospects over time.

Several pieces of evidence together suggest that these results are unlikely to be driven by a net increase in labor demand. The capacity of electricity supplied to households was too small for even mid-size firms (South African Department of Minerals and Energy, 2006); there is no correlation between land gradient and growth in major sources of female employment (schools and domestic worker employers); and there is no evidence of cross-community employment spillovers, which would be one outcome of increasing labor demand. While my research design cannot rule out that electricity lowered the costs of opening new businesses, thereby creating new jobs, the weight of evidence points to household electrification operating as a labor-saving technology shock to home

[5]In that paper, no employment data are available to measure the impact of dams on labor market outcomes.

[6]This is partly due to industrial restructuring in the 1990s, particularly in male-dominated commercial agriculture and mining sectors situated outside of rural areas.

production in rural areas.

The channels through which household electrification affects female employment are related to other household constraints.[7] I show that treatment probability is significantly higher in flatter, middle-poor communities. These middle-poor communities are just the ones more likely to respond to the new option to use electricity. They contain households that have not yet invested in alternative modern technologies like gas or kerosene (as richer areas have), and compared to the poorest communities, more households switch towards the new service when it arrives. Relatedly, I also find that women most affected by the expansion are those who have more flexibility to respond to the new home-production technology. These are women in their thirties and forties, who are less likely to live with young children requiring full-time care.

This paper contributes to our understanding of 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.[8] Studies which do not measure employment effects could be missing important economic impacts, particularly when the infrastructure has a home-production bias (for example, water and sanitation services).[9] Second, the results highlight short-run heterogeneity in the effects of infrastructure across communities and groups of women. This information helps policy-makers to predict the distributional effects of infrastructure.

Third, the fact that inference from a comparison of treated and non-treated areas is misleading highlights the importance of designing an identification strategy robust to omitted variables' bias

[7]It is possible, that employment effects are driven through improved health through access to cleaner technology. Indoor air pollution emanating from traditional fuels used in cooking has been associated with respiratory diseases and infant morbidity and mortality (see WHO at www.who.int/indoorair/en/ and Rosenzweig et al (2006)). Health effects are less likely to drive the employment response 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 settings where much of the indoor air pollution research is conducted; 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 South African Time Use Survey and Rosenzweig et al (2006).)

[8]See Cutler and Miller (2005) for the effects of clean water technology in the USA; 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 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 of 38 percent.

[9]Field (2007) highlights a similar point in her paper in which she studies how new property rights in urban Peru release time spent on monitoring assets for more market work.

and selection issues. These are critical issues in all infrastructure studies. The strategy used in this paper, which relies on administrative and spatial data to model project allocation (see data section below), is likely to be feasible in other networked-infrastructure settings.

Finally, the results of this paper add to a vast literature on female labor force participation and economic development.[10] Recent work by Greenwood et al (2005) argues 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.[11] The response of female employment to household electrification in South Africa represents complementary and current evidence that access to infrastructure may have similarly large effects on the extensive margin of female employment in a developing country setting.

The paper begins with a brief discussion of the effects of a positive shock to home production technology on labor supply and describes where responses to electrification might be largest. Sections 3 and 4 describe the data and context of South Africa's electrification. Section 5 outlines the empirical strategy and sections 6 and 7 present main results and robustness checks. Section 8 investigates some of the channels through which electrification affects employment, section 9 discusses the implications of the employment results and section 10 concludes.

2 Effects of technology shocks in home production

In a Becker-type home-production model (Becker, 1965), households consume commodities that are produced with a combination of market goods and home time, and market goods can only be bought by supplying time to the labor market. Improvements in the technology of home production can affect the intensive and extensive margins of labor supply in ambiguous ways in this model (Gronau, 1986). To work through the intuition for this, consider a household that consumes two commodities: meals and clothing. Meal production is time-intensive and clothing production is market goods-intensive. Assume that the household has non-homothetic preferences (non-linear Engel curves), so that as income increases, the demand for clothing increases relatively more than the

[10] 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)).

[11] Although see Bailey and Collins (2006) for some counter-evidence using variation in electrification use rates over time.

demand for meals.[12]

The arrival of infrastructure for domestic electricity may be characterized as a positive shock to time productivity.[13] Labor-saving electrification increases the effective amount of labor available for producing commodities: it reduces the need to fetch wood, speeds up cooking time and allows households to shift activities from daytime into night-time. Since meals production is time-intensive, this shock increases meals productivity more than clothing productivity.

After electricity arrives, substitution and endowment effects operate in different directions on the household's decision to supply time to the market and home production. A higher marginal product of time in meals production encourages the household to move time out of clothing production and into meals. However, the new infrastructure increases the effective amount of time available to the household. The demand for all normal goods rises in response to this increase in household endowment, and the demand for income-elastic clothing rises relatively more. Since clothing is market-goods intensive, the household may decide to work more in the market in order to consume more of this commodity. Within this framework, it is possible for the household to produce more meals with less total time in home production after the technology shock. However, the net effect of the substitution and endowment effects on labor supply are ambiguous.

Differences in responses to this technology shock will be linked to heterogenous preferences for time- and market-intensive commodities (which we can't measure) and differences in initial home production technology (some of which we can measure) which determine how large an effect the given technology shock can have. In a more complex model, households engaging in fewer home-time intensive activities (e.g. child-care) will be more able to respond to a technology shock by shifting labor into the market. Regardless of whether labor supply to the market increases or decreases, we would expect the technology shock to have a larger effect on female time, as women are typically the primary home producers.

One question worth asking is: How plausible is it that household electrification affects the

[12]It seems reasonable to assume that as people get wealthier, expenditure on non-food items increases relatively more than expenditures on food items. Houthakker (1957) shows cross-country evidence that the demand for food is income inelastic while clothing is income elastic. Deaton and Subramanian (1996) estimate income-inelastic demand for calories in India.

[13]This is similar to Michael (1982) who models the impact that human capital has on non-market productivity. A related interpretation is that electrification reduces necessary home activities (collecting wood) that are complementary to time in the home, and so releases this time for other activities.

amount of time in home production in rural South Africa? Since fuel-wood collection is one of the most time-consuming home production activities undertaken by women (as evidenced in time-use data), large effects are possible (Budlender et al, 2001). Potential beneficiaries also expect to experience large effects: in a 1990 volume of the South African journal Reality, researchers note that Africans without electricity expected its arrival to lengthen the day for productive activities and ease household work. Households that get connected to the grid do exhibit large increases in ownership of appliances that improve the efficiency of home production (kettles, fridges and electric lighting), [14] and there is some evidence that access to electricity directly reduces the physical burden of home work.[15]

Of course, new access to electricity may directly affect market production and the demand for labor. Wage data over time and space could provide evidence consistent with a net labor supply (if wages fall) or demand (if wages rise) effect of electrification. Unfortunately, my data contain no wage information, nor do any surveys capturing wages contain enough spatial information to be useful. Instead, I argue that both the context of the roll-out and evidence from the data are consistent with household electrification having a net labor supply impact. Since new access was limited to a low level of service, not even medium-size businesses could be supported (South African Department of Minerals and Energy, 2004). In addition, if electrification predominantly affected labor demand, we should expect (i) no gender or age asymmetry in employment responses (ii) different employment effects when comparing treated areas to non-adjacent control areas that are at less risk of experiencing labor demand spillovers and (iii) growth in the number of major employers in response to treatment. None of these effects are evident in my data.

[14] Author's own calculations from the KwaZulu-Natal Income Dynamics Study, a small panel study of households from 1993 and 1998. These data do not contain enough geographic information to be used for the main analysis.

[15] Wittenberg (2007) shows that African women in households with electricity have significantly higher body mass index than women in households without, and that this positive relationship remains even after controlling for household expenditures and the presence of a television.

3 Electrification roll-out in South Africa

3.1 Details of the program

Under apartheid, many African households were denied access to basic services, particularly in homeland areas.[16] By 1990, most economic units and white settlements were electrified and the political concerns of the 1980s had led to extensive electrification 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 eliminating this backlog a development priority.[17]

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).[18] 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 in to the industry was not necessary to provide full access to previously disadvantaged communities (Gaunt, 2006). As a result, Eskom met the 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 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 Amps or could upgrade to a 20 Amp supply for a fee of about ZAR40 (USD6.00).[19] The majority of Eskom's 3 million rural customers

[16] Homelands were pockets of land designated for African settlement which functioned as labor reserves for the white economy under apartheid. In 1994, all homelands were legally reintegrated into South Africa (Christopher, 2001).

[17] Details in this section were collected from a combination of written sources (Gaunt, 2003; Eskom, 1996) and personal 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.

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

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

upgraded to the 20 Amp supply (Gaunt, 2003). The default supply was sufficient for television, radio, two lights and several small kitchen appliances. The upgraded supply could simultaneously power more small appliances including a fridge and a small water heater (South African Department of Minerals and Energy, 2004).

This subsidized roll-out represented a change in the option to use electricity. Households still had to pay for the use of this service by purchasing electricity credits loaded on to pre-paid cards from local stores. This suggests that the behavior of individuals in poor households may be less responsive to this new infrastructure. In 1999, the cost of household electricity was equivalent to \$0.039 per kilowatt hour (kWh).[20] Using estimates of load demand from Eskom reports, most rural households would have used between 35 and $60kWh$ per month, translating into between \$1.37 and \$2.34 per month (Gaunt, 2003) or 1.8 percent of median monthly household income in rural KZN in 1995.[21] Despite this user cost of electricity, industry experts agreed that "Electric lighting was synonymous with the roll-out"[22], and that the NEP did reach poor households.

3.2 Selection

Almost by definition, networked infrastructure (whether fixed-line telephony, roads, rail, electricity, piped water or waterborne sanitation) requires that consumers be connected in some order, and identical consumers can seldom be connected simultaneously. In the context of the NEP, local political pressures and connections costs each played a role in project prioritization.

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.[23] As there is no way

force behind the current, and voltage*amperage gives the wattage, or total measure of power.

[20]This was ZAR0.246. 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)

[21]Median monthly household income for Africans living in rural KZN comes from the Income and Expenditure Survey 1995.

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

[23]See Christopher (2001), and Piper (1999) for a discussion of the political landscape in KZN. Johnston (1997) show how the role of tribal authorities in rural local government structures contributed to postponements of local elections in the area.

to measure these political factors in my data, I treat them as omitted variables.

Annual reports and interviews with planning engineers also point to a 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.[24] 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 to cross and the softer the soil, the cheaper it is to lay power lines and erect transmission poles (Eskom, 1996; West et al, 1997).

In my data set, I have assembled measures of each of these cost factors. Distance from the grid and household density are important controls, as both 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 which I discuss further in section 5.

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 Appendix 1. For some parts of the analysis, I also use the 10% micro data for 1996 and 2001. These data are reported at the household level, but do not contain enough geography to identify areas treated with electricity projects which is necessary for

Khan et al (2006) claim that "There is no doubt that any continued contestation and outbreaks of violence will continue to hamper service delivery such as houses, electricity, water and roads". They also describe how variation in levels of education and ability among tribal authorities affected their ability to lobby government for effective service delivery.

[24]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."

the main analysis.

Community aggregate data (where a community is a Census sub-place) provide full population totals for each year for different combinations of variables. A community is small, roughly equivalent to a US Census tract. The median number of households is 146 in 1996, and 187 in 2001. Ninety-five percent of communities have no more than 750 households. Key variables 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. Results are not weighted, as all variables are derived from the full population Census.[25]

Employment questions across Census waves are similar. The 2001 employment definition is somewhat more expansive 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 are unlikely to be problematic as long as reported part-time work does not differentially contribute to new employment in areas treated by virtue of having flatter gradient.

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

To assign treatment 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 treatment status with project data is that we can identify when a community gets new access to infrastructure rather than rely on time variation in electricity use patterns which is likely to be strongly correlated with wealth.

An area is treated ($T = 1$) if the community had its first electricity project between 1996 and 2001 (inclusive) and untreated ($T = 0$) 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

[25] The aggregate data are adjusted by Statistics South Africa for under-count (Personal Communication with Piet Alberts, Senior Statistician in the Census department of Statistics South Africa, May 2007).

conduct a placebo test in support of the exclusion restriction. Two other treatment measures are constructed for sensitivity tests: time since treatment (T_{time}) and treatment exposure ($T_{connect}$) that calculates the cumulative proportion of households connected between 1996 and 2001.

Finally, I construct measures of land gradient for each community using digital elevation data. For sensitivity analysis, I also construct an additional instrument containing information on whether a community is in the likely area of expansion of the electricity grid. Electricity is delivered along distribution lines connecting substations to each other. Substations (shown as triangles in Figures 1 and 2) are necessary pieces of infrastructure for converting power supplies to appropriate voltages for household use. I assume that it is easier to expand the network by setting up transmission lines between substations than it is to build new substations. This is reasonable, since new substations take between three and five years to build and my study period is only five years long. Also, in the mid-1990s, there was enough capacity in the system to build lines out from existing substations.[26] Using the spatial data, I connect each substation to all other substations in the region and create a variable ($Transect$) indicating whether any part of a community lies on any one of these connector lines.

Rural former homeland areas in KwaZulu-Natal (KZN) are selected to be part of the sample. 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 rural electrification effects 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. Rural areas are also simpler to analyze than urban areas. Eskom is the sole distributor of power for rural areas, whereas in urban areas, distribution rights are shared with local municipalities. 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 the most likely source of confounding in rural settings, and I control for this in the empirical work.

[26] NetGroup (2006) describe how this excess capacity had been exhausted by 2006.

Boundaries of the sample of 1,992 rural communities and the spatial distribution of land gradients are illustrated in Figure 1. The fragmentation that characterized the former KwaZulu is evident; the apartheid government forcefully resettled Africans to areas deemed inhospitable for white settlement, wherever those were (Christopher, 2001). Gradient varies widely across the region and is very steep in some areas: the middle of the province is colloquially named the "Valley of a Thousand Hills". The average gradient of the area falls into the Food and Agriculture Organisation category of "strongly sloping" (FAO, 1998).

Several important features of project placement are evident in Figure 2. Treated areas are not all positioned close to 1996 grid infrastructure, and many areas adjacent to the grid are control areas. Being close to the original grid is neither necessary nor sufficient for subsequent electrification although we will see that it does raise the probability of treatment. Additionally, treated areas are not all clustered near towns– proximity to a town is not necessary for treatment. There is also a good distribution of treated areas across the entire province. This allows for the inclusion of district fixed-effects that can absorb differences in growth rates across local labor markets.

5 Empirical strategy

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.[27] T_{jdt} is an indicator variable for whether a community has had an electrification project by time period t. If treatment T_{jdt} was randomly assigned across communities, we could estimate the average treatment effect α_2 by OLS:

$$y_{jdt} = \alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \lambda_d t + \epsilon_{jdt} \quad (1)$$

where μ_j is a community fixed-effect, δ_j is a community trend, λ_d is a district trend and ϵ_{jdt} is remaining idiosyncratic error. With two years of data, it is not possible to control for the community trend term, δ_j.[28]

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

[28] This linear relationship between employment and treatment is adopted for simplicity. Appendix 2 discusses results using

Electricity projects are unlikely to be randomly assigned across space or time, and positive or negative selection on community- and district-level unobservable characteristics is possible. To eliminate the community fixed-effect, re-write equation (1) in first differences so that ΔT_{jdt} is 1 if the community has an electricity project in between t and $t+1$, otherwise 0. Without more years of data, the community trend δ_j is still problematic.

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

Even in first-difference form, there are reasons to expect that OLS will not provide the correct answer to the question: what is the causal effect of electrification on employment?

Positive selection on community trend (δ_j) could occur and $\hat{\alpha}_{2,OLS}$ would be biased upwards if electricity projects are allocated to communities growing faster for unobservable reasons. This is a typical concern when estimating effects of infrastructure.[29] However, reports suggest that Eskom was not cherry-picking wealthier areas and that the expansion was in fact an unprofitable undertaking for the company (Gaunt, 2003; South African Department of Minerals and Energy, 2001).

Negative selection on the community trend is also possible if projects are targeted to more disadvantaged areas.[30] Since electrification was driven by a socio-political compact between Eskom and the newly-elected government, political concerns for disadvantaged communities could have directed some of the placement. In addition, during this period, South Africa's economy is restructuring and jobs are being lost in certain sectors and areas. If the areas that are losing jobs are also the ones targeted for projects, $\hat{\alpha}_{2,OLS}$ would be biased downwards.

Measurement error in the treatment variable ΔT_{jdt} presents a third practical challenge for estimating (2). Since Eskom region boundaries do not line up with Census boundaries, treatment is assigned 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

an alternative specification for the outcome variable that allows for all variables to impact employment in a non-linear fashion. This specification produces qualitatively similar results to the linear model.

[29] de V. Cavalcanti (2007) develop a growth model in which higher female labor force participation actually leads to increases in government expenditure on services that improve the efficiency of home production. They argue that female demand for services drives these investments.

[30] Banerjee and Somanathan (2007) report that gains in access to public goods in India appear to be allocated to the more politically mobilized disadvantaged groups.

are assigned full treatment status when only a small fraction of households in the area are treated. In addition, non-NEP electrification was still occurring during this time in areas where households were willing to pay for their connections. Finally, not all households treated with a connection may be able to afford this new option. With measurement error in a binary variable, the estimate of the treatment effect $\hat{\alpha}_{2,OLS}$ would be downwards biased, and IV will tend to be upwards biased as long as the measurement error is not too large.[31]

I take two approaches to dealing with these three sources of bias. I include a vector of baseline (measured in 1996) covariates (X_{jd0}) to control for some factors affecting a community's growth path (δ_j). These include 1996 household density, fraction of households living in poverty, adult sex ratio (female/male), fraction of female-headed households, distance to the 1996 grid, distance to the nearest road and town, fraction of adults that are white or Indian and measures of adult educational attainment in the area.[32] The Census income measure is crudely captured (in intervals and only at the household level) and so the share of female-headed households and the female/male sex ratio provide additional information about the extent of poverty in a community.[33]

Confounding trends and unmeasured political factors are still of concern.[34] Most individuals in this area are Zulu and Zulu-speaking, precluding the construction of an ethic- or linguistic-heterogeneity index which is often used to analyze political allocation of public goods. To overcome this issue, I instrument for program placement using average community land gradient (Z_j). The system of equations to be estimated is then:

$$\Delta y_t = (y_{jdt+1} - y_{jdt}) = \alpha_1 + \alpha_2 \Delta T_{jdt} + X_{0jt}\beta + \lambda_d + \delta_j + \Delta\epsilon_{jdt} \quad (3)$$

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

[31]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 arising from a mis-measured binary treatment variable.

[32]The September 2001 Labor Force Survey shows that white and Indian adults between ages 20 and 70 with at least a grade 8 level of education are disproportionately likely to be employers.

[33]Standing et al (1996) argue that these two variables are appropriate poverty indicators in former homeland areas. Since many of these areas historically sent male migrant labor to the mines, sex ratios were (and still remain) highly skewed towards women, and many households are headed by women whose partners migrate for work. These rural economies rely heavily on migrant remittances and pensions for income, and are some of the poorest communities in the country.

[34]One alternative would be to use a third difference to eliminate unobservable economic growth trends. This is not possible with only two waves of data. It is also not sensible in this context where the transition to democracy occurred in 1994, bringing with it new national governance and national policies.

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 to an electrification project.

In the data, we will see that gradient is correlated with some of the baseline covariates. To bolster confidence in the research design, I test for whether gradient is correlated with outcomes for areas treated before 1996 and with changes in other development services (water, sanitation). I also test for whether gradient is correlated with changes in possible sources of employment. In each case, the data cannot reject zero correlation.

The obvious concern with using land gradient as an instrument is that it may affect agricultural outcomes or characteristics of individuals settling in steep and flat areas.[35] In rural KZN, the direct impact of gradient on agricultural productivity and agricultural employment growth is limited, since most people are not farming. Under 10 percent of employed individuals are involved in agriculture (see Figure 3, constructed from 1996 Census micro-data).[36] In addition, validity of the instrument in (3) is threatened only if non-random sorting of individuals across flat and steep areas resulted in differential employment growth, outside of the effects of new electrification. For example, if flat areas attract more productive people over time for reasons unrelated to the expansion of the grid, the exclusion restriction may be violated. Mobility within homeland areas during this time period is limited by a lack of property titles and the role that tribal authorities (rather than the market) play in allocating land.[37] Using information on former place of residence, I show that results do not appear to be driven by differential migration.

Conditional on instrument validity, it is worthwhile considering the interpretation of $\alpha_{2,IV}$. This parameter captures the local average treatment effect (LATE) of electricity projects on employment

[35]Gradient is sometimes used as a control in the estimation of 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).

[36]Simkins (1981) documents the collapse of farm yields in homeland areas under the pressure of rapid population expansion due to forced resettlements and high birth rates. This occurred in the 1960s. Standing et al (1996) and Aliber (2001) describe how agriculture does not generate the majority of income in the homelands. Under one-third of households in this area have access to land for farming (own calculations, October Household Survey 1996).

[37]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 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).

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. However, individuals aggregate to communities and so community composition will drive 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. In addition, employment "returns" to electrification may differ by gradient itself, leading to larger estimated employment effects for marginal than for average communities. Flatter areas have lower fixed commuting costs, so people in flatter areas always face a higher outside wage, net of transportation costs.[38] Individuals with higher outside wages will then be more likely to respond when electricity arrives, as they are initially closer to the participation margin.[39] Hence, there are several reasons to expect that IV estimates will be different than average treatment effects for a more general population.

6 Main Results

6.1 Summary statistics

Table 1 presents summary statistics on baseline covariates for the sample of 1,992 communities. Overall, these areas are poor: 61 percent of households live on less than 6,000ZAR per year (approximately USD840 at a 2006 USD/ZAR exchange rate).[40] On average, over half of households are female-headed and the female/male adult sex ratio is well over 1. These variables reflect the historical function of these homelands as migrant-labor communities.

Column (4) of Table 1 reports differences in means by treatment status. Compared to control areas, treated communities are somewhat less poor, have higher fractions of female-headed households, high-school educated men and women and are about 2.8 kilometers (1.7 miles) closer to

[38]Working individuals are significantly more likely to take public taxis or trains to work if they live in flatter areas (Census 2001).

[39]That gradient is correlated with transportation costs is one additional 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 a discussion of how high transportation costs driven by terrain ruggedness allowed some parts of Africa to escape the effects of the slave trade by being hard to reach. To test whether access to roads is driving the response in KZN, I restrict to areas without any main roads. Although the coefficient on treatment falls slightly in the female employment regressions, there is no significant difference in results (not shown here, available from author upon request).

[40]This is the cut off for households in the two lowest income brackets reported in the Census. No household per capita income measure is available to compute a finer poverty measure.

the nearest road and town. Given that low average cost areas were prioritized for projects, it is not surprising that treated areas are higher density in 1996, are about 4.5 kilometers (2.8 miles) closer to the nearest substation, and have a 2.4-degree flatter average gradient than control areas.

Table (1) illustrates some of the difficulties with inferring causality by comparing treatment and control groups, as well as the hope that an IV strategy promises. In a randomized experiment, all observable characteristics should be balanced across treatment and control groups. However, treated areas are significantly closer to towns and roads and more densely settled. The last column shows that gradient does an excellent job of balancing the community poverty rate, the distance to town and road variables as well as distance to the grid.

Male and female employment rates and population totals by treated and control areas are shown in Table 2. The main outcome variable is the employment to population rate of Africans aged 15 to 59. 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. Comparing changes in employment rates in treated areas to the same change in control areas, the unadjusted estimate for women is not different from zero. For men, it is a statistically significant -2 percentage points.

Population is also growing faster in treated than in control areas, even though they begin with higher populations. Treated areas grow at about 6 percent per year and control areas are growing at about 3 percent. Communities are small, so a 3 percent change in population over five years is an increase of about 30 people in the median community. Both in- and out-migration are occurring over this period (Leibbrandt et al, 2003), and migration itself could be a response to treatment.[41]

Two striking points emerge from Table 2: employment rates are low, and are falling for men. In these low employment communities, many households are supported by migrant remittances or state old-age pensions.[42] Large reductions in male employment in treated relative to control areas reflects what was happening more broadly in the South African labor market during the 1990s. Banerjee et al (2006) document large shifts in the composition of jobs away from commercial agricultural and

[41]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.

[42]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.

mining sectors, and towards service and retail sectors. These changes were largely a continuation of trend, impacting on male-dominated sectors. Falling male employment rates in Table 2 reflect less any negative impact of electrification and more the fact that electricity is being placed in areas doing more poorly over time.

A second key feature of the South African labor market during this period is that large numbers of African men and women entered the market (Banerjee et al, 2006; Casale and Posel (2004)). On average, female labor force participation increased by 10 percentage points between the mid-1990s and 2001, with even larger increases in rural areas.[43] The types of new jobs created during this time were predominantly low skill and in the informal sector (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.

Labor market opportunities are therefore changing for all men and women over this period. Table 2 reflects that this restructuring is correlated with treatment status. Treated areas are closer to towns and formal jobs, near existing infrastructure and have higher household density– all factors which raise the community's exposure to these changes in industry. Political reasons for targeting household services investments towards disadvantaged areas would reinforce this correlation. Since very little of the restructuring was associated with changes in subsistence agriculture, there is less concern that gradient is similarly correlated with these changes. The key point is that the impact of electricity as identified by the gradient instrument should be interpreted as the impact on employment within a context of changing economic opportunities. This not unlike the way that the effects of changing constraints on female labor market participation in the US have been interpreted.[44]

6.2 OLS and IV main results

First-stage estimates for assignment to treatment are presented in Table 3.[45] A one standard deviation increase in gradient (about 10 degrees) reduces the probability of being treated by 4

[43] Here, labor force participation refers to employed and unemployed women looking for work.

[44] For example, effects of new fertility-control technologies or falling appliance prices are analyzed during periods in which World Wars are happening, social norms are changing, and the structure of jobs available for women are being altered. See for example, Bailey (2006), Goldin and Katz (2000) and Greenwood et al (2005).

[45] Results from a logit model of the treatment are very similar to these linear probability model results.

percent. The size of the coefficient does not change with the addition of more controls and precision improves, particularly with district fixed-effects. When restricting to areas where no households had electricity in 1996, the gradient coefficient is slightly larger. A community that is transected by a line connecting any two substations is also significantly more likely to be treated. The other definitions of treatment convey the same message: a 10 degree increase in gradient reduces the probability of being treated early in the period by 12 percentage points, and decreases the fraction of households treated by 2 percentage points. For the main analysis, I use the treatment indicator variable and instrument with mean gradient since this provides the strongest first stage.[46]

Other cost coefficients have the expected signs: a one standard deviation increase in distance from the grid (about 13 kilometers) reduces the probability of treatment by 2 percent. A one standard deviation increase in household density (30 households) per square kilometer increases the probability of treatment by 1 percent.

The first stage provides mixed evidence on whether treated areas are selected on wealth. While areas with more female-headed households (i.e. poorer) are less likely to be treated, areas with more white and Indian adults (i.e. richer) are also less likely to be treated.[47] The community poverty rate and sex-ratio variables also have positive signs in most specifications, suggesting that treatment is being assigned to poorer areas. This lack of strong evidence for project placement going to richer areas is consistent with the overarching socio-political motivation for the roll-out.

Electricity projects did change patterns of household fuel use. Each cell in Table 4 presents coefficients from a different OLS or IV regression of (3), where the outcome is the change in the fraction of households using electricity for lighting or cooking. Average electrification rates rise by 23 percentage points more in treated areas. Reliance on wood for cooking falls by 4.2 percentage points and cooking with electricity rises by 6 percentage points. In areas induced to be treated by virtue of having a flatter gradient, use of electric lighting increases by a substantial and significant 71 percentage points, wood use falls by 28 percentage points and cooking with electricity rises by 24 percentage points.

[46]Appendix 2 contains some sensitivity tests for different sub-samples, definitions of treatment and instrument combinations.

[47]Note that coefficients on education and share of Indian and white adults are small once scaled by the appropriate mean in Table 1.

Treated regions do not appear to have differential changes in other basic services that could affect home production. In fact, IV results for water services 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.[48] These results demonstrate two points: gradient does not capture access to development projects more generally, and instrumented employment responses could be large, since the effects of treatment on household fuel use are much larger than treatment-control comparisons.

Employment effects for men and women are consistent with this latter point. Tables 5 and 6 present OLS and IV results for African women and men. The dependent variable is the change in sex-specific employment rate between 1996 and 2001. Column (1) reflects the falling employment rates from Table 1. Adding controls and district fixed-effects increases the coefficient on treatment slightly. Female employment grows faster in poorer places, indicated by the positive and significant coefficients on community poverty rate, sex ratio and female-headed households. It is reassuring that the coefficients on variables other than treatment are consistent in sign and magnitude for male and female regressions. For instance, the poverty variables all have the same sign and significance in OLS and IV results.

IV estimates of the treatment effect are substantially larger than OLS estimates, and significantly positive for women. Since gradient is correlated with some of the control variables as Table 1 indicated and the first stage is strongest with other controls absorbing residual variation, it is preferable to focus on results in columns (6) to (8). Female employment increases by 13.5 percentage points in areas induced to get the treatment by gradient. 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.[49] The AR test for women strongly rejects zero and the confidence interval is wider, between 5 and 40 percentage points. Male employment increases by a substantially smaller 4.2 percentage points, and this is not significantly different from zero under either test. The reduced form for men is weak; there is

[48] Hemson (2004) explains that some of this worsening service provision is related to lack of maintenance of systems in some municipalities between 1996 and 2001.

[49] These tests are conducted because the F-statistic is just below 10 once all controls are accounted for. Heteroscedasticity-robust Anderson-Rubin confidence intervals have correct coverage properties in the presence of weak instruments while standard Wald tests do not. See Moreira and Cruz (2005), Mikusheva and Poi (2006) and Chernosukov and Hansen (2007).

essentially no relationship between gradient and employment for men. This observation is consistent with electricity having effects on the employment of primary home production workers and facilitating women's rather than men's entry into the market. The difference in the male and female employment effect is significant at the 14% level.[50] This positive, significant effect for female employment persists when using $T_{connect}$ or T_{year} as the treatment variable, and when using both gradient and the $Transect$ indicator as instruments (see Appendix 2, Table A2).[51]

In examining national trends in labor force participation in the 1990s, Casale and Posel (2004) argue that the increase in women's participation occurs because male employment (and hence household income) is falling. While this could be possible more generally, it is not clear why this effect would operate differentially across steep and flat areas. Moreover, if women increase participation on the extensive margin because men are losing jobs, we should expect to see negative and significant IV results for men at the same time that the female employment results are positive and significant. The results in columns (4)-(8) in Table 6 show that this is not the case.[52]

6.2.1 Measurement error in the treatment variable

Measurement error in the treatment variable could contribute to the difference between OLS and IV coefficients. OLS will underestimate the effect of treatment on outcomes when there is a negative covariance between δ_j and ΔT_{jt} (which I have argued is likely) and in the presence of measurement error. However, the valid IV that is uncorrelated with $\delta_j + \Delta\epsilon_{jt}$ will tend to be correlated with non-classical measurement error in the binary variable ΔT_{jt}. In this situation, even if the instrument deals with the omitted variables bias, the measurement error in ΔT_{jt} leads to an upwards biased IV estimator.[53]

[50] 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.

[51] In further sensitivity tests, I estimate median regression models of changes in male and female employment rates to reduce the impact of outliers on results. To adjust for endogeneity of treatment, I pursue a control function approach as suggested by Lee (2007) in which I include the residual from a first stage linear regression of treatment on gradient and other controls in the median regression of employment outcomes. I find that the female result is somewhat smaller, shrinking to about 0.08 (but still significant at the 5% level) and the male result is not significantly different from zero. The male-female difference is significant at the 12% level.

[52] The data do not contain good measures of wages or income values. For treated areas, the fraction of households living above the poverty line does increase, but not significantly so in either OLS or IV specifications. Results not shown; available upon request.

[53] This result is conditional on the measurement error in treatment not being too extreme (Kane et al, 1998).

To detect how much of the difference in OLS and IV results is due to measurement error, I restrict to samples where ΔT_{jt} should be measured with less error. Table 7 reproduces the main result for females in the full sample, and results for successive sample limitations. To identify communities where projects translated into large changes in home production, I exclude treated areas with less than a 10 percent change in coverage of electric lighting, and treated areas where the connection rate was under 80 percent of households. Under the first restriction in columns (3) and (4), the OLS coefficient rises substantially and the IV coefficient is slightly smaller than the main result at 12.6 percentage points. Columns (5) and (6) impose the second restriction. Again, the OLS estimate is large and positive and the IV result is almost identical to the main result, although neither is significant due to the smaller sample size.

Effects estimated under the OLS specification for these sub-samples are just shy of 2 percentage points for female employment and the IV effects are still larger. This provides some evidence that measurement error in the treatment variable is an issue. However, this measurement error alone is unable to account for the entire gap between OLS and IV estimates. OLS results are much more likely confounded by an unobserved community level effect.[54]

7 Threats to validity

7.1 Do flatter areas have different labor demand trends?

If employment rates in steep and flat areas evolve differently, the gradient IV would be invalid. Checking for differential trend is difficult without more years of data. This is where having the administrative data on electricity projects from 1990-2007 is helpful for conducting a placebo test. These data identify which areas are treated 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 been treated with an electricity project. By implication, a reduced-form relationship between employment growth and gradient *after* treatment would lead us to think that gradient has a direct effect on employment growth.

[54] I show in Appendix 2, that using continuous treatment variables which are less likely subject to the same type of non-classical measurement error still produces differences between OLS and IV results for female employment.

To test this, I select the sample of areas treated prior to 1996 ($N = 406$) and run an OLS regression of employment growth between 1996 and 2001 on the full set of controls, and gradient. Columns (1) and (2) of Table 8 contain the results of this placebo test. The coefficient on gradient is small (-0.001) and insignificant. There is no evidence of any such reduced-form relationship for male or female employment. The fact that we cannot reject the hypothesis that gradient has no effect on employment growth after treatment boosts confidence in the research design.

7.2 Do flatter areas have contemporaneous labor demand shocks?

Another potential threat to the validity of this research design 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.[55] For example, businesses may expand in homeland areas after the end of apartheid for reasons unrelated to electrification.

Information about the major employers of women in the area is useful for testing for some of these confounders. Figure 3 shows that professional occupations and elementary occupations contributed the majority of female employment in these areas. Data from the 10% micro Census sample (not shown here) indicate that 75 percent of African women in rural KZN working as professionals or associate professionals are teachers, while 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. Labor demand shocks in these two industries are the most likely candidates for confounding IV estimates of electrification effects.[56]

Using two waves of the South African Schools Register of Needs that fall just before each Census wave (1995 and 2000), I construct a variable measuring the change in the number of schools in each community over time.[57] Over the five-year period, the number of schools increases by 19 percent,

[55] A negative labor demand shock in steep areas would have the same confounding effect.

[56] In this rural area, it is also plausible that "high skilled" women could have time freed up from home production to take a teaching job. Census micro data indicate that 30 percent of high skilled women (with at least a high school certificate) live in households where the main cooking fuel is wood, and half of them live in households using candles for lighting. Such women are also involved in home production activities as the September 2001 Labor Force Survey indicates: almost 20 percent of women with a high school certificate (minimum requirement for teaching) report collecting wood for their households, conditional on anyone in the household collecting wood.

[57] The Schools Register of Needs provides GPS coordinates of each school. This allows me to allocate schools to communities using the Census spatial data.

from 1,770 to 2,801. This creates a higher demand for teachers across the province.

Table 8 shows results from a regression of the change in the number of schools on community gradient and all other controls. There is no significant relationship between gradient and the growth in schools over time. While school placement (and hence teacher hiring) is probably related to the distribution of children in space, this distribution does not appear to be correlated with gradient.

As a second indirect check that female employment is not being driven by an expansion of demand that happens to occur in flatter areas, I proxy for "new employment opportunity" using the change in the proportion of adult population that is Indian or white. These are the individuals most likely to hire household workers (Dinkelman and Ranchhod, 2007). The number of Indian and white adults is not changing differentially across areas of different gradient, as column (4) of Table 8 shows. There is no apparent increase in the number of potential employers of domestic workers in areas where electricity is rolling out.

It would be ideal to perform similar tests using a measure of the number of other firms in the area over time. However, these two tests alone are still very informative, since employed women are most likely to be working as teachers or domestic workers. While more jobs are being created over the period at the low end (Casale and Posel, 2004) and in the public sector, there is no evidence to show that these job openings are occurring differentially in flatter areas.

8 Channels

8.1 Does electrification stimulate demand for labor?

In these small communities, an electricity project that generates new demand for labor by stimulating the growth of firms is likely to have spill-over effects into neighboring areas. Spillovers could be positive or negative. For example, if firms create jobs for people living in neighboring areas, the positive spillovers in these control areas would dampen treatment effects. If people move out of neighboring control areas and in to treated areas to get one of the new jobs, this negative spillover would amplify treatment effects. In both cases, the treatment effect is the sum of an incumbents' effect and a spill-over effect. In both cases, OLS and IV coefficients should be different when adjacent control areas most susceptible to these spillovers are excluded from the analysis.

To test this, OLS and IV regressions are re-estimated after excluding control areas within a one- and five-kilometer radius of a treatment area.[58] 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: a coefficient of 0.114 could not be rejected in the full sample. This result suggests that there are no strong spillover effects between communities. Combined with the fact that roll-out was driven by household targets and capacity was too small to stimulate even mid-size manufacturing or service enterprizes (South African Department of Minerals and Energy, 2004), the lack of evidence for spillovers suggests that the net employment effect of electrification operates primarily through a labor supply channel.

8.2 Heterogeneous Effects: Marginal Communities

IV estimates apply to marginal 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 treated community. Recall that the expansion of infrastructure did not entail free electricity. So, one way in which marginal communities could differ is that they could contain more households able to switch to using electricity when the new service arrives.

The Census provides only a crude measure of community poverty. To decompose the wealth characteristics of communities most affected by the gradient instrument, I combine the three poverty indicators into a poverty index and consider the characteristics of communities in each quintile of this index. For the sample of communities in the steepest half of the gradient distribution, I predict the probability of treatment using baseline poverty rate, the female/male sex ratio and the 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.[59]

The graph in Figure 4 shows the fraction of the predicted poverty quintile treated, for communities in the flattest and steepest halves of the gradient distribution. Areas with higher predicted values of the treatment are more likely to actually be treated (both lines slope upwards)

[58] 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.

[59] This procedure follows Card (1995) and Kling (2001).

and flatter areas are systematically more likely to be treated (there is a gap between the two lines). The third and fourth quintiles are most likely to have treatment probability manipulated by the instrument. In Table 10a, the contribution of each quintile to the IV estimate is given by the IV weight: this is the amount that each quintile contributes to the computation of the IV estimate.[60] Together, the middle quintile and the second richest quintile contribute over 70 percent to the IV result.

Why might middle-quintiles in particular have large employment effects? These communities contain households that have large changes in home production technology when electricity arrives. Middle-poor areas are initially less likely to be using electricity or other modern fuels than richer areas, as columns (1) and (2) of Table 10b indicate. They also appear to be more likely to switch to using electricity when it arrives than poorer communities. Columns (4), (5) and (6) of Table 10b present within-quintile reduced-form coefficients from regressions of the change in fuel use for different home production activities. These columns indicate large increases in the use of electricity and large decreases in reliance on wood for cooking in flatter middle-poor areas.[61] This is consistent with the result in column (6) of Appendix 2 Table A2. When restricting to the set of communities where no household had electricity in 1996, the female employment effect of treatment is smaller, at 7.7 percentage points. The smaller point estimate is sensible, since half of the communities in this restricted sample are drawn from the poorest two quintiles of the poverty index, compared to only one-third of communities in the full sample. In areas completely without power in 1996, there are simply fewer households able to respond to the new access to electricity. Finally, column (7) of Table 10b indicates that the female employment result is indeed driven by women living in middle- and second-richest quintile communities.[62]

[60] The weights are constructed as in Kling (2001) and their computation is explained in the table notes.

[61] 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.

[62] 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 10a will give the IV coefficient.

8.3 Heterogeneous Effects: 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. For example, child-care responsibilities raise the value of a woman's time at home, and in the absence of pre-school care (which most of these rural areas do not have), 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).

Census micro data from 1996 give some indication of which women are more likely to live with a child younger than age 9. Figure 5 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 younger and older women.[63] After age 30 and up to about age 50, the probability of a woman living with a child who requires constant care falls substantially. We should expect the employment effects of electrification to be largest for women in these age groups.

To investigate this channel, 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 of the age cohorts: the estimated coefficients sum to the main treatment coefficient in column (8) of Table 5. Table 11 presents OLS and IV coefficients on the treatment dummy for separate regressions.[64] IV results are larger and positive for each age group, but significant only for women in their thirties and late forties. Employment grows by 3.9 percentage points for women between the ages of 30 and 34, by 2.6 percentage points for the 35 to 39 year old group and by a smaller but still significant 1.9 percentage points for the older age group. Together, these age groups account for 65 percent of the total female employment result.

Since aggregate data do not allow me to identify exactly which women have children of different ages, I capture these other demands on female time as the ratio of the number of children ages 5 to 14 in the community in 2001 to the number of households in 2001. These children will have been

[63]The allocation of young children to households with older women is a common pattern in South Africa, where pension-age women care for grandchildren in skip-generation households.

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

ages 0 to 9 in the period between 1996 and 2001– just those ages that require full-time care. I control for this historical variable in the main regression and interact it with treatment and gradient to examine heterogenous treatment effects on female employment. Table 12 presents results.

This more direct test of the child-care channel asks a lot of the aggregate data. However, there is some evidence in both OLS and IV results that the treatment effect is attenuated in areas with a higher fraction of young children to households. At the mean of this variable, the interacted coefficient implies that a 1 percentage point increase in the ratio of young children to households reduces the treatment effect by 0.5 percentage points (-0.537*0.01). Adding the interaction coefficient at the mean to the treatment coefficient, the treatment effect of electricity projects is just over 9 percentage points. At least in the short-run, additional constraints on female time appear to reduce the impact of a new home production technology.[65]

8.4 Migration

Out-migration from rural areas is occurring during this period of roll-out.[66] Cross et al (1998) also document rural-to-rural migration in KZN in the 1990s and show that some of this is migration towards areas with better infrastructure and amenities. Each flow could alter the composition of the population in treated and non-treated communities, and contribute to employment effects in OLS and IV specifications.

Aggregating community-level information to the district level (as this is the only level of geography available in the micro-data), the 10% sample of the 2001 Census micro data indicate substantial out-migration rates: as a percentage of the 1996 population, 15 percent of men and 10 percent of women report out-migration during the five-year period.[67] These out-migration rates are,

[65] Over a longer period of time, fertility may adjust in response to more efficient home production. Greenwood et al (2005) suggest that in the USA, home appliances reduced the cost of child-care and encouraged families to have more children, contributing towards the post-war baby boom. Fertility may also decline as alternative ways of using time (leisure) open up. In a five year period, such large changes in fertility are unlikely. I tested for a fertility effect of treatment using the number of young children per women as the outcome variable, and found no evidence of this.

[66] Leibbrandt et al (2002) find that men with some education tend to leave rural areas, and both the least and most educated men remain.

[67] 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 in the 10 percent sample. Results not shown.

however, not significantly different by gradient.[68]

Table 13 presents differences in population growth rates in treatment and control areas. Even numerically small increases in population can translate into large percentage changes in these small areas. The first two columns indicate that treated areas have significantly higher population growth than control areas, both in the OLS and IV results. Over the five-year period, treated areas grow 25.8 percent faster than control areas, and this growth is 400 percent in the IV result.[69]

Ideally, we could isolate employment growth for in-migrants and incumbents. However, the aggregate Census data does not allow me to identify the migrant status of employed individuals. Instead, to move part way towards an understanding of how much in-migration contributes to the employment result, I assume all individuals who report themselves as recent in-migrants 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. This new variable captures lower bound changes in employment rates for incumbents only.[70]

Table 13 provides results for men and women. For women, OLS and IV results are remarkably similar across the full definition of employment and the migrant-excluded definition. A Hausman test on the treatment coefficient across each specification of the female employment variable cannot reject that they are the same. Once again, the AR test rejects a zero effect for women but now with a tighter confidence interval from 0.05 to 0.3. This suggests that differential in-migration of employed women cannot account for the entire female employment effect. For men, redefining the employment variable in this way raises the IV point estimate somewhat. Male employment is 8.4 percentage points higher in treated regions compared to non-treated regions, but still not statistically significantly different from zero.

Within the limitations of the aggregate Census data, there is some evidence that in-migration of individuals towards treated areas may be an additional response to electricity projects. Given

[68]Ideally, we could test directly whether gradient predicted out-migration from communities in my sample rather than using the more aggregated individual-level data. Unfortunately, the community data do not contain information on prior place of residence, and so this test is not possible.

[69]Pirouz (2004) documents a 50% increase in the number of households in South Africa over the 1995 to 2002 period, with a concomitant reduction in average household size.

[70]It should be noted that the in-migration data are far from perfect. Individuals are asked "Were you living in this place 5 years previously?" This leaves room for a wide interpretation of 'this place'. It is still useful doing this exercise, as a significant difference in results would indicate a substantial migration response to treatment.

reported preferences for household services, this effect is not surprising.[71] However, there is no strong evidence to suggest that in-migration or out-migration contaminates the IV estimates.

9 Implications of employment results

IV results suggest that in a non-treated community with the median number of adult women in 1996 (N=254), a 13.5 percentage-point increase in female employment raises the number of women working from 18 to 52. If we assume this 13.5 percentage point increase applies to the entire treatment group (rather than marginal communities only), this translates into an increase of 22,487 newly employed women out of the baseline female population of 166,574. This is 1.1 percent of the estimated 2 million new jobs created across the country over the period (Casale and Posel, 2004).

To perform a careful cost-benefit analysis of KZN's rural electrification, additional information about benefits (impacts of electrification on deforestation, and of lighting on safety, on literacy, on human capital accumulation) and costs (environmental costs of electricity generation, maintenance costs, dead weight loss of raising revenue through taxation) would be required. Nevertheless, using the estimated employment responses along with information on connections costs, we can perform a back-of-the-envelope calculation to assess whether this rural electrification was a worthwhile use of public funds.

Table 14 sets out the relevant parameters used for this calculation. With approximately 470,000 households connected during the period, and an average connections cost of ZAR2,533 (in 2001 South African Rands), the total cost of the project is close to ZAR1.2 billion.

Using median unskilled annual earnings (ZAR3,600) and assuming that most individuals will work on average only one-quarter of a full-time job (if, on on average, two to three hours per day are saved in fetching wood and cooking), then each job earns ZAR900 per year. Summing up over all jobs, earnings generated by this project amount to just over ZAR20 million. Without applying any discount rate, the value of the stream of benefits over five years is just over ZAR100 million. Even summing benefits over a 15 year period generates only ZAR300 million in earnings– discounting would further lower this benefit stream. Hence, it appears that electrification would not generate

[71] In a recent household survey conducted in a rural part of the country by Fort Hare Institute of Social and Economic Research (2007), individuals ranked electricity in the home as the second most important service (after water).

positive social returns if we value only the employment facilitated by the new infrastructure.

One intuitive alternative to valuing benefits is to assume that all women in treated areas have their time freed up by between two and three hours per day with the arrival of electricity. Some of these women export this time to market work, while others may take the additional time as leisure. If we value the total number of potential hours saved by electrification at the same median wage for unskilled workers (assuming that the women who take this extra time as leisure will value this leisure at least at their market wage) then the project creates annual benefits of almost ZAR150 million. Again, without applying any discount factor, the project only yields a positive social return after ten years. Using this ten year time frame, the provisional internal rate of return for this investment is just over 5%.[72] This is far below the typical range of discount rates used in World Bank infrastructure project evaluations (usually 10-12%).

10 Conclusion

This paper uses the experience of household electrification in South Africa to measure the direct effects of public infrastructure on employment in rural labor markets. I combine hand-collected administrative and spatial data on electricity project roll-out with aggregate Census data to estimate large increases in electric lighting and cooking, and reductions in wood-fueled cooking over a five-year period. Consistent with one prediction from a simple model of technological change in home production, female employment rises by 13.5 percentage points in treated areas, and there are no significant effects for male employment. The female employment response is driven by middle-poor communities that initially rely heavily on wood for cooking and are able to respond more to the new service. Effects are also larger for women in their thirties and forties, and there is some evidence that this is related to fewer child-care responsibilities at these ages.

Although I cannot rule out a labor demand channel for this employment effect, it is unlikely that electricity-driven increases in labor demand explain all of this response. Electricity projects provided power supplies too small for industrial use, and there is little evidence of demand-related spillovers across communities. Before extrapolating this result to other energy-poor societies, it would be

[72] The internal rate of return is the discount rate that sets the discounted flow of benefits equal to the discounted flow of costs.

important to test for similar effects in areas with more equal sex ratios, given the history of male out-migration for work in KZN. However, if the primary channel through which electricity operates is by freeing women from home production, we should see similar effects, even if more men are present. In addition, these results must be interpreted as the response to household electrification during a period of economic restructuring. This type of interpretation is not unusual in the literature on historical labor force participation of women; changing constraints on women's ability to work generally occur within the context of broader changes in the economy. In most countries, infrastructure expansion will also be accompanied by changing economic conditions, and the challenge is to find ways to measure effects within this environment. In my study, a placebo test provides some evidence that general changes in the South African labor market are unlikely to be confounding the IV estimates of the effects of electrification.

These results represent some of the first pieces of evidence on the impact of infrastructure for rural electrification in a developing country. They highlight the importance of measuring employment effects in infrastructure evaluations more generally. The findings also suggest that paying attention to heterogeneous treatment effects across different types of communities could yield insights into which groups benefit from infrastructure expansion in the short-run. 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 provide more actual variation in programs than legal changes. Finally, even though a crude analysis of direct benefits and costs suggests that the investment may not have been worthwhile over a short time-frame, the finding that female employment in rural areas of South Africa responds to household electrification contributes to a large literature on how patterns of female labor force participation have responded to changing constraints in developed countries. More detailed individual-level data collected in countries where these constraints are currently being relaxed will help us to learn more about longer-run impacts of this infrastructure on labor markets: for example, what the effects of household infrastructure are on the intensive margin of work, how long it takes the poorest women to accumulate complementary appliances that make them more productive in the home, and what jobs and occupations women first choose to enter when they are able to use their labor outside of home production.

References

Akee, Richard, "Road to Palau," *IZA Working Paper Number 2452*, 2006.

Aliber, Michael, "Study of the incidence and nature of chronic poverty and development policy in South Africa: An overview," Technical Report, Programme for Land and Agrarian Studies, School of Government, University of the Western Cape 2001.

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. Manuscript.

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

Bank, The World, "Lighting Africa Initiative," Online: www.lightingafrica.org 2007.

Barnard, Jenny, "An Investigation into Using GIS in Electrification and Network Planning in Rural Kwazulu-Natal," Technical Report, GISDevelopment.net 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), 4490476.

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 2001.

Canning, David, "A Database of World Infrastructure Stocks," *World Bank Policy Research Working paper*, 1998, (1929).

Card, David, *Aspects of Labour Market Behaviour: Essays in Honor of John Vanderkamp*, Toronto: University of Toronto Press,

Cattaneo, Matias, Sebastian Galiani, Paul Gertler, Sebastian Martinez, and Rocio Titiunik, "Housing, health and happiness," Technical Report 4214, World Bank Policy Research Working Paper 2007.

Charmes, Jacques, *Gender, Time Use and Poverty in Sub Saharan Africa*, 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.

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

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

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.

Deutsche Presse-Agentur, "Middle East, North Africa to see large energy investments," Online Newspaper June 2006.

Dinkelman, Taryn and Vimal Ranchhod, "Compliance with minimum wage laws when penalties and enforcement are weak: Domestic workers in South Africa," 2007. Manuscript, University of Michigan.

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

EnergyNet Limited, "Rural Energy 21: Bringing Power to Africa," in "in" $www.energynet.co.uk/Infosite/images/RE21_report_2004.pdf$ 2004.

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.

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.

Galiani, Jim Levinsohn Ingrid Woolard Abhijit Banerjee Sebastian, "Why is unemployment risen in the new South Africa?," 2006. Working paper.

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." Department of Electrical Engineering, University of Cape Town May 2003.

__ , "Person communication," May 31 2006. Head of Department of Electrical Engineering at University of Cape Town.

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 Guillaume Vandenbrouke, "The baby boom and the baby bust," *American Economic Review*, 2005, *95* (1), 183–207.

__ **,** __ **, and Mehmet Yorukoglu**, "Engines of Liberation," *Review of Economic Studies*, 2005, *72*, 109–133.

Gronau, Reuben, *Handbook of Labor Economics*, Vol. I, Elsevier Science Publishers, 1986.

Hemson, D., "Beating the backlog: meeting targets and providing free basic services," Technical Report, Human Sciences Research Council 2004. Position paper for the National Treasury.

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

Jimenez, Emmanuel, *Handbook of Development Economics, ch 43*, Vol. 3B, Elsevier Science,

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," *NBER Working Paper 7235*, 1998.

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.

Lee, Sokbae, "Endogeneity in quantile regression models: A control function approach," *Journal of Econometrics*, 2007, *141*, 11311158.

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.

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.

Michael, Robert T., "Education in non-market production," *Journal of Political Economy*, 1973.

Mikusheva, A. and C. B. 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, M. and L.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.

Netgroup, "Universal Access Program - Electrification Stream," Presentation to Eskom Eastern Region, Planning Workshop 3 October 2006.

Nunn, Nathan and Diego Puga, "Ruggedness: The blessing of bad geography in Africa," March 2007. Working paper.

of Energy, United States Department, "Average retail price of electricity to ultimate customers by end-use sector," Technical Report, Department of Energy 2007.

Piper, Laurence, "Democracy for a bargain: the 1999 election in KwaZulu-Natal," *Politikon*, 1999.

Pirouz, Farah, "Have labour market outcomes affected household structure in South Africa? A preliminary descriptive analysis of households," October 2004. Conference on African Development and Poverty Reduction: The Macro-Micro linkage, South Africa.

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." Mimeo, Brown University.

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," June 2006. Working paper.

Saghir, Jamal, "Energy and Poverty: Myths, Links and Poverty Issues," Technical Report, The World Bank 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. Under submission.

South African Department of Minerals and Energy, "National Electrification Programme

Summary Evaluation Report: 1994-1999," Technical Report, South African Department of Minerals and Energy 2001.

__ , "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*, International Labour Office, 1996.

Statistics South Africa, *Census Mapping Manual 2000.*

Tavares, Jose and Tiago V. de V. Cavalcanti, "Women prefer larger governments: Female labor supply and public spending," 2007. Mimeo, Universidade Nova de Lisboa, Portugal.

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 Nations, "The Energy Challenge for Achieving the Millenium Development Goals," Technical Report, United Nations 2005. Accessed on 8/8/2006.

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. Mimeo, South African Labour and Development Research Unit.

West, N A., B. Dwolatzky, and A.S. Meyer, "Terrain-based routing of distribution cables," *IEEE Computer applications in power*, January 1997.

Wittenberg, Martin, "The Weight of Success: The Body Mass Index and Economic Well-being in South Africa," October 2007. Working Paper, South African Labour and Development Research Unit, University of Cape Town.

Appendix 1: Data

A: Census data

Census community data 1996 and 2001: Available from Statistics South Africa (at www.statssa.gov.za). Proprietary software (Supertable) 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 addition, Statistics South Africa provided me with 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, only rough measures of employment are captured. 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, and contains between 30,000 and 50,000 households in the KZN region.
- Sub-district or main place: these correspond to loose groupings of towns and surrounding areas. These boundaries have no administrative meaning.
- Community or sub-place: this is the lowest unit of analysis possible in the 2001 Census data. Average community size is small (200-250 households on average).

The 2001 community boundaries define the main unit of analysis. I aggregate (smaller) 1996 areas up to (larger) 2001 boundaries, assuming a uniform distribution of people over the 1996 areas that span 2001 boundaries.

As in most countries, boundaries in South Africa have shifted over time.[73] There are two aspects of these boundaries that make working with the Census data challenging. First, the 1996 data is available at the Enumeration Area (EA) level, which is smaller than a US Census tract. These areas contain up to about 250 households. The 2001 Census data is not available at the EA level for confidentiality reasons - the data is only released at the Sub-Place level (SP) which is an aggregation of 2001 EA's (and more like a US Census tract). In addition, between 1996 and 2001, some EA boundaries were re-drawn, meaning that some of the 1996 EA's span the 2001 EA boundaries. Statistics South Africa notes that EA boundaries should never cut across existing administrative boundaries, and all "social boundaries should be respected".[74] 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".[75]

[73]See Christopher (2001) for a good discussion of these changes.
[74]StatsSA (2000)
[75]StatsSA (2000: 21, 26).

This suggests that the 2001 EA's are more appropriate settlement areas than the 1996 EA's. Since I aggregate up to sub-place level anyway, any 1996 EA's that were merged together to make a 2001 EA do not pose a problem. Rather, it is the split EA's that may lie partially within a sub-place that could be problematic. I create the panel in the following ways, using spatial software (ArcGIS 9.2): I assign to each 2001 SP all of the 1996 EA's with which it intersects. This is a many-to-many mapping, as some SP's will contain more than 1 EA and some EA's will fall into multiple SP's. For each EA, I calculate the proportion of the EA polygon are that falls inside each SP. I use this proportion as a weight to assign some of the 1996 EA data to the 2001 SP for EA's that span 2001 boundaries. In order for this matching exercise to yield correct measures of sub-place aggregates, I assume a uniform distribution of people over the 1996 EA. Once the panel of areas has been created, I use the matched identifiers to create Census aggregate data in 1996 and 2001.

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 be assigned to district boundaries (but not sub-district or community boundaries).

B: 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. I used digital elevation model data to construct measures of average land gradient using GIS software (ArcMap 9.1). The procedure works as follows: for each pixel on the image representing a 90m interval, there is an associated elevation (above sea level) point. The elevation data are captured digitally by a radar system that flew onboard the Space Shuttle Endeavor in February of 2000. For each pixel, the maximum rate of change is calculated between itself and its 8 adjacent neighbors. Mean gradient per community is created by averaging over these measures across all pixels falling inside each Census community. I also calculate the variance of gradient points for each community, the range and the majority of points in each area. Gradient is measured in degrees from 0 (flat) to 90 degrees (vertical).

Other measures of proximity: Eskom's 1996 grid network was provided to me by Steven Tait. I observe the geographic location of all power lines from the highest voltage (400kV) to the lowest voltage (33kV) in this year. I also observe the position of each sub-station, a necessary piece of infrastructure for stepping down electrical current to household-use voltage. I spatially merge the grid information with the Census geography 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. These distances are then merged with the aggregate Census data.

C: Electricity project data

I collected these data from Sheila Brown at Eskom. This list consists of the number of pre-paid electricity connections per Eskom area by year from 1990 to 2007. The year of treatment is defined as the year in which a community experienced a spike in household connections in one year. This indicates a concentration of 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, I first mapped the project data to a physical location (using a spatial database of transformer codes that corresponded to project codes) and then matched these locations back to Census regions. A list of Census communities containing these generated treatment variables will be made available on my web site, once I have permission to publish these data.

D: Schools Register of Needs

These data are available from the South African Department of Education. A facilities survey was conducted for all schools in South Africa in 1995 and 2000. I use the GPS co-ordinates for each school to match schools to Census community boundaries. Each community is assigned the total number of schools in each year, and the change in the number of schools between the five years.

Appendix 2: Empirical specification and sensitivity analyses

Sensitivity to Non-linear Specification

The main equation is:

$$\Delta y_{jdt} = \alpha_1 \Delta T_{jt} + \alpha_2 + \delta_j + \lambda_d + \Delta \epsilon_{jdt} \tag{A-1}$$

The partial derivative in which I am interested is the aggregate employment elasticity in response to new access to electrification infrastructure:

$$\frac{\partial \Delta y_{jdt}}{\partial \Delta T_{jdt}} = \frac{\partial y_{jdt}}{\partial T_{jdt}} = \alpha_1 \tag{A-2}$$

In this linear model, the treatment effect has the same marginal impact everywhere. A more realistic specification would allow the treatment to have different effects depending on initial employment rate. Specifying y in logistic form would satisfy this requirement.

$$y_{jdt} = \frac{e^{\alpha_0+\alpha_1 T_{jt}+\alpha_2 t+\mu_j+\delta_j t+\lambda_d t+\epsilon_{jdt}}}{(1+e^{\alpha_0+\alpha_1 T_{jt}+\alpha_2 t+\mu_j+\delta_j t+\lambda_d t+\epsilon_{jdt}})} \tag{A-3}$$

Re-writing as $(\frac{y_{jdt}}{/(1-y_{jdt})})$ and taking logs of both sides delivers the linear form which can then be differenced to eliminate μ_j:

$$ln(\frac{y_{jdt}}{(1-y_{jdt})}) = \alpha_0 + \alpha_1 T_{jt} + \alpha_2 t + \mu_j + \delta_j t + \lambda_d t + \epsilon_{jdt} \tag{A-4}$$

$$\Delta(ln(\frac{y_{jdt}}{1-y_{jdt}})) = \alpha_1 \Delta T_{jt} + \alpha_2 + \delta_j + \lambda_d + \Delta \epsilon_{jdt} \tag{A-5}$$

With this transformation of the dependent variable, we can still implement OLS and IV; but calculating average marginal effects of the treatment is difficult. If we let $\Lambda(.) = \frac{e^{\alpha_0+\alpha_1 T_{jt}+\alpha_2 t+\mu_j+\delta_j t+\lambda_d t+\epsilon_{jdt}}}{1+e^{\alpha_0+\alpha_1 T_{jt}+\alpha_2 t+\mu_j+\delta_j t+\lambda_d t+\epsilon_{jdt}}}$. Then,

$$\frac{\partial y_{jdt}}{\partial T_{jt}} = \alpha_1 \Lambda(.)(1-\Lambda(.)) \tag{A-6}$$

The marginal effect for each community depends on initial values of $\Lambda(.)$. This marginal effect is difficult to calculate since neither α_0 nor μ_j is estimated. As an alternative, I calculate the marginal effect for the average community by using $\bar{y_{jd0}}$. $\hat{\alpha_1}\hat{\Lambda}(.)(1-\hat{(}\Lambda(.)))$*100 is then the *percentage point* increase in community employment in response to electrification, for communities at the average employment rate in period 0. Estimates using the non-linear specification of the dependent variable will differ from the linear specification when the data are not all close to 0.5. In my sample, very few areas have employment rates near or above 0.5, and many areas have values clustered close to zero.

Table A1 compares marginal effects for the linear and logistic different specifications below. Qualitatively, the results from the two models are similar. For women, the OLS results are fairly similar, while IV results are much larger in the logistic model, but still significant. Since many communities are at low initial employment, these areas have the potential to experience large increases in employment. For women, the lower bound of the AR confidence interval is the same in linear and logistic models, while the upper bound is larger in the logistic model; both of these confidence intervals strongly reject zero.

As in the linear model, not too much weight should be placed on the male results, since there is no reduced form relationship between male employment and gradient in the logistic model either.

OLS and IV estimates are insignificant in the non-linear specification.

Male and female effects are significantly different at the 17% level.

Sensitivity to Treatment, Sample and Instrument Definition

In Table A2, I present results for different definitions of treatment (year of exposure to an Eskom project and the fraction of households connected over the period), different sub-samples (including only areas without any electricity in 1996), and different instruments (using the interaction of gradient at the community and district level, and using gradient and the transect indicator). The IV results are generally consistent with a positive female employment response and no significant male employment response (male results not shown, available on request).

Table A1: Comparing Specifications of the Dependent Variable: Average Marginal Effects

	Women				Men			
	Linear model		Logistic model		Linear model		Logistic model	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.001	0.135**	-0.005	0.236***	-0.011*	0.042	-0.006	0.215
	(0.005)	(0.062)	(0.006)	(0.124)	(0.006)	(0.068)	(0.009)	(0.148)
Poverty rate	0.031***	0.028*	0.058***	0.052**	0.064***	0.063***	0.077***	0.072***
	(0.011)	(0.015)	(0.017)	(0.025)	(0.016)	(0.018)	(0.028)	(0.033)
Female-headed hh's	0.034	0.019	-0.005	-0.032	0.240***	0.234***	0.210***	0.185***
	(0.023)	(0.030)	(0.031)	(0.045)	(0.033)	(0.036)	(0.015)	(0.06)
Sex ratio (F/M)	0.024**	0.040***	0.035***	0.065***	0.001	0.007	0.000	0.027
	(0.009)	(0.013)	(0.014)	(0.023)	(0.012)	(0.016)	(0.022)	(0.031)
N	1992	1992	1992	1992	1992	1992	1992	1992
F-stat on Z		8.870		8.870		8.870		8.870
C.I.	[-0.01,0.01]	[0.01,0.26]	[-0.02,0.01]	[-0.01,0.48]	[-0.02,0]	[-0.09,0.17]	[-0.02,0.01]	[-0.08,0.5]
AR C.I.		[0.05,0.4]		[0.05,0.79]		[-0.05,0.25]		[-0.01,0.45]

(1) Marginal effects for the logistic specification are reported at the sample average employment rate in 1996, $\hat{\beta}^{*}\bar{y} * (1 - \bar{y})$. The average female employment rate in the first period is 0.069 and for men is 0.136.
(3) Standard confidence intervals for $\hat{\beta_T}$ appear underneath the F-statistic on gradient.
(4) Anderson-Rubin confidence intervals (robust to heteroscedasticity) appear in the final row.
(5) Robust standard errors, clustered at the sub-district level.

Table A2: Sensitivity Analysis: Treatment Definition, Sub-samples and Instrument Definitions

	Alternative definitions of treatment				Alternative sub-samples		Additional instruments	
	Year of exposure		Fraction connected		Areas w.o. electricity, 1996		Gradient and Transect	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Δ female employment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-0.001	0.045	0.006	0.247	0.001	0.077	0.001	0.112
	(0.001)	(0.027)*	(0.008)	(0.127)*	(0.008)	(0.044)*	(0.005)	(0.047)**
N	1,992	1,992	1,992	1,992	368	368	1,992	1,992
R^2	0.10		0.10		0.11		0.10	
CI	[-0.003-0.002]	[-0.01-0.1]	[-0.01-0.02]	[0-0.5]	[-0.015-0.017]	[-0.01-0.16]	[-0.008-0.01]	[0.02-0.204]
AR CI								

(1) Dependent variable is the change in female employment rate.
(2) In columns (1), (2), treatment is 0,1,2,3,4,5 for number of years ago community was treated, 0 if never treated. Mean of this variable is 0.60. In columns (3), (4), treatment is fraction of 1996 households connected between 1996 and 2001. Mean of this variable is 0.10. In columns (5) and (6), treatment is an indicator variable and the sample is restricted to areas that had no households with electricity in 1996.
In columns (7) and (8), the treatment variable is an indicator of treatment, and the instruments are gradient and the transect indicator.
(3) Robust standard errors, clustered at sub-district level.

Table 1: Average Community Covariates in 1996, by Treatment Status and by Gradient

Covariates	All (1)	Treatment (2)	Control (3)	Δ_{T-C} (4)	$\hat{\beta}_{Gradient}$ (5)
Poverty rate	0.61	0.59	0.61	-0.021	0.001
	(0.19)	(0.17)	(0.20)	(0.016)	(0.001)
Female-headed households	0.55	0.55	0.55	-0.004	0.000
	(0.13)	(0.12)	(0.13)	(0.011)	(0.000)
Adult sex ratio (f/m)	1.48	1.41	1.49	-0.082***	0.002**
	(0.29)	(0.25)	(0.30)	(0.023)	(0.001)
Share Indian/white adults	0.02	0.01	0.02	-0.016***	-0.001
	(0.21)	(0.03)	(0.23)	(0.007)	(0.001)
Kms to road	38.31	36.07	38.85	-2.789**	0.081
	(24.60)	(24.10)	(24.70)	(1.387)	(0.093)
Kms to town	39.01	36.80	39.55	-2.751***	-0.091
	(18.29)	(15.32)	(18.91)	(1.030)	(0.093)
Men with high school	0.06	0.08	0.06	0.016***	0.000**
	(0.05)	(0.05)	(0.05)	(0.003)	(0.000)
Women with high school	0.07	0.08	0.06	0.021***	0.000
	(0.05)	(0.06)	(0.05)	(0.004)	(0.000)
Household density/sq. km	20.67	30.76	18.21	12.549***	-0.465***
	(29.50)	(48.15)	(22.07)	(4.077)	(0.159)
Kms from grid	19.32	15.68	20.21	-4.53***	0.014
	(13.46)	(9.96)	(14.04)	(1.725)	(0.062)
Land Gradient - mean	22.26	20.33	22.73	-2.408*	
	(9.90)	(8.56)	(10.14)	(1.167)	
Gradient - std. dev.	10.84	10.32	10.97	-0.653	0.336***
	(3.92)	(3.82)	(3.93)	(0.490)	(0.011)
Gradient - range	52.51	50.26	53.06	-2.793	1.276***
	(15.75)	(15.34)	(15.81)	(2.037)	(0.047)
N communities	1992	391	1601		1992

(1) Cells in columns (1)-(3) are community level means (s.d.). Column (4) shows mean T-C and column (5) shows coefficients from regressions of each covariate on gradient, controlling for all other covariates and district fixed effects. Robust s.e. in columns (4) and (5) clustered at sub-district level.

(2) Sample is tribal KZN communities not treated before 1996; treatment is 1 if first Eskom project occurred between 1996 and 2001, else 0.

(3) Variable definitions (measured in 1996 and at community level): Poverty rate is fraction of households earning below ZAR6,000 per year. Sex ratio is number of African females (ages 15-59) over number of males (ages 15-59). Female-headed households expressed as fraction of all households. Number of Indian/white adults as a fraction of all adults. Distances to nearest road, town,sub-station are straight-line kilometer distances from community centroid to nearest object. Men/women with at least least high school as a share of all men/women. Household density is per square kilometer. Land gradient statistics created in ARCMAP at the sub-place level.

Table 2: Average Community-Level Outcomes in 1996 and 2001, by Treatment Status

	Year (1)	Mean (2)	Min (3)	Max (4)	Treatment (5)	Control (6)	Δ_{T-C} (7)
Female Employment Rate	1996	0.07	0.00	0.91	0.09	0.07	0.02***
		(0.08)			(0.07)	(0.08)	(0.00)
	2001	0.07	0.00	0.80	0.08	0.07	0.02***
		(0.07)			(0.07)	(0.08)	(0.00)
	Δ_t	0.00			-0.004	0.00	-0.004
		(0.00)			(0.00)	(0.00)	(0.00)
Male Employment Rate	1996	0.14	0.00	0.99	0.16	0.13	0.03***
		(0.12)			(0.11)	(0.12)	(0.01)
	2001	0.10	0.00	0.83	0.11	0.10	0.01**
		(0.10)			(0.09)	(0.10)	(0.01)
	Δ_t	-0.04			-0.05	-0.03	-0.02***
		(0.00)			(0.01)	(0.00)	(0.01)
Adult Females	1996	356.07	1	4,553	426.02	338.98	87.04***
		(347.84)			(379.53)	(337.58)	(19.53)
	2001	421.45	6	3,392	560.88	387.39	173.49***
		(401.50)			(494.99)	(367.37)	(22.32)
	Δ_t	65.38			134.86	48.41	86.45***
		(11.90)			(31.54)	(12.47)	(16.84)
Adult Males	1996	253.15	1	3,135	312.53	238.65	73.89***
		(261.74)			(291.11)	(252.03)	(14.68)
	2001	310.17	3	2,770	427.37	281.55	145.82***
		(314.09)			(403.18)	(281.00)	(17.42)
	Δ_t	57.02			114.84	42.90	71.94***
		(9.16)			(25.15)	(9.43)	(13.24)
N Communities		1,992			391	1,601	1,992

(1) Columns (1), (2), (5), (6) contain variable means (s.d.).

(2) Mean differences (s.e.) are shown for Δ_{T-C} (column) and Δ_t (row).

(3) Treatment is 1 first if Eskom project occurred between 1996 and 2001, else 0.

(4) All variables constructed for Africans only, adults are aged 15-59.

Table 3: Assignment to Treatment using Various Treatment Definitions - OLS

Dependent variable	Treatment Indicator [1/0] (1)	(2)	(3)	(4)	(5)	(6)	Year treated (7)	Fraction treated (8)
Gradient*10	-0.040	-0.040	-0.040	-0.040	-0.050	-0.050	-0.120	-0.020
	(0.020)*	(0.020)**	(0.010)***	(0.010)***	(0.02)***	(0.014)***	(0.050)**	(0.010)**
Kms to grid*10		-0.050	-0.020	-0.020	0.017	-0.009	-0.010	-0.020
		(0.020)**	(0.020)	(0.020)	(0.023)	(0.023)	(0.060)	(0.010)
Household density*10		0.020	0.010	0.010	0.012	0.013	0.050	0.000
		(0.000)***	(0.010)**	(0.010)**	(0.012)	(0.005)**	(0.020)***	(0.000)
Poverty rate		0.034	0.034	0.029	0.073	0.041	0.016	0.046
		(0.066)	(0.067)	(0.066)	(0.068)	(0.065)	(0.212)	(0.044)
Adult sex ratio (f/m)		0.350	0.134	0.124	0.084	0.078	-0.080	0.055
		(0.118)***	(0.104)	(0.104)	(0.108)	(0.103)	(0.365)	(0.075)
Female-headed hh's		-0.164	-0.117	-0.111	-0.087	-0.110	-0.344	-0.029
		(0.048)***	(0.038)***	(0.038)***	(0.052)*	(0.038)***	(0.416)*	(0.025)
Indian/white adults		-0.693	-0.576	-0.571	53.016	-0.651	-1.789	-0.320
		(0.256)***	(0.250)**	(0.230)**	(60.482)	(0.236)***	(0.168)***	(0.151)**
Kms to road*10		0.000	-0.010	-0.010	-0.004	-0.008	0.200	0.000
		(0.010)	(0.010)	(0.010)	(0.008)	(0.010)	(0.300)	(0.000)
Kms to town*10		0.020	0.010	0.010	-0.016	0.011	-0.400	0.000
		(0.010)	(0.010)	(0.010)	(0.016)	(0.015)	(0.500)	(0.010)
Men with high school		-0.041	0.090	0.087	-0.237	0.107	-0.265	0.273
		(0.451)	(0.399)	(0.387)	(0.469)	(0.386)	(1.253)	(0.221)
Women with high school		0.836	0.726	0.783	1.157	0.730	2.493	0.219
		(0.419)**	(0.395)*	(0.375)**	(0.469)**	(0.376)*	(1.280)*	(0.209)
Change in water access				0.016	0.109	0.014	-0.265	0.020
				(0.045)	(0.065)*	(0.044)	(0.159)*	(0.041)
Change in toilet access				0.178	0.444	0.169	0.931	0.083
				(0.086)**	(0.272)	(0.085)**	(0.382)**	(0.047)*
Transect Indicator						0.160		
						(0.069)**		
District FE	N	N	Y	Y	Y	Y	Y	Y
Sample	All	All	All	All	No elec	All	All	All
Mean of dependent var.	0.196	0.196	0.196	0.196	0.187	0.196	0.332	0.096
N	1,992	1,992	1,992	1,992	477	1,992	1,992	1,992
R^2	0.009	0.075	0.168	0.169	0.187	0.174	0.150	0.199
F-stat on instrument(s)	3.610	5.400	9.060	8.870	9.670	7.540	4.920	6.000
Prob>F:	0.060	0.020	0.000	0.000	0.000	0.000	0.030	0.010

(1) Dependent variables: columns (1)-(6) is indicator for treatment (1 if the area had a project in between 1996 and 2001); column (7) is number of years ago the project was completed (1,2,3,4 and 5 for up 5 years before 2001, 0 if no project); column (8) is fraction of 1996 households connected between 1996 and 2001.

(2) Land gradient in degrees, all distances in kilometers.

(3) In column (6), "Transect" is 1 if community lies on a straight line connecting any two sub-stations in the region, otherwise 0.

(4) Sample in column (5) is restricted to set of areas where no households had electricity in 1996.

(5) All controls measured in 1996, except change in access to water and flush toilet.

(6) Change in service access: change in fraction of households with access to water in the house or < 200m away, or with a flush toilet.

(7) Robust standard errors clustered at sub-district level. Ten district fixed-effects included in columns (3) to (7).

Table 4: Effects of Electricity Projects on Household Energy Sources and Other Services

	OLS		IV			
Dependent variable is Δ_t	No controls	Controls	No controls	Controls	Mean $\Delta\bar{y}_t$	N
	(1)	(2)	(3)	(4)	(5)	(6)
A: Lighting with Electricity	0.258***	0.233***	0.661***	0.713***	0.08	1,992
	(0.031)	(0.031)	(0.233)	(0.232)		
B: Cooking with Wood	-0.049***	-0.042***	-0.305	-0.283*	-0.03	1,992
	(0.012)	(0.012)	(0.197)	(0.148)		
C: Cooking with Electricity	0.069***	0.059***	0.281**	0.241**	0.03	1,992
	(0.008)	(0.008)	(0.125)	(0.099)		
D: Water nearby	-0.03	0.009	-0.449*	-0.287	0.01	1,992
	(0.028)	(0.023)	(0.251)	(0.231)		
E: Flush Toilet	0.003	0.01**	0.042	0.095	0.09	1,992
	(0.006)	(0.005)	(0.080)	(0.067)		

(1) Cells contain treatment coefficients (robust standard errors clustered at sub-district level) from regressions of dependent variable on treatment and all explanatory variables.
(2) Dependent variable is change in use of each type of service. Service use in each year is calculated as follows: share of households with electricity as main source of lighting; share of households using wood as main source of cooking fuel, share of households using electricity as main energy source for cooking; share of households with a water source in the house or within 200m; share of households with a flush toilet. Sample mean in column (5).
(3) Treatment is 1 if first Eskom project occurred between 1996 and 2001, otherwise 0.
(3) Excluded instrument is mean community gradient.
(4) Other controls: 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 D (E).
(5) Anderson-Rubin confidence intervals (not shown here) are positive and reject 0 for electric lighting and cooking; are negative and reject zero for wood cooking and cannot reject zero for water and toilet access.

Table 5: Effects of Electrification on Female Employment

Δ_t employment rate Females	OLS (1)	OLS (2)	OLS (3)	OLS (4)	IV (5)	IV (6)	IV (7)	IV (8)
Treatment	-0.004	0.000	0.002	0.001	0.045	0.091	0.136	0.135
	(0.005)	(0.005)	(0.005)	(0.005)	(0.055)	(0.062)	(0.064)**	(0.062)**
Kms to grid *10		0.004	0.004	0.004		0.001	0.001	0.001
		(0.003)	(0.003)	(0.003)		(0.004)*	(0.000)	(0.000)
Household density *10		0.000	0.000	0.000		-0.001	-0.002	-0.002
		(0.001)	(0.001)	(0.001)		(0.002)	(0.001)	(0.001)
Poverty rate		0.032	0.035	0.031		0.028	0.031	0.028
		(0.011)***	(0.011)***	(0.011)***		(0.013)**	(0.016)**	(0.015)*
Female-headed hh's		0.036	0.039	0.034		0.008	0.022	0.019
		(0.023)	(0.023)	(0.023)		(0.032)	(0.030)	(0.030)
Sex ratio (F/M)		0.020	0.020	0.024		0.036	0.038	0.040
		(0.010)**	(0.010)**	(0.009)**		(0.015)**	(0.013)***	(0.013)***
Indian/white adults		-0.495	-0.485	-0.482		-0.433	-0.413	-0.410
		(0.270)*	(0.269)*	(0.256)*		(0.271)	(0.263)	(0.255)
Kms to road*10		0.001	0.001	0.001		0.000	0.001	0.002
		(0.001)	(0.001)	(0.001)		(0.001)	(0.002)	(0.002)
Kms to town*10		-0.004	-0.003	-0.004		-0.006	-0.004	-0.005
		(0.002)**	(0.002)	(0.002)*		(0.003)**	(0.000)	(0.003)*
Men with h/s		0.150	0.161	0.159		0.146	0.139	0.137
		(0.104)	(0.105)	(0.092)*		(0.102)	(0.101)	(0.094)
Women with h/s		-0.180	-0.195	-0.153		-0.257	-0.290	-0.257
		(0.115)	(0.116)*	(0.100)		(0.120)**	(0.114)**	(0.108)**
Change in water access				0.028				0.026
				(0.007)***				(0.010)***
Change in toilet access				0.111				0.085
				(0.058)*				(0.058)
District FE?	N	N	Y	Y	N	N	Y	Y
N	1,992	1,992	1,992	1,992	1,992	1,992	1,992	1,992
R^2	0.000	0.067	0.075	0.100				
Standard 95% C.I.	[.0-.0]	[.0-.0]	[.0-.0]	[.0-.0]	[-.06-.2]	[-.03-.2]	[.01-.26]	[.01-.26]
AR 95% C.I.							[.05-.4]	[.05-.4]

(1) Outcome variable is change in employment rate of African females aged 15-59.
(2) Robust standard errors in parentheses, clustered at sub-district level.
(3) Treatment is 1 if community had the first Eskom project between 1996 and 2001, otherwise 0.
(4) Excluded instrument is land gradient.
(5) All controls (except treatment and change in access to other services) measured in 1996. Change in access to water and toilets measured between 1996 and 2001. 10 District fixed effects in (3),(4),(7),(8).
(6) Standard confidence intervals are provided for IV results as well as confidence intervals from the Anderson-Rubin test. The AR test is robust to weak instruments and was implemented to be robust to heteroscedasticity.

Table 6: Effects of Electrification on Male Employment

Δ_t employment rate Males	OLS (1)	OLS (2)	OLS (3)	OLS (4)	IV (5)	IV (6)	IV (7)	IV (8)
Treatment	-0.018	-0.016	-0.010	-0.011	-0.053	0.053	0.041	0.042
	(0.008)**	(0.006)**	(0.006)	(0.006)*	(0.080)	(0.080)	(0.068)	(0.068)
Kms to grid *10		0.009	0.006	0.006		0.012	0.007	0.008
		(0.004)**	(0.004)	(0.004)		(0.005)**	(0.005)	(0.005)
Household density *10		0.002	0.002	0.002		0.000	0.001	0.001
		(0.002)	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)
Poverty rate		0.066	0.068	0.064		0.063	0.066	0.063
		(0.018)***	(0.017)***	(0.016)***		(0.020)***	(0.018)***	(0.018)***
Female-headed hh's		0.235	0.243	0.240		0.213	0.237	0.234
		(0.031)***	(0.033)***	(0.033)***		(0.041)***	(0.036)***	(0.036)***
Sex ratio (F/M)		0.002	-0.001	0.001		0.015	0.005	0.007
		(0.011)	(0.011)	(0.012)		(0.019)	(0.015)	(0.016)
Indian/white adults		-0.077	-0.055	-0.052		-0.030	-0.027	-0.024
		(0.275)	(0.270)	(0.257)		(0.280)	(0.273)	(0.262)
Kms to road*10		0.002	0.000	0.000		0.001	0.000	0.001
		(0.001)	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)
Kms to town*10		-0.008	-0.003	-0.003		-0.009	-0.003	-0.004
		(0.002)***	(0.003)	(0.003)		(0.003)***	(0.003)	(0.003)
Men with h/s				-1.407		-1.399	-1.468	-1.491
				(1.258)		(1.258)	(1.241)	(1.343)
Women with h/s				1.405		0.615	0.712	0.997
				(1.314)		(1.551)	(1.451)	(1.506)
Change in water				0.278				0.273
access				(0.009)***				(0.009)***
Change in toilet access				0.826				0.723
				(0.747)				(0.737)
District FE?	N	N	Y	Y	N	N	Y	Y
N	1,992	1,992	1,992	1,992	1,992	1,992	1,992	1,992
R^2	0.005	0.152	0.169	0.179				
Standard 95% C.I.	[.0-.0]	[.0-.0]	[.0-.0]	[.0-.0]	[-.2-.1]	[-.1-.2]	[-.1-0.2]	[-.1-0.2]
AR 95% C.I.							[-.05-0.25]	[-.05-0.25]

(1) Outcome variable is change in employment rate of African males aged 15-59.

(2) Robust standard errors in parentheses, clustered at sub-district level.

(3) Treatment is 1 if community had the first Eskom project between 1996 and 2001, otherwise 0.

(4) Excluded instrument is land gradient.

(5) All controls (except treatment and change in access to other services) measured in 1996. Change in access to water and toilets measured between 1996 and 2001. 10 District fixed effects in (3),(4),(7),(8).

(6) Standard confidence intervals are provided for IV results as well as confidence intervals from the Anderson-Rubin test. The AR test is implemented to be robust to heteroscedasticity.

Table 7: Contribution of Measurement Error in Treatment to Female Employment Result

Δ_t female employment rate	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)
Treatment	0.001	0.135	0.010	0.126	0.012	0.136
	(0.005)	(0.062)**	(0.007)	(0.065)**	(0.009)	(0.092)
Sample is all control areas and treated areas	Treated areas		w/ > 10% Δ electricity		w/ >80% coverage	
N	1,992	1,992	1,619	1,619	1,420	1,420

(1) Dependent variable is change in employment rate of African females aged 15-59.
(2) Each coefficient (standard error) is from a separate regression. Robust standard errors in parentheses, clustered at sub-district level.
(3) Treatment is 1 if community had first project between 1996 and 2001, else 0. Columns (1) and (2) replicate the coefficient on treatment from Table 5.
Columns (3) and (4) restrict the sample to all control areas and treated areas with a 10% or larger change in fraction of households using electric lighting. Columns (5) and (6) restrict the sample to all control areas and treated areas where Eskom connected at least 80% of households between 1996 and 2001.
(4) All controls (except treatment 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, change in proportion of households with water close by, change in proportion of households with flush toilets and 10 district fixed-effects.

Table 8: Placebo Experiment and Reduced Form for Employers of Women - OLS

Dependent variable is change in	Placebo Experiment		Δ Female employers	
	Female Employment	Male Employment	Schools (N)	Indian/white Adults
	(1)	(2)	(3)	(4)
Gradient*10	-0.001	-0.001	0.002	0.001
	(0.001)	(0.001)	(0.013)	(0.001)
Sample	Treated<1996	Treated<1996	All	All
N	406	406	1,992	1,992
R^2	0.162	0.351	0.050	0.032

(1) Dependent variable in columns (1) and (2) is change in employment rate for adult Africans.
In column (3) it is the change in the number of schools in a community between 1996 and 2001.
In column (4) it is the change in the fraction of Indian/white adults.
(2) Robust standard errors in parentheses, clustered at sub-district level.
(3) In columns (1) and (2), sample is areas that had projects prior to 1996. Full sample included in columns (3) and (4)
are areas treated between 1996 and 2001 or never treated.
(4) All controls (except treatment and change in access to other services) measured in 1996: distance to the grid, household density, community poverty rate, adult sex ratio (F/M), fraction female-headed households, share of Indian/white adults, distance to road and town, share of adult men and women with at least high school, change in proportion of households with water close by, change in proportion of households with flush toilets. and 10 district fixed-effects. In the column (4), the level of Indian/white adults
in the community is excluded from regression.

Table 9: Testing for Spill-overs on Female Employment by Excluding Adjacent Control Areas

	Treatment coefficient		
	OLS	IV	N
Δ_t Female Employment	(1)	(2)	(3)
Full Sample	0.001	0.135	1,992
	(0.005)	(0.062)**	
Sample excludes:			
Control areas <1km from treated areas	-0.005	0.104	1,656
	(0.006)	(0.061)*	
Control areas <5km from treated areas	-0.005	0.114	1,374
	(0.008)	(0.097)	

(1) Dependent variable is change in employment rate of African women aged 15-59.

(2) Each coefficient (standard error) is from a separate regression. Robust standard errors in parentheses, clustered at sub-district level.

(3) Treatment is 1 if community had first Eskom project between 1996 and 2001, otherwise 0.

(4) Successive sample restrictions exclude control communities which fall partly/wholly inside inside an [X] kilometer radius of an area treated prior to 2001.

(5) All controls (except treatment 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, change in proportion of households with water close by, change in proportion of households with flush toilets and 10 district fixed-effects.

Table 10a: Contribution of Each Poverty Quintile to IV Estimate

Quintiles of Predicted Poverty Index	Sample Fraction in Quintile (1)	Variance of Gradient by Quintile (λ_q) (2)	$\Delta\hat{T}\|_q$ (3)	IV weight (w_q) (4)
Richest quintile	0.24	0.20	0.039 (0.045)	0.13
Second Richest Quintile	0.21	0.20	0.109** (0.040)	0.32
Middle Quintile	0.22	0.21	0.126** (0.038)	0.39
Second Poorest Quintile	0.18	0.21	0.056* (0.033)	0.15
Poorest Quintile	0.16	0.20	0.007 (0.037)	0.01

(1) Predicted poverty quintile is assigned as follows: for communities in the steepest half of the gradient distribution, I project the treatment indicator 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 sample.
(2) λ_q is the estimated conditional variance of the gradient dummy (1 =flat, 0 =steep)within each quintile: $\hat{E}(P[Z|x,q][1-P(Z|x,q)|q])$.
(3) For each quintile, $\Delta\hat{T}|q = \hat{E}(E((T|z=1,x,q) - \hat{E}(T|z=0,x,q)|q))$ is the estimated difference in treatment probability across top and bottom halves of gradient distribution, controlling for co-variates. The reported coefficient is on the interaction of this gradient dummy with each predicted quintile.
(4) w_q is the weight that each quintile contributes to the IV estimate. It is computed (as described in Kling (2001)) across the columns (1)-(3): $\omega_q = \frac{[(1)_q*(2)_q*(3)_q]}{\sum_q[(1)_q*(2)_q*(3)_q]}$

Table 10b: Household Energy Use by Poverty Quintile: Baseline and Over Time (1996 to 2001)

	Fuel Use in Home Production: Fraction using [X] in 1996			Δ_t Fuel Use in Home Production: Within-Quintile Difference by Gradient			Δ_t Employment by Gradient	
	Electric Lighting	Electric Cooking	Wood Cooking	Electric Lighting	Electric Cooking	Wood Cooking	Female	Male
Quintile of Poverty Index	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Richest Quintile	0.193	0.098	0.636	0.054*	0.027***	-0.029**	-0.003	-0.015
	(0.280)	(0.167)	(0.299)	(0.032)	(0.011)	(0.017)	(0.008)	(0.010)
Second Richest	0.086	0.038	0.762	0.055***	0.026***	-0.040***	0.013	-0.001
	(0.189)	(0.085)	(0.249)	(0.023)	(0.010)	(0.014)	(0.006)*	(0.008)
Middle Quintile	0.061	0.030	0.815	0.055***	0.022***	-0.024***	0.010	0.000
	(0.162)	(0.096)	(0.208)	(0.026)	(0.008)	(0.011)	(0.005)*	(0.007)
Second Poorest	0.049	0.023	0.851	0.009	0.007	-0.013	0.009	0.006
	(0.149)	(0.078)	(0.188)	(0.017)	(0.008)	(0.013)	(0.005)	(0.007)
Poorest Quintile	0.013	0.007	0.900	0.007	0.004	0.000	0.006	0.002
	(0.060)	(0.023)	(0.151)	(0.017)	(0.006)	(0.018)	(0.006)	(0.008)

(1) Columns (1)-(3) present quintile means, columns (4)-(8) present coefficients from regression of outcomes on all controls, and interactions of gradient dummy and predicted poverty quintile.
The gradient dummy is 1 for areas in the flattest half of the gradient distribution, 0 if in the steepest half.
(2) Predicted poverty quintile is assigned as follows: for communities in the steepest half of the gradient distribution, I project the treatment indicator 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 sample.
(3) All controls (except treatment and change in access to other services) included and 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 road and town, share of adult men and women with at least high school, change in proportion of households with water close by, change in proportion of households with flush toilets and 10 district fixed-effects.
Columns (7) and (8) are akin to reduced-form regressions of outcome variables on values of the instrument.

Table 11: Effects of Electrification on Female Employment: Age-Specific Treatment Effects

	OLS	IV
Dependent variable	(1)	(2)
Δ female employment, all ages	0.001	0.135**
	(0.005)	(0.062)
Δ_t female employment, ages 15-19	0.000	0.005
	(0.000)	(0.007)
Δ_t female employment, ages 20-24	0.000	0.009
	(0.001)	(0.014)
Δ_t female employment, ages 25-29	-0.001	0.02
	(0.001)	(0.014)
Δ_t female employment, ages 30-34	0.000	0.039***
	(0.001)	(0.018)
Δ_t female employment, ages 35-39	0.001	0.026**
	(0.001)	(0.014)
Δ_t female employment, ages 40-44	0.001	0.013
	(0.001)	(0.012)
Δ_t female employment, ages 45-49	0.001	0.019***
	(0.001)	(0.009)
Δ_t female employment, ages 50-54	-0.001	0.003
	(0.001)	(0.006)
Δ_t female employment, ages 55-59	0.000	0.004
	(0.001)	(0.005)

(1) Robust standard errors in parentheses, clustered at sub-district level.

(2) Dependent variable is change in fraction of employed African females in age group [X] over all females.

(3) Treatment is 1 if community had first project between 1996 and 2001, else 0.

(4) Other controls in 1996: distance to the grid, household density, community poverty rate, fraction female-headed households, adult sex ratio (F/M), 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 proportion of households with water close by, change in proportion of households with flush toilets and 10 district fixed-effects.

(5) Excluded instrument is average land gradient.

(6) N=1,992 in each regression.

Table 12: Effects of Electrification on Female Employment Related to Children

	OLS	IV
Δ_t Female Employment	(1)	(2)
Treatment	0.035	0.623
	(0.022)	(0.271)**
Treatment*Children/Household ratio	-0.040	-0.607
	(0.024)*	(0.346)*
Children/Household ratio	-0.002	0.076
	(0.017)	(0.040)*
Mean of Children/Household ratio	0.874	0.874
Interaction at mean Children/Household ratio	-0.035	-0.531
	(0.004)	(0.303)*
Total effect at mean Children/Household ratio	0.000	0.092
	(0.004)	(0.069)
N	1,992	1,992
R^2	0.10	

(1) Each cell is a coefficient (s.e.) from a different regression. Dependent variable is change in proportion of employed African females.
(2) Robust standard errors in parentheses, clustered at sub-district level.
(3) Treatment is 1 if community had first Eskom project between 1996 and 2001, else 0.
(3) Children/Household is the ratio of children ages 5 to 14 in 2001 to the number of households in 2001. Most of these children will not have been enrolled in school in between 1996 and 2001. Mean (std. dev.) of this variable is 0.874 (.20).
(4) Excluded instruments are average land gradient and interaction of gradient with the ratio of children aged 5 to 14 to the number of households measured in 2001.
(5) Other controls: distance to the grid, household poverty rate, adult sex ratio, fraction of female-headed households, proportion of Indian/White adults, distance to road and town, proportion of men and women with high school, in proportion of households with access to water close by and to flush toilets, and 10 district fixed-effects.

Table 13: Effects of Electrification on Population Growth and Employment of Incumbents

	Δ_t Log Population			Δ_t Female Employment: Excluding In-Migrants		Δ_t Male Employment: Excluding In-Migrants	
	OLS (1)	IV (2)	OLS (3)	OLS (4)	IV (5)	OLS (6)	IV (7)
Treatment	0.258	4.012		0.001	0.135	-0.010	0.084
	(0.062)***	(1.495)***		(0.005)	(0.071)*	(0.006)*	(0.069)
Kms to grid*10	-0.030	0.047	-0.003	0.002	0.005	0.003	0.005
	(0.030)	(0.091)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Household density*10	-0.059	-0.113	-0.006	0.000	-0.002	0.002	0.001
	(0.030)**	(0.040)***	(0.003)**	(0.001)	(0.001)	(0.001)*	(0.002)
Poverty rate	-0.165	-0.255	-0.140	0.024	0.021	0.051	0.049
	(0.132)	(0.269)	(0.132)	(0.009)***	(0.014)	(0.014)***	(0.016)***
Female-headed hh's	0.041	-0.371	0.126	0.047	0.032	0.248	0.238
	(0.236)	(0.462)	(0.234)	(0.019)**	(0.026)	(0.031)***	(0.035)***
Adult sex ratio (F/M)	-0.103	0.358	-0.086	-0.003	0.013	-0.018	-0.006
	(0.087)	(0.218)	(0.086)	(0.008)	(0.012)	(0.012)	(0.015)
Indian/white adults	-4.248	-2.226	-4.515	0.071	0.143	0.544	0.595
	(0.820)***	(1.661)	(0.865)***	(0.209)	(0.215)	(0.169)***	(0.178)***
Kms to road*10	0.003	0.020	0.000	0.001	0.002	0.001	0.002
	(0.014)	(0.039)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Kms to town*10	-0.009	-0.037	0.000	-0.003	-0.004	-0.003	-0.004
	(0.022)	(0.066)	(0.002)	(0.002)*	(0.003)	(0.003)	(0.003)
Men with h/s	0.766	0.163	0.512	0.069	0.047	-0.105	-0.121
	(0.833)	(1.781)	(0.834)	(0.066)	(0.087)	(0.090)	(0.105)
Women with h/s	0.304	-2.607	0.533	-0.097	-0.201	0.054	-0.019
	(0.649)	(1.999)	(0.629)	(0.076)	(0.104)*	(0.099)	(0.122)
Change in water access	0.156	0.116	0.179	0.009	0.008	0.011	0.010
	(0.065)**	(0.186)	(0.067)***	(0.007)	(0.009)	(0.009)	(0.010)
Change in toilet access	-0.178	-0.908	-0.193	0.010	-0.016	0.004	-0.015
	(0.251)	(0.534)*	(0.257)	(0.031)	(0.035)	(0.065)	(0.067)
Gradient*10			-0.016				
			(0.003)***				
N	1992	1992	1992	1992	1992	1992	1992
R^2	0.072			0.025		0.138	
Standard C.I.	[0.14-0.38]	[1.08-6.94]			[0.05-0.3]		[0-0.3]
AR confidence interval		[2.15-6.9]			[0.05-0.35]		[0-0.3]

(1) Dependent variable in columns (1)-(3) is change in log African population, in columns (6)-(9) is change in employment rate of African females or males where the numerator has been adjusted downwards for the count of all African adults who report they have moved to the area in the 5 years before the relevant Census year.

(2) Robust standard errors in parentheses, clustered at sub-district level.

(3) Treatment is 1 if community had first project between 1996 and 2001, else 0.

(4) Other controls: distance to the grid, household density, community poverty rate, adult sex ratio (F/M), fraction female-headed households, share of Indian/white adults, distance to nearest road and town, share of adult men and women with at least high school, change in proportion of households with water close by,change in proportion of households with flush toilets and 10 district fixed-effects.

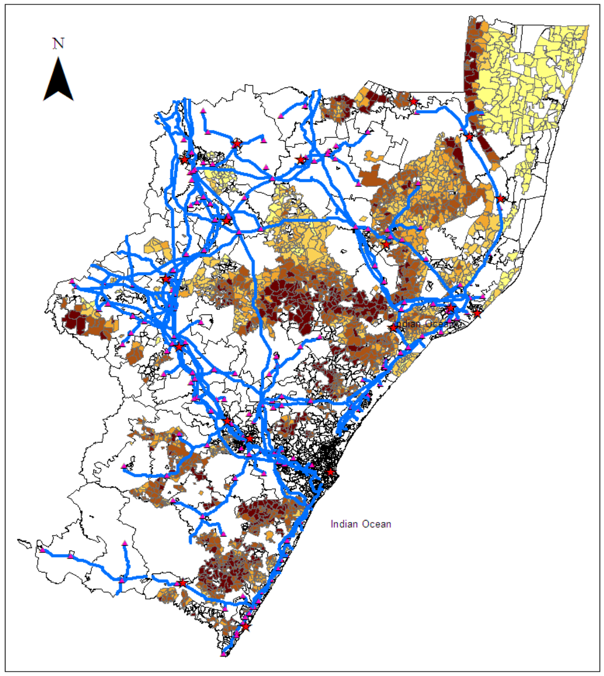
(5) Mean number of adult female (male) in-migrants in 1996 is 15.02 (10.07), in 2001 is 11.56 (8.60).

Table 14: Social returns to rural electrification

(1)	Total number of connections 1996-2001	470,000
(2)	Average cost per connection	2,533
(3)	Total fixed costs (1)*(2)	1,190,510,000
(4)	Estimated increase in employment (%-point)	0.135
(5)	Total female population in treated areas, 1996	166,574
(6)	Total number newly employed across all treated areas (4)*(5)	22,487
(7)	Median unskilled earnings in rural areas, annual (ZAR300/month)	3,600
(8)	Fraction of time worked	0.25
(9)	Total earnings per job, annual (8)*(7)	900
(10)	Total earnings over all treated areas and all jobs, annual (9)*(5)	20,238,741
(11)	Total value of time saved for all women (9)*(5)	149,916,600
(12)	Total (undiscounted) stream of benefits from employment	
	over five years	101,193,705
	over ten years	202,387,410
	over fifteen years	303,581,115
(13)	Total (undiscounted) stream of benefits from total time saved	
	over five years	749,583,000
	over ten years	1,499,166,000
	over fifteen years	2,248,749,000

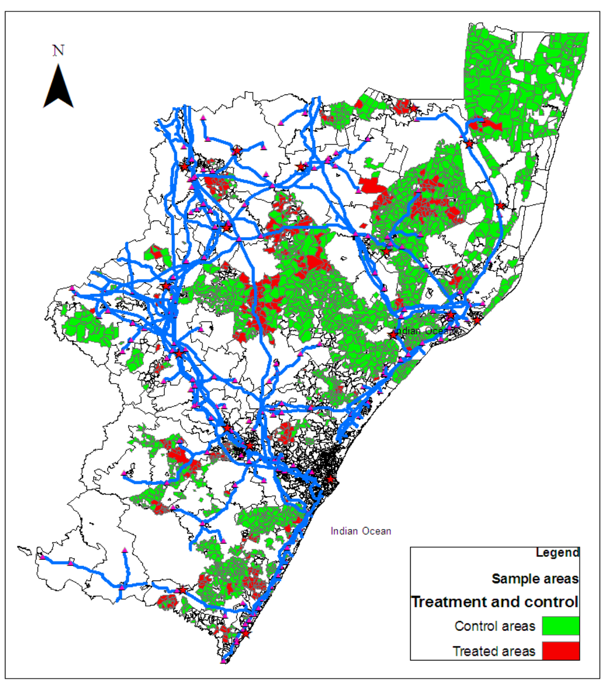
All monetary amounts measured in South African Rand (ZAR)
In (7), median earnings are for rural African women in elementary occupations, taken from the 2001 South African Labor Force Survey (February).
In (8), fraction of time worked is 0.25 under the assumption that 2 hours of work/day are freed up from collection of wood.

Figure 1: Gradient in Sample Areas: KwaZulu-Natal, South Africa



All shaded areas included in sample. Steeper areas are shaded dark, flatter areas are shaded light (see electronic version for color). Lines depict electricity grid lines in 1996, triangles are electricity substations in 1996 and stars represent towns. N=1,992.

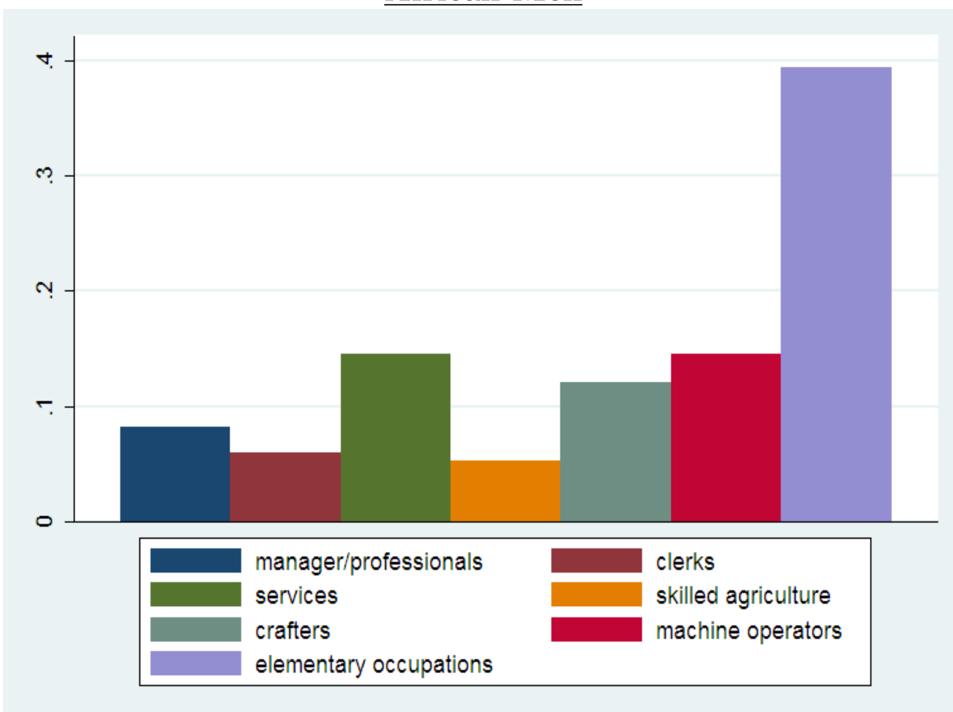
Figure 2: Treated and Control Areas: KwaZulu-Natal, South Africa



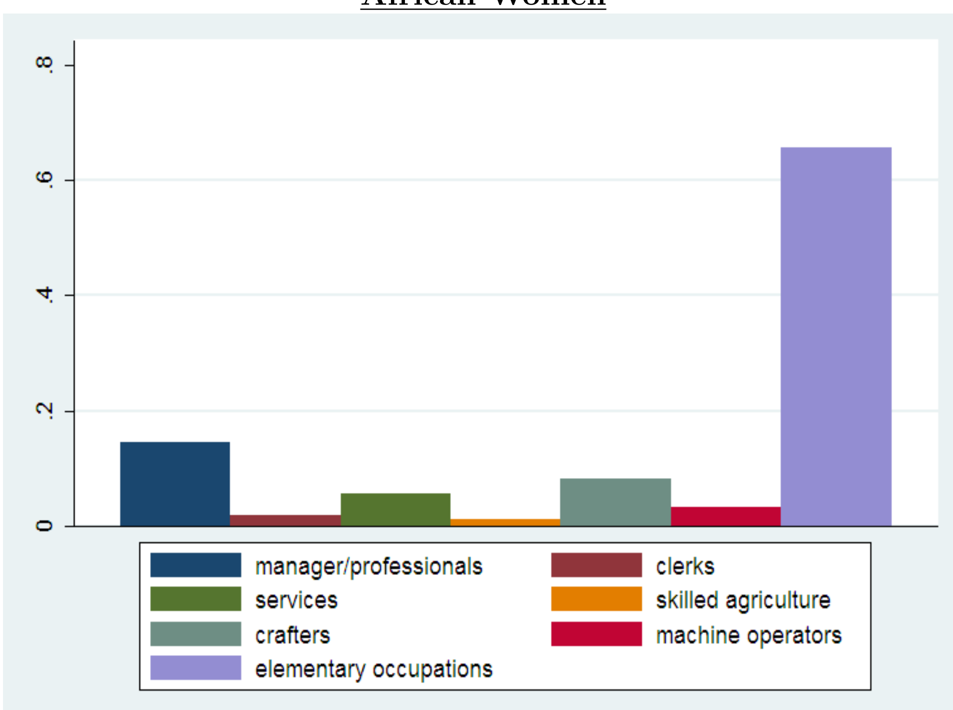
Shaded areas are in the sample: dark shaded areas are treated with an Eskom project between 1996 and 2001, lighter shaded areas are treated after 2001 or not at all (see electronic version for color). Lines represent electricity grid lines in 1996, triangles are electricity substations in 1996 and stars represent towns. N=1,992, N_T = 391, N_C = 1,601.

Figure 3: Distribution of Occupation Groups for the Employed, OHS 1996

African Men

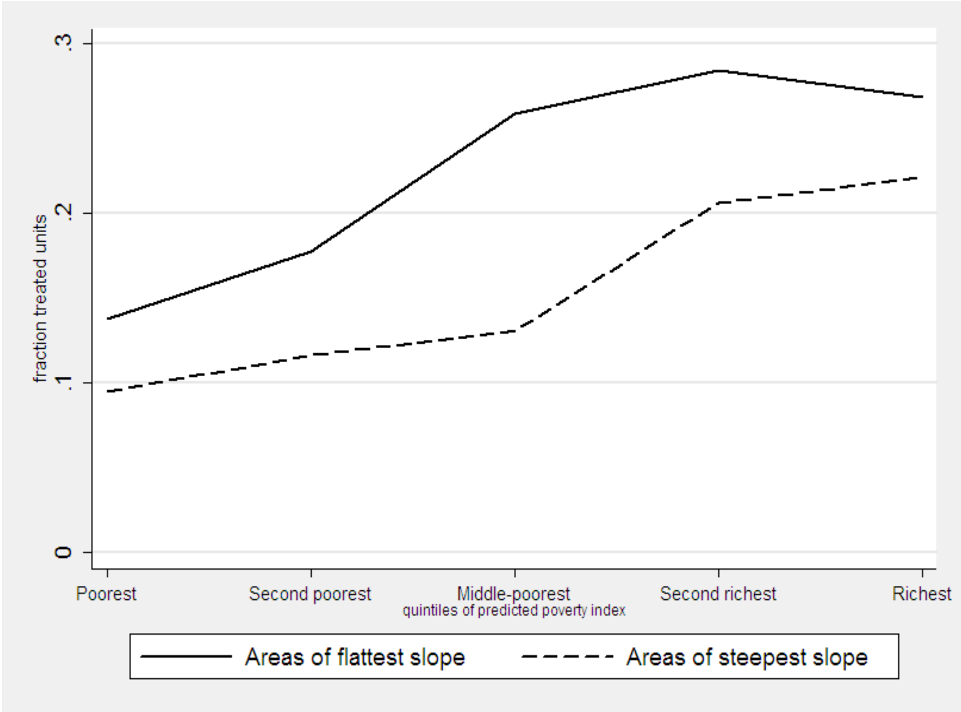


African Women



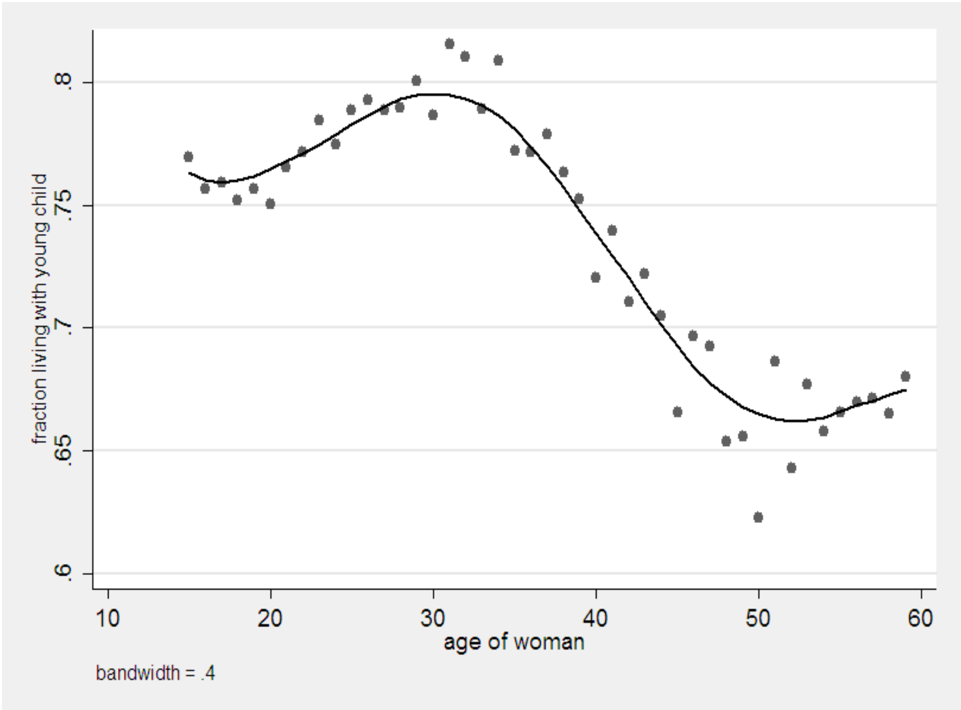
Groups are defined as number of men/women employed in each occupation over all employed men/women. Sample includes employed African men and women in rural KZN. (Weighted) data are from South African October Household Survey 1996. N_{men} =282, N_{women} =270.

Figure 4: Effect of Gradient on Treatment Probability, by Poverty Quintile



Lines show fraction of each predicted poverty quintile that is treated, by top and bottom half of the gradient distribution. See notes for Table 10a 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 treatment probability the most.

Figure 5: 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 Census 10% micro data and include African women aged 15-59 living in rural KwaZulu-Natal. N=116,381 collapsed to 44 age-specific data points.