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

Taryn Dinkelman*
University of Michigan

August 3, 2008

Abstract

This paper establishes that new access to public infrastructure affects both home production technologies and market employment in a developing country. I identify these effects by exploiting variation in electricity project placement and timing from South Africa's mass roll-out of rural household electricity. I estimate district fixed-effects models of employment growth and fuel-use and instrument for project placement using land gradient. Within five years, treated areas substitute sharply towards electricity in home production and IV employment results are asymmetric by gender: female employment increases by a significant 13.5 percentage points, while there are no significant male effects.

*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, Raymond Fisman, Ben Keyes, David Lam, Jim Levinsohn, Claudia Martinez-A, Justin McCrary, Nancy Qian, 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 following universities, organizations and conferences for their valuable input: Brown, California-Los Angeles, Dartmouth, Duke, the Harris School, the Harvard School of Public Health, the Kennedy School, KwaZulu-Natal, Michigan, Michigan State, the Microeconomics of Growth Network, Northwestern, Penn State, Princeton, Resources for the Future, Washington-St Louis and Western Ontario. Data collection was funded through a National Science Foundation Doctoral Dissertation Grant and an Eva Mueller New Directions in Demography research award. I also acknowledge support from the Fogarty International Center of the U.S. National Institutes of Health (D43TW000657) and the Population Studies Center at the University of Michigan. All errors are my own.

1 Introduction

Electricity is pervasive in all industrialized countries and largely absent in 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)). Perhaps because of its symbolic association with modernization, electricity is often part of electoral promises made to constituencies across Africa. While such political promises have not always been honored, access to electricity and other modern energy sources is likely to expand in many poor countries over the next several decades.[1] Some of this planned expansion targets industrial needs while other "Bottom of the Pyramid" initiatives focus on making cheap modern energy available to 250 million Africans by 2030.[2] Regardless of the type of infrastructure expansion, the effects of this new access to modern energy on economic growth, health, human capital accumulation and labor market outcomes are largely unknown.

In this paper, I use the experience of post-apartheid electrification roll-out in rural South Africa to analyze the impact of new access to modern energy sources on one microeconomic outcome of considerable interest: the ability of the poor to use their labor resources for market production. If households substitute towards electricity once infrastructure becomes available, potentially large amounts of time spent in fuel-wood collection and in food preparation with traditional fuels could be saved for other economic activities (Charmes (2005), Saghir (2005)).[3] In South Africa in the mid-1990s, over 80% of households collected wood for household needs; over three-quarters of those collecting wood were women; and on average, these individuals spent the equivalent of two working days per week in fuel-wood collection (Budlender et al, 2002).[4] Using data from this environment in

[1]For example, Ellen Johnson-Sirleaf promised and delivered a return of electricity to Monrovia during her Liberian presidential campaign while promises of widespread electrification by Olusegun Obansanjo, the former president of Nigeria, were largely unfulfilled during his 8 years in office. See the New York Times, July 26 2007 "One new light in Libera" and All Africa News, 2 December 2007, "Nigeria: Local official brings power to the people". For details on planned electrification investments, see EnergyNet Limited (2004) and Sustainable Development Network (2007) as well as "Middle East, North Africa to see large energy investments" by Deutsche Presse-Agentur, accessed June 30, 2006 at news.monstersandcritics.com/energywatch/news/article_1172084.php.

[2]World Bank commitments to energy infrastructure in sub-Saharan Africa rose from $447 million in 2001 to $790 million in 2007. See also the World Bank's Lighting Africa initiative www.lightingafrica.org.

[3]See also "The Energy Challenge for Achieving the Millenium Development Goals" at http://esa.un.org/un-energy/pdf/UN-ENRG/20paper.pdf accessed 8 August 2006.

[4]Food preparation in rural areas consumes on average 3 hours per day. Data from South African's 1997 October Household Survey.

rural KwaZulu-Natal (KZN), this paper provides some of the first evidence in support of claims that access to modern energy sources can release time into the labor market. My results also highlight some of the factors which are important for facilitating the entry of women into the labor market– a central aspect of the development experience of most economies.

The roll-out of grid infrastructure in South Africa provides a good opportunity to evaluate the effects of household electrification on market employment and household fuel use. In 1993, one year 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 on industrial connections (Gaunt, 2003).

Even with this focus on households, evaluating the effects of electrification is tricky. A large literature on the relationship between infrastructure and economic growth acknowledges that infrastructure spending may be targeted at growth centers, or towards areas that are lagging behind, but politically important. This means that estimates of the effects of infrastructure obtained from a comparison of treated and non-treated areas may be biased upwards or downwards.[5] Furthermore, teasing out the effects of infrastructure on economic variables is difficult without adequate controls for trends in economic conditions. 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 employment 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. I match this administrative and spatial data with two waves of Census data covering rural KwaZulu-Natal (KZN), a province containing one-fifth of the population of South Africa.[6] Using this two-wave panel of communities, I compare changes in

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

[6] Electricity projects are likely to have larger impacts in rural areas where reliance on wood is high in the baseline period.

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 at the community level, I use community land gradient to instrument for project placement between 1996 and 2001. Gradient directly affects the average cost per household 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.

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 over the five-year period. Even greater shocks to home production technology occur in communities in which treatment probability is most sensitive to gradient: instrumental variables estimates are three to seven times larger than ordinary least squares (OLS) estimates.

The simple comparison of treated and non-treated areas strongly suggest that projects are targeted to 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 in the OLS results.[7] 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 stretching from 5 to 40 percentage points, indicating a non-trivial positive impact on women's employment. Given the baseline female population of 166,574 in treated areas, this translates into an increase of 22,487 newly employed women. Effects for men are not significantly different from zero in the IV specifications. The female results are notable, since over the same period national unemployment rates are rising.

Gender asymmetries in employment responses are plausible in this context: home production is a domain in which women typically perform most activities and so are more likely to be affected by the rural electrification. In addition, as Gronau (2006) shows for the case of Russia, women have a

Administrative data on these projects is also only available for KZN province.

[7] Reductions in employment are partly the result of industrial restructuring in the 1990s, particularly in male-dominated commercial agriculture and mining sectors.

much higher elasticity of substitution between work in the market and work at home than men.[8] When less time is required in home production due to technological improvements in the form of electricity, women are more likely than men to substitute this time towards market work.

Several pieces of evidence together suggest that the female employment results are unlikely to be driven by a net increase in labor demand in response to the new household electrification. The capacity of electricity that was supplied to households was too small to be of use to large or 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; and there is no evidence of cross-community employment spillovers, which would be one outcome of increasing labor demand at the firm level. In addition, a placebo test for areas treated prior to 1996 provides no evidence that gradient is directly associated with employment prospects over time. 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 production in rural areas, which led to a corresponding increase in female labor supply.

Measured employment effects are not uniform across all women; rather, the channels through which household electrification affects female employment are connected to other household constraints.[9] I show that treatment probability is significantly higher in flatter, middle-poor communities. Households in these middle-poor communities are just the ones more likely to respond to the new option to use electricity. They contain households that have largely not started to invest in alternative modern technologies like gas or kerosene (as richer areas have), and compared to the poorest communities, they contain more households able to switch towards the new service when it

[8]Gronau (2006) uses longitudinal time-use data from Russia that spans the transition from a controlled to a market economy to show that men and women react differently to employment fluctuations during a recession. When time spent in market work falls during a recession, women split this extra time between home work and leisure while men substitute most of the time towards leisure. This implies that women have a higher elasticity of substitution between market and home work than men do.

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

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 the literature on the effects of physical infrastructure in developing countries in several ways. First, it places a new emphasis on employment effects in an area where the current focus is largely on poverty, health and education outcomes.[10] Studies which do not measure employment effects could be missing important economic impacts, particularly when the infrastructure has a home-production bias, as in the case of water and sanitation services.[11] Second, the results highlight short-run heterogeneity in the effects of infrastructure across types of communities and types of women. This information is useful for predicting the short-run 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 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, is likely to be feasible in other networked infrastructure settings.

Finally, the results of this paper add to a vast and newly invigorated literature on female labor force participation that connects aspects of economic development, labor economics and economic history.[12] Recent simulations by Greenwood et al (2005) show that price reductions in household appliances contributed over 50 percent of the rise in female labor force participation in the US between 1900 and 1980, and Coen-Pirani, Leon and Lugauer (2008) provide microeconomic evidence that the expansion of dryers and freezers in particular explains about half of the increase in married

[10]See Cutler and Miller (2005) for the effects of clean water technology in the US; Loshkin and Yemtsov (2005) for effects of infrastructure upgrades in Georgia, Russia; Duflo and Pande (2006) on the effects of Indian dam construction; Cattaneo et al (2007) for the effects of cement floors in Mexico. Existing evidence on how infrastructure affects work and wages is limited, but suggests that effects could be large. Banerjee et al (2007) find that Chinese wages are 37 percent higher in areas transected by railroads while Akee (2006) estimates the effects of road construction on wage employment at 27 percent and on agricultural employment at 38 percent.

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

[12]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)).

women's labor force participation in the US between 1960 and 1970.[13] The response of female employment to household electrification in South Africa represents complementary 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 we might expect responses to electrification to 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 and section 9 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 demand for meals.[14]

The arrival of infrastructure for domestic electricity may be characterized as a positive shock to the productivity of time spent in both activities.[15] Labor-saving electrification increases the effective amount of labor available for producing commodities: it reduces the need to fetch wood, speeds up

[13]Although see Bailey and Collins (2006) for some counter-evidence using variation in electrification use rates between 1930 and 1960.

[14]There is evidence to show that as people get wealthier, expenditure on non-food items increases relatively more than expenditures on food items and that the demand for food and calories is income inelastic (Houthakker (1957), Deaton and Subramanian (1996)).

[15]This is similar to Michael (1982) who models the impact that human capital has on non-market productivity.

cooking time and allows households to shift activities from daytime into nighttime. Since the production of meals is more time intensive than the production of clothing, this shock increases a household's productivity in making meals more than it increases their clothing productivity.

After electricity arrives, substitution and endowment effects operate in different directions on the household's decision about how to allocation time across market work 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 endowment, and the demand for income-elastic clothing rises relatively more under the assumption of non-homothetic preferences. Since clothing is market-goods intensive, the household may decide to supply more labor to the market in order to consume more of this commodity. It is therefore possible for the household to produce more meals and more clothing with less total time in home production after the technology shock. This also means that the net effect of the substitution and endowment effects on labor supply to the market is ambiguous.

Differences in responses to this technology shock across households will be linked to heterogenous preferences for time- and market-intensive commodities (which we can't measure) and to 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 activities intensive in home-time (e.g. child-care) will be more able to respond to a technology shock by shifting labor into the market, but also experience smaller positive endowment shocks when electricity arrives. 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 amount of time spent in home production in rural South Africa? Fuel-wood collection is one of the most time-consuming home production activities undertaken by women (as evidenced in time-use data) and large effects are certainly 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. High-skilled women are also possibly directly affected by this new infrastructure: Census micro data indicate that 30 percent of women with at least a high school certificate live in households where the main cooking fuel is wood, and half of these women live in households using candles for lighting. Household survey data collected in the same area as my study sample indicates that newly connected households report large uptake of appliances that improve the efficiency of home production (electric tea kettles, refrigerators and electric lighting),[16] and there is some evidence that access to electricity directly reduces the physical burden of home work.[17]

Of course, new access to electricity may directly affect market production and the demand for labor. Access to electricity may change the nature of enterprizes that can operate in these rural areas. 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 that do measure wages contain enough spatial information to be useful. Since new access was limited to a low level of service, medium-size businesses would not have benefited from this roll-out (South African Department of Minerals and Energy, 2004). In addition, if electrification predominantly affected labor demand, we should expect different employment effects when comparing treated areas to non-adjacent control areas that are at less risk of experiencing labor demand spill-overs and growth in the number of major employers in response to treatment. Neither of these effects are evident in my data. With this context as background and these pieces of evidence, I argue that the employment results are consistent with household electrification having a net labor supply impact.

3 Electrification roll-out in South Africa

3.1 Details of the program

Under apartheid, many African households were denied access to basic services. This was true particularly in homeland areas, which were pockets of land designated for African settlement and

[16]Author's own calculations from the KwaZulu-Natal Income Dynamics Study, a small panel study of households from 1993 and 1998.

[17]See Wittenberg (2007) for the relationship between household electricity and body mass index in South Africa.

which functioned as labor reserves for the white economy.[18] In 1994, all homelands were legally reintegrated into South Africa (Christopher, 2001) and the South African government was again responsible for basic service provision in these areas. By 1990, most economic units and white settlements were electrified and the political concerns of the 1980s had led to extensive electrification of commercial white farms in rural areas (Gaunt, 2003). In contrast, high-voltage lines carrying power from generation plants to white farms and towns often transected homelands that were themselves without power. At the time of the first democratic elections in 1994, over two-thirds of African households had no access to electricity. The National Electrification Programme (NEP) made the elimination of this backlog a development priority.[19]

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).[20] Since Eskom was a parastatal and a monopolist in electricity generation during this period, internal support for the roll-out was partly strategic. Eskom was interested in signalling to the government that introducing competition to the industry was not necessary to provide full access to previously disadvantaged communities.[21] As a result, Eskom met their connections targets in most years. Between 1993 and 2003, over 10 billion Rands (about USD1.4 billion) was spent on household electrification and over 470,000 households (28% of all KZN households) were electrified in KZN province alone.

Once an area had been targeted for electrification, each household was fitted with the basic connection package, consisting of an electric circuit board, a pre-payment meter, three plug points and one light bulb. Households received a default supply of 2.5 amperes or could upgrade to a 20 ampere supply for a fee of about ZAR40 (USD6.00).[22] The majority of Eskom's 3 million rural

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

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

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

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

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

customers upgraded to the 20 ampere supply (Gaunt, 2003). Default supply was sufficient for television, radio, two lights and several small kitchen appliances; an upgraded supply could power more appliances simultaneously including a refrigerator and a small water heater (South African Department of Minerals and Energy, 2004).

This subsidized roll-out represented a change in the option to use electricity. Industry experts agreed that "Electric lighting was synonymous with the roll-out"[23], and that the NEP did reach poor households. However, households still had to pay for the use of this service by purchasing electricity credits loaded on to pre-paid cards. In 1999, household electricity cost $0.039 per kilowatt hour (kWh).[24] Estimates of load demand from Eskom reports suggest that most rural households would have used between 35 and 60kWh per month which translates into energy expenses of between $1.37 and $2.34 per month (Gaunt, 2003) or 1.8 percent of median monthly household income in rural KZN in 1995. Because of this non-zero marginal cost, the poorest households may have been least responsive to this new technology in the short-run.

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 an important 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.[25] As there is no way to measure these political factors in my data or how they may have affected allocation of projects, I

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

[24]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)

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

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.[26] These cost factors are central to the identification strategy in this paper. The bulk of electrification cost is in laying distribution lines out from electricity sub-stations to households. Three main factors reduce the cost of these distribution lines: proximity to existing sub-stations and power lines; higher density settlement; and land gradient and terrain. The less of an incline the land has, the fewer hills and valleys and the softer the soil, the cheaper it is to lay power lines and erect transmission poles (Eskom, 1996; West et al, 1997).

In my data set, I have assembled measures of each of these cost factors. Distance from the grid and household density are important control variables 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 that addresses the multiple biases arising from selection on unobservables and confounding trends. I discuss further motivation for this particular instrumental variable 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 Data Appendix 1. For some parts of the analysis, I also use the 10% micro Census data for 1996 and 2001. These Census data are reported at the household level, but the public-release data do not contain enough geography to identify areas treated with

[26]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."

electricity projects which is necessary for the main analysis.

The primary unit of analysis in this paper is the community. A community is small, roughly equivalent in size to a US Census tract. The median number of households per community is 146 in 1996, and 187 in 2001. Ninety-five percent of communities have no more than 750 households. Key variables captured in the aggregate community-level Census data include the fraction of households with electricity in each year, the fraction of African adults by age group and labor market state and the fraction of households living below a poverty line. I also create a household density measure using the land area of the community. Results are not weighted, as all variables are derived from the full population Census.[27]

Questions about employment are similar across Census waves. 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 that are treated by virtue of having a 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 some information about 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 the second of the electrification 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 I can identify when a community gets new access to infrastructure rather than rely on time variation in the use of electricity at the household level, which is likely to be strongly correlated with changes in 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

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

project post-2001. Areas with pre-1996 projects are excluded from the main analysis; there are 406 of these out of the total 2,398 tribal areas in the sample (17 percent). I use these communities to conduct a placebo test in support of the exclusion restriction. Two other treatment measures are constructed for sensitivity tests: time since treatment and treatment exposure 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. Gradient is the third main factor affecting the average cost of a household connection.

Rural ex-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 the effects of rural electrification since the appropriate technology for supplying smaller power loads to rural areas had been developed by the mid-1990s. Since rural households are more likely to rely on time-consuming traditional fuels than urban households, we should see larger effects of electrification in this group. Census micro data from 1996 confirms that 63.4 percent of rural households used wood for cooking, whereas only 2.7 percent of urban African households did so. Expansion of service in rural areas is also more straightforward to analyze than urban areas, since Eskom is the sole distributor of power for rural areas. In urban areas, distribution rights are shared with local municipalities, which complicates decision-making about roll-out. In addition, there are potentially fewer economic confounders in rural areas than in urban areas in the first years after the end of apartheid. Access to other development services is one source of confounding in rural settings and I control for this in the empirical work. Migration is another challenge to identification and I address this issue in detail in section 8.

Boundaries of the sample of 1,992 rural communities and the spatial distribution of land gradients are illustrated in Figure 1. The geographical fragmentation that characterized the former KwaZulu is evident; the apartheid government forcefully resettled Africans to areas deemed inhospitable for white settlement, wherever those happened to be (Christopher, 2001). Gradient varies widely across the region and is very steep (dark shaded areas) in some areas. The Food and Agriculture Organisation categorizes the average gradient of this area as "strongly sloping" (FAO, 1998).

Several important features of project placement are evident in Figure 2, which shows the distribution of (dark shaded) treatment and control areas. Treated areas are not all positioned close to the 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 electrification between 1996 and 2001, although the first stage shows that proximity does raise the probability of treatment. Proximity to a town is also not necessary for treatment. Finally, treated areas are distributed across the province rather than clustered in one area. This allows for the inclusion of district fixed-effects that can absorb aggregate differences in employment growth rates across local labor markets. In the next section, I describe how pairing the variation in land gradient with the variation in project status helps to identify the effects of rural electrification.

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.[28] 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 of electrification α_2 by ordinary least squares:

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

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 at the community level, it is possible to include the community fixed effect μ_j and the district trend term $\lambda_d t$ but not δ_j, the community trend term.[29]

As outlined in section 3.2, 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.

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

[29]A linear relationship between employment and treatment is adopted for simplicity. Appendix 2 discusses results using an alternative specification for the outcome variable that allows for all variables to impact employment non-linearly. This logistic specification produces qualitatively similar results to the linear model.

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

Since $\delta_j + \Delta\epsilon_{jdt}$ are unobserved in this formulation, there are reasons to expect that estimating (2) by 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 the economic effects of infrastructure. 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 well 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 (Banerjee et al (2006)). If the areas that are losing jobs are also the ones targeted for projects, $\hat{\alpha}_{2,OLS}$ would be biased downwards. With only two years of data, we cannot identify α_2 separately from the community trend δ_j.

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 are assigned full treatment status when only a fraction of households in the area are treated. In addition, non-NEP electrification continued during this period in areas where households were willing to pay for their connections. 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 biased upwards, as

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

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 covariates (X_{jd0}) to control for some factors affecting a community's growth path (δ_j), where the baseline year is 1996. These include household density, fraction of households living in poverty, distance to the 1996 grid, distance to the nearest road and town, fraction of adults that are white or Indian (as a proxy for local employers) and measures of adult educational attainment in the area. Since the Census crudely captures income– in intervals and only at the household level– I also include two measures often used to provide additional information about the extent of poverty: the share of female-headed households and the female/male sex ratio (Standing et al, 1996).

Confounding trends in community-level employment and unmeasured political factors are still of concern.[32] Most individuals in this area are Zulu and Zulu-speaking, precluding the construction of an ethnic- or linguistic-heterogeneity index which is often used in the analysis of the political allocation of public goods. To overcome these issues, I instrument for program placement using average community land gradient (Z_j). The system of equations to be estimated is then:

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

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

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

To bolster confidence in the research design, I test for whether gradient is correlated with outcomes for areas treated before 1996, and whether it is correlated with changes in other development services including access to water and to flush toilets. I also test for whether gradient is correlated with changes in two possible sources of employment. In each case where we might be

[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.

[32]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 the context in which the 1994 transition to democracy heralded new national governance and national policies.

concerned that gradient could have a direct effect on outcomes, the data cannot reject zero correlation.

One obvious concern with using land gradient as an instrument in a rural setting is that it may affect agricultural outcomes or characteristics of individuals settling in steep and flat areas.[33] 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 appendix Table A2, constructed from 1996 Census micro-data).[34] In addition, validity of the instrument in column (3) is threatened only if non-random sorting of individuals across flat and steep areas resulted in differential employment growth, independent of the effects of new electrification. 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 titling and the role that tribal authorities (rather than the market) play in allocating land.[35] However, to check that differential migration flows are not driving the results, I use information on former place of residence to estimate (3) and (4) for the subset of incumbents. The main results are robust to this re-specification of the outcome variable.

Conditional on instrument validity, it is important to consider the interpretation of $\alpha_{2,IV}$. This parameter captures the local average treatment effect (LATE) of electricity projects on employment growth at a community level. It is typical to think about LATE's in terms of marginal effects for individuals who are affected by the instrument. In this paper, individuals aggregate to communities and so community composition will drive marginal effects. If individuals living in flatter areas can better afford electricity once it arrives, then a larger than average treatment effect may be measured for these areas. Or, if individuals living in flatter communities have fewer home production demands, they may also respond more to the arrival of the new technology. In addition, employment "returns"

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

[34]See Simkins (1981), Standing et al (1996) and Aliber (2001) for a description of historical and current agricultural conditions in ex-homeland areas.

[35]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).

to electrification may differ by gradient itself, leading to larger estimated employment effects for marginal than for average communities. Flatter areas always have lower commuting costs, so individuals in flatter areas always face a higher net wage. Since these individuals are initially closer to the employment participation margin, they will always be more likely to respond when electricity arrives.[36] These reasons imply that we might expect IV estimates to be larger than average treatment effects for a more general population. I explore several of these avenues for treatment effect heterogeneity in the last part of the paper.

6 Main Results

6.1 Summary statistics

Table 1 presents summary statistics on baseline (1996) 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).[37] On average, over half of households within a community are female-headed and the female/male adult sex ratio is well over 1. These variables reflect the historical function of the homelands as migrant-labor communities.

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

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

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

observable characteristics should be balanced across treatment and control groups. However in column (4) treated areas are significantly closer to towns and roads and more densely settled. These characteristics may capture some, but not all, differences across communities related to treatment status. In column (5), I instead compare values of each covariate across steep and flat areas. I regress each covariate individually on the treatment dummy, controlling for all other covariates. The results in column (5) show 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 in 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 inclusive. Column (2) indicates that over the period, employment rates fall by 4 percentage points for men in these areas (row 6). Female employment rates remain steady on average across communities but low, at 7 percent (first three rows). Employment is uniformly higher in treatment than in control communities in the baseline period (rows (1) and (4)). Comparing changes in employment rates in treated areas to the same change in control areas (column 7, row 3), the unadjusted estimate for women is not different from zero. For men, it is a statistically significant -2 percentage points (column (7), row (6)).

Both in- and out-migration are occurring over this period (Leibbrandt et al, 2003). Migration could contaminate any comparison of outcomes across treatment and control areas or it could be a response to treatment.[38] Some of these changes are evident in the lower part of Table 2. Population growth is faster in treated than in control areas, even though treated areas begin with higher populations (column (7), rows (7), (8), (10) and (11)). Treated areas grow at about 6 percent per year and control areas grow at about 3 percent.[39] Communities are small, so a 3 percent change in population over five years is an increase of about 30 people in the median community.

Two striking points emerge from Table 2: employment rates are very low for both men and women, and they are falling for men between 1996 and 2001.[40] The large reductions in male

[38] 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.

[39] For treated areas, the increase in the number of women is 136/426; for men, the increase over 5 years is 114/312. Averaging this increase over the five years between the Census, it translates into about 6% growth per year for treated areas. A similar calculation can be done for control areas.

[40] In these low employment communities, many households are supported by migrant remittances or state old-age pensions (Standing et al, 1996). The 1996 micro Census data show that 32 percent of African households in rural KZN contain either

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 mining sectors, and towards service and retail sectors. These changes were largely a continuation of trend, impacting heavily 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, the fraction of women working or looking for work increased by 10 percentage points between the mid-1990s and 2001, with even larger increases in rural areas. 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.

These general trends in the South African economy imply that labor market opportunities are changing for all men and women. Table 2 reflects that this restructuring is correlated with treatment status. Treated areas are closer to towns and formal jobs, near existing infrastructure and are more densely settled– all factors that raise the community's exposure to the industrial restructuring of the late 1990s. Political reasons for targeting household services investments towards disadvantaged areas would reinforce this correlation. All of these factors caution against interpreting the simple difference-in-differences comparisons as causal and recommend an IV strategy. Since very little of the restructuring was associated with changes in subsistence agriculture, there is no reason to think that gradient is correlated with these changes. However, even the IV estimates should be interpreted within the context of the broader economic changes occurring in South Africa at the time.

6.2 OLS and IV main results

First-stage estimates for assignment to treatment are presented in Table 3.[41] In columns (1) to (5), the outcome variable is the treatment indicator. A one standard deviation increase in gradient

a pensioner or a migrant worker, and 20 percent of households receive remittances from outside the household.

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

(about 10 degrees) reduces the probability of being treated by 4 percent. Across columns, the size of the coefficient does not change with the addition of more controls and precision improves, especially once district fixed-effects are added in column (3). When restricting to areas where no households had electricity in 1996 (column (5)), the gradient coefficient is slightly larger. 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 (column (6)), and decreases the fraction of households treated by 2 percentage points (column (7)). For the main analysis, I use the treatment indicator variable and instrument with mean gradient (the specification in column (4)).

The other two cost coefficients do 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, although this is not significant when other controls are added. A one standard deviation increase in household density (30 households) per square kilometer increases the probability of treatment by 1 to 2 percent and this is robust across specifications.

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

In order for electrification to change patterns of employment through the channel of reduced time in home production, households need to switch out of traditional fuels when their communities are connected to the grid. Electricity projects did indeed change patterns of household fuel use and quite dramatically, as Table 4 indicates. Each coefficient reported in the table is from a separate OLS or IV regression, where the outcome variable is the change in the fraction of households using electricity for lighting or cooking or using wood for cooking. Columns (1) and (3) do not contain any additional controls while columns (2) and (4) report results from regressions containing all relevant control variables. Average electrification rates rise by 23 percentage points more in treated than in control areas in the OLS comparison in column (2). In the same column, we see that reliance on wood for cooking falls by 4.2 percentage points and cooking with electricity rises by 6 percentage points. Column (4) indicates that 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. Both the OLS and IV regressions illustrate substantial shifts towards using electricity for home production and the IV results are substantially larger than OLS estimates.

To check that gradient is not simply picking up easier access to all types of services that could make home production easier, rows (4) and (5) of Table 4 present results for two additional outcome variables: the change in fraction of households with access to piped water in the home or within 200 meters of the house, and the change in fraction of households with a flush toilet at home. There is no evidence that treated regions experience differential changes in these basic services. In fact, the IV results for water services in column (5) and (6) are in the opposite direction to what we would expect if gradient was simply a noisy measure of wealth: slightly flatter areas have larger reductions in access to water sources close by, although these estimates are not significant once all controls are added. The change in access to flush toilets is not systematically associated with treatment status or with land gradient. Table 4 demonstrates 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 in the IV specifications than in the treatment-control OLS comparisons.

Employment effects for men and women are consistent with this latter point. Tables 5 and 6 present the main OLS and IV results for African women and men.[42] In each column, the dependent variable is the change in sex-specific employment rate between 1996 and 2001. The coefficient on treatment in column (1) reflects the falling employment rates from Table 1: there is no significant change in female employment across treated and control areas while male employment falls by 1.8% in treated areas, relative to control areas. Adding controls and district fixed-effects in columns (2) and (3) increases the coefficient on treatment slightly, with the female employment effect still not

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

significantly different from zero and male employment becoming less negative and less statistically significant. The positive, significant coefficients on community poverty rate, sex ratio and female-headed households in both tables indicate that female and male employment is growing faster in poorer places. It is reassuring that the coefficients on variables other than treatment are largely consistent in sign and magnitude for male and female regressions. For instance, the poverty variables all have the same sign and significance across OLS and IV results, suggesting that gradient is not correlated with community poverty.

IV estimates of the treatment effect are substantially larger than OLS estimates and significantly positive for women as seen in columns (5) to (8). Since gradient is correlated with some of the control variables as Table 1 indicated and the F-statistic on the excluded variable in the first stage is larger with other controls absorbing residual variation, it is preferable to focus on results in columns (6) to (8). In these columns, female employment increases by 13.5 percentage points in areas induced to get the treatment by gradient.[43] The Anderson-Rubin (AR) test for the significant of the treatment for female employment strongly rejects zero and the confidence interval is wider than the standard confidence interval, 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 the standard test or the AR test (column (8), Table 6).[44] This observation is consistent with a model in which electricity has effects on the employment of primary home production workers and facilitates women's rather than men's entry into market work. The difference in the male and female employment effects estimated in Table 5 and 6 is significant at the 14% level.[45] IV results for female employment are also more sensitive to the inclusion of district fixed effects than the male employment results (compare columns (6) and (7) in each table). This difference may reflect that the market for female labor is less spatially integrated than the market for male labor, since, if labor was perfectly mobile across districts, we should not see substantial differences in the effects of

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

[44] The reduced form for men shows no correlation between gradient and employment.

[45] 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.

electrification with or without controls for district fixed effects.

The 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).

6.3 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_{jdt} (which I have argued is likely) and when ΔT_{jdt} is measured with error. However, the valid IV that is uncorrelated with $\delta_j + \Delta\epsilon_{jt}$ will tend to be correlated with any non-classical measurement error in the binary variable ΔT_{jdt}. In this situation, even if the instrument deals with the omitted variables bias, the measurement error in ΔT_{jdt} could lead to an upwards-biased IV estimator.[46]

To detect how much of the difference in OLS and IV results is due to measurement error, I restrict to samples where ΔT_{jdt} should be measured with less error. The first two columns of Table 7 reproduce the main result for females in the full sample while columns (3) to (6) present results for successive sample limitations. To identify communities where projects had greater coverage, I exclude treated areas with less than a 10 percent change in coverage of electric lighting, and treated areas where the connection rate between 1996 and 2001 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, indicating that some measurement error in the treatment variable is present. 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 statistically significant due to the smaller sample size.

Although effects estimated under the OLS specification for these sub-samples are just 1

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

percentage point or higher for female employment, they are still well smaller than the IV results. The measurement error in the treatment dummy 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.

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. However, 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.

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.[47]

[47] I perform a similar placebo test comparing the change in employment rates in communities that are scheduled to be treated after 2001 to communities that are never treated with an electricity project. There are again no significant differences (results not shown).

7.2 Do flatter areas experience 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. For example, businesses may expand in flatter ex-homeland areas after the end of apartheid for reasons unrelated to electrification.

Since our effects are significant for female but not male employment, information about the major employers of women in the area is useful for testing whether labor demand is directly expanding in these flatter areas. Appendix Table A2 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) 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, so labor demand shocks in these two industries are the most likely candidates for confounding IV estimates of electrification effects.

Using two waves of the South African Schools Register of Needs (1995, 2000) that capture the location of schools, I construct a variable measuring the change in the number of schools in each community over time. Over the five-year period, the number of schools across the rural KwaZulu-Natal area increases by 19 percent, from 1,770 to 2,801. This 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 (column (3)). While school placement (and hence teacher hiring) is probably related to the distribution of children in space, this placement 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 fraction of adult population that is Indian or white. These are the individuals most likely to hire household workers (Dinkelman and Ranchhod, 2008). The number of Indian and white adults is not changing differentially across areas of different gradient, as column (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.

While it would be ideal to perform similar tests using a measure of the number of other firms in the area over time, the tests related to schools and white and Indian households as a source of labor demand are still informative since employed women are most likely to be working as teachers or domestic workers. While more jobs are being created at the low end and in the public sector during the 1990s(Casale and Posel, 2004), there is no evidence that these job openings are occurring differentially in flatter parts of KZN.

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 spillover effects into neighboring areas. Spillovers could be positive or negative. For example, if firms create jobs for people living in neighboring areas, positive spillovers in these control areas would dampen treatment effects. If people move out of neighboring control areas towards 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 spillover effect. In both cases, OLS and IV coefficients should be substantively 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.[48] 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

[48] 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.

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 supports the idea electrification increased employment primarily through a labor supply channel.

8.2 Heterogeneous effects of electrification related to income

IV estimates identify effects for communities which are cheaper to electrify by virtue of having a flatter gradient. In these communities, female employment may be more responsive to electrification than in an average 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 home production technologies when the new service arrives.

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

The graph in Figure 3 shows the fraction of the predicted poverty quintile that is treated, for communities in the flattest and steepest halves of the gradient distribution. Both lines slope upwards, indicating that areas with higher predicted values of the poverty index (i.e. richer) are more likely to actually be treated. The gap between the two lines shows that flatter areas are systematically more likely to be treated than steeper areas. The middle-poorest and second-richest quintiles are most likely to have treatment probability manipulated by the instrument which we can see from the larger gap between the lines occurring at these quintiles. This larger gap is also seen in column (3) of Table 10 which shows the difference in treatment probability for flat versus steep areas, within each quintile. In column (4), I compute the contribution of each quintile to the final IV estimate by calculating a weight: the middle quintile and the second richest quintile together contribute over 70 percent to the IV result.[49]

Why might middle-quintiles in particular have large employment effects? These communities contain households that experience large changes in home production technology when electricity arrives. Middle-poor areas are initially less likely to be using electricity than richer areas and more reliant on wood for cooking, as columns (1) and (2) of Table 11 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 11 present within-quintile reduced-form coefficients from regressions of the change in fuel use for different home production activities on a gradient dummy (1 is flat, 0 is steep). These columns indicate large increases in the use of electricity and large decreases in reliance on wood for cooking in flatter areas of middle-poor, second-richest and richest areas.[50] Finally, column (7) of Table 11 indicates that the female employment result is indeed driven by women living in middle- and second-richest quintile communities: the effects for these communities are large, positive and significant and are weighted most heavily in the final IV results.[51]

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

Women who have additional home-production responsibilities are less likely to be able to respond to new access to electricity, even though their productivity at home may be substantially enhanced by the use of electricity. For example, child-care responsibilities raise the value of a woman's time at home and in the absence of pre-school care (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,

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

[50]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.

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

not shown). Children also create work at home though, and so the more children in the house that require child-care, the more time can potentially be saved with access to more a efficient power source.

Census micro data from 1996 give some indication of which women are more likely to live with a child younger than age 9. Figure 4 is a lowess-smoothed graph of the fraction of women of each age living with at least one child aged 9 or under. The graph is drawn for African women between ages 15 and 59 living in rural areas of KZN and shows a clear distribution of youngest children to households with both younger and older women.[52] 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 might 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 age cohort: the estimated coefficients sum to the main treatment coefficient in column (8) of Table 5. Figure 5 presents the IV coefficients (and standard errors) on the treatment dummy for separate regressions.[53] 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.[54]

Since aggregate data do not allow me to identify exactly which women have children of which ages, I capture these other demands on female time as the ratio of the number of children ages 5 to 14 living in the community in 2001 to the number of households in 2001. These children will have been ages 0 to 9 in the period between 1996 and 2001– just those ages that require full-time care before formal school enrollment. I control for this historical variable in the main regression and

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

[53]Results for men and OLS results for women are not shown as the treatment coefficient was never significant for any cohort.

[54]The coefficient on treatment for the group of women aged 40-44 is large and positive (0.012) but not quite statistically different from zero.

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 where the ratio of young children to households is higher. 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 0.6 percentage points (-0.607*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 in the home appear to reduce the employment impact of a new home production technology.[55]

8.4 The role of migration

Out-migration from rural areas is occurring during this period of roll-out.[56] Cross et al (1998) also document rural-to-rural migration in KZN in the 1990s and show that part of this migration is towards areas with better infrastructure and amenities. Each of these flows could alter the composition of the population in treated and non-treated communities and contribute to the employment effects.

We can get some sense of the prevalence of migration flows at a more aggregated level using the 10% sample of the 2001 Census micro data. These data provide information on the district that an individual was living in as of 1996 and the district in which they report living in 2001. Matching reports of out-migrants from each district to the total number of people living in each district in 1996, there is evidence of substantial out-migration from the rural KZN districts in our sample. Fifteen percent of men and 10 percent of women report out-migration during the five-year period.[57]

[55] 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 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 for this.

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

[57] 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 available from the author on request.

These out-migration rates are, however, not significantly different by gradient.[58]

Table 13 presents differences in population growth rates in treatment and control areas, which tell us something about net migration at the community level. Since these areas are small, numerically small increases in population can translate into large percentage changes. The first two columns of the table indicate that treated areas have significantly higher population growth than control areas, both in the OLS and IV results. Over the five-year period, the population in treated areas grows 25.8 percent more than in control areas, and this growth is 400 percent under the IV specification.

Ideally, we could isolate employment growth for in-migrants and for incumbents. Although it is not possible to identify the migrant status of employed individuals in the Census, I assume that all individuals who report themselves as recent in-migrants are employed. I redefine the dependent variable by excluding this total number of recent in-migrants from the numerator of the employment rate in 1996 and in 2001. The new outcome variable is therefore a conservative measure of changes in employment rates for incumbents only.[59] This helps to move part way towards understanding the contribution of in-migration to the total employment result.

Table 13 provides results separately for men and women in columns (3) to (6). For women, OLS and IV results are remarkably similar across the full definition of employment and the migrant-excluded definition. Using a Hausman test to compare the treatment coefficient for female employment defined in the original way and in this more conservative manner, we cannot reject the hypothesis that the coefficients are the same. This is the case for both the OLS and the IV specifications. Once again, the Anderson-Rubin test rejects a zero treatment effect for women in column (5), with a confidence interval from 0.05 to 0.45. 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 8.2 percentage points higher in treated regions compared to non-treated regions, but still not statistically significantly different from zero.

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

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

Within the limitations of the aggregate Census data then, there is some evidence in columns (1) to (3) that in-migration of individuals towards treated areas may be an additional response to electricity projects. Given reported preferences for household services, this effect is not surprising.[60] However, there is no strong evidence to suggest that in-migration or out-migration contaminates the IV estimates of employment for women.

9 Conclusion

This paper uses a period 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 the use of 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 and second-richest communities that initially rely heavily on wood for cooking and are able to respond more when the new service becomes accessible. Effects are also larger for women in their thirties and forties, and there is some evidence that this is related to fewer child-care responsibilities at these ages.

These results represent some of the first pieces of evidence on the impact of infrastructure for rural electrification in a developing country.[61] They highlight the importance of measuring employment effects in infrastructure evaluations more generally and of paying attention to existing conditions of economic restructuring in understanding these measured effects. 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.[62]

[60] 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.

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

[62] Effects of new fertility-control technologies or falling appliance prices in the US have been analyzed over periods characterized by World Wars, changing social norms and alterations in the structure of jobs available for women. See for example,

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 in institutionally weak environments.

Finally, the finding that female employment in rural areas of South Africa responds to household electrification contributes to a large literature on how female labor force participation responds 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 microeconomic aspects of economic development and 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 types of jobs and occupations women first choose to enter when they are able to use their labor outside of home production.

Bailey (2006), Goldin and Katz (2000) and Greenwood et al (2005).

References

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

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

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

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

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

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

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

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

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

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

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

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

Card, David, "Using geographic variation in college proximity to estimate the returns to schooling," in L.N. Christofides et al, ed., *Aspects of labour market behaviour: Essays in honor of*

John Vanderkamp, Toronto: University of Toronto Press, 1995, pp. 201–221.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Field, Erica, "Entitled to work: Urban property rights and labor supply in Peru," *Quarterly Journal of Economics*, November 2007, *122* (4), 1561–1602.

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

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

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

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

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

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

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

Greenwood, Jeremy, Ananth Seshadri, and 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.

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

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

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

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

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

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

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

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

Leibbrandt, Murray, Laura Poswell, Pranushka Naidoo, Matthew Welch, and Ingrid Woolard, "Measuring recent changes in South African inequality and poverty using 1996 and 2001 Census data," 2003. Center for Social Science Research Working Paper 05-94.

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

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

Michael, Robert T., "Education in non-market production," *Journal of Political Economy*, 1973, *81* (2), 306–327.

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

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

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

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

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

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

Qian, Nancy, "Missing women and the price of tea in China: The effect of relative female income on sex imbalance," 2006. Brown University $www.chass.utoronto.ca/\ brandt/ECO2703/papers/Qian_2006_Missing_mimeo.pdf$.

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

Saghir, Jamal, "Energy and Poverty: Myths, Links and Poverty Issues," Technical Report, The World Bank $http://esa.un.org/un-energy/pdf/UN-ENRG/20paper.pdf$ 2005. Energy

Working Notes No 4.

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

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

South African Department of Minerals and Energy, "National electrification programme summary evaluation report: 1994-1999," Technical Report, South African Department of Minerals and Energy *www.dme.gov.za/pdfs/energy/electricity/summary*$_o$*bjectives.pdf* 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*, Geneva: International Labour Office, 1996.

Statistics South Africa, *Census Mapping Manual 2000.*

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

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

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

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

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

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

Wittenberg, Martin, "The weight of success: The body mass index and economic well-being in South Africa," October 2007. South African Labour and Development Research Unit

$http://ipc.umich.edu/edts/pdfs/wittenberg.pdf$.

Appendix 1: Data

Census data

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

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

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

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

Boundaries for communities from the 2001 Census define the main unit of analysis. Since boundaries have shifted over time (Christopher, 2001), the 1996 (smaller) areas are aggregated up to the (larger) 2001 boundaries.[63] The matched identifiers from this panel of areas are used to extract Census aggregate data in 1996 and 2001. For each 1996 EA, the proportion of the EA polygon area that falls inside each 2001 community is calculated. This proportion is used as a weight to assign a proportion of the 1996 EA data to the 2001 community. The key assumption in this process is that people are uniformly distributed over 1996 EA's.

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

Geographic data

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

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

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

Electricity project data

Data on Eskom projects in KwaZulu-Natal were provided by Sheila Brown. The project list details the number of pre-paid electricity connections per Eskom area in each year from 1990 to 2007. The

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

year of treatment is defined as the year in which a community experienced a spike in household connections (concentrated project activity). Areas are referenced by name and village code. Eskom's planning units do not line up accurately with Census regions. To match project data to Census regions, the project data were first mapped to a physical location (using a spatial database of transformer codes that corresponded to project codes) and these locations were then merged back to Census spatial boundaries.

Schools Register of Needs

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

Appendix 2: Sensitivity analyses

The main equation in levels is :

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

and in first differences:

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

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

$$\frac{\partial \Delta y_{jdt}}{\partial \Delta T_{jdt}} = \frac{\partial y_{jdt}}{\partial T_{jdt}} = \alpha_2 \tag{A-3}$$

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 the initial employment rate. Specifying y_{jdt} in logistic form would satisfy this requirement.

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

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:

$$ln(\frac{y_{jdt}}{(1-y_{jdt})}) = \alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \lambda_d t + \epsilon_{jdt} \tag{A-5}$$

$$\Delta(ln(\frac{y_{jdt}}{1-y_{jdt}})) = \alpha_1 + \alpha_2 \Delta T_{jdt} + \delta_j + \lambda_d + \Delta\epsilon_{jdt} \tag{A-6}$$

With this transformation of the dependent variable, OLS and IV can still be implemented but calculating the average marginal effects of the treatment is tricky. Let $\Lambda(.) = \frac{e^{\alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \lambda_d t + \epsilon_{jdt}}}{1 + e^{\alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \lambda_d t + \epsilon_{jdt}}}$, then :

$$\frac{\partial y_{jdt}}{\partial T_{jdt}} = \alpha_2 \Lambda(.)(1 - \Lambda(.)) \tag{A-7}$$

The marginal effect for each community depends upon the initial values of $\Lambda(.)$. This marginal effect is difficult to calculate since neither α_0 nor μ_j is estimated. As an approximation, I calculate the marginal effect for the average community by using $\bar{y_{jd0}}$. Thus, $\hat{\alpha_2}\hat{\Lambda(.)}(1 - \hat{(}\Lambda(.)))$*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 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 statistically significant. Since many communities have low initial employment rates, these areas have the potential to experience large increases in employment. For women, the lower bound of the Anderson-Rubin confidence interval is the same in linear and logistic models, while the upper bound is larger in the logistic model; both confidence intervals strongly reject zero treatment effect.

As in the linear model, not too much weight should be placed on results for African men, since there is no reduced form relationship between male employment and gradient in either model. OLS and IV estimates are insignificant in the non-linear specification. Male and female effects are significantly different at the 17% level.

Table A1: Average marginal effects for linear and logistic specifications

	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	1,992	1,992	1,992	1,992	1,992	1,992	1,992	1,992
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.

(2) Significant at the p<0.01***, p<0.05** and p<0.1* levels.

(3) Standard confidence intervals for $\hat{\beta_T}$ appear underneath the F-statistic on gradient (Z).

(4) Anderson-Rubin confidence intervals (robust to heteroscedasticity) appear in the final row.

(5) Robust standard errors, clustered at the sub-district level.

Appendix Table A2: Distribution of employment across occupations, Census 1996

	Employed men			Employed women		
Occupation category	Fraction	s.d.	N	Fraction	s.d.	N
Manager	0.014	(0.033)	1,988	0.011	(0.041)	1,984
Professional	0.126	(0.154)	1,988	0.321	(0.266)	1,984
Clerk	0.020	(0.051)	1,988	0.028	(0.060)	1,984
Service worker	0.103	(0.101)	1,988	0.042	(0.084)	1,984
Skilled agriculture	0.047	(0.087)	1,988	0.038	(0.092)	1,984
Subsistence agriculture	0.001	(0.009)	1,988	0.001	(0.013)	1,984
Crafter	0.169	(0.142)	1,988	0.048	(0.089)	1,984
Machine operator	0.130	(0.113)	1,988	0.016	(0.052)	1,984
Elementary occupation	0.178	(0.158)	1,988	0.338	(0.239)	1,984

(1) The table presents the fraction of adult African men and women who are employed in 1996, in each of 9 different occupations groups, averaged across communities. Data are from 1996 Census community data.

(2) Managers include managers, legislators and senior officials.

(3) Professional workers include technicians and associate professionals such as teachers, nurses and other public officials.

(4) Clerks include administrative and clerical workers.

(5) Service workers include shop and sales assistants and retail sector workers.

(6) Skilled agriculture include includes all commercial farmers; subsistence farming is farming for own account.

(7) Crafters include extraction and building trade workers, metal, machinery and related trades, handicraft, printing and related trade workers.

(8) Machine operators include manufacturing workers and assemblers.

(9) Elementary occupation workers include drivers, domestic workers, messengers and unskilled farm laborers.

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)
Fraction female-headed hh's	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)
Fraction Indian and white adults	0.02	0.01	0.02	-0.016***	-0.001
	(0.21)	(0.03)	(0.23)	(0.007)	(0.001)
Kilometers to town	38.31	36.07	38.85	-2.789**	0.081
	(24.60)	(24.10)	(24.70)	(1.387)	(0.093)
Kilometers to road	39.01	36.80	39.55	-2.751***	-0.091
	(18.29)	(15.32)	(18.91)	(1.030)	(0.093)
Fraction men with high school	0.06	0.08	0.06	0.016***	0.000**
	(0.05)	(0.05)	(0.05)	(0.003)	(0.000)
Fraction women with high school	0.07	0.08	0.06	0.021***	0.000
	(0.05)	(0.06)	(0.05)	(0.004)	(0.000)
Households per km^2	20.67	30.76	18.21	12.549***	-0.465***
	(29.50)	(48.15)	(22.07)	(4.077)	(0.159)
Kilometers 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)	
Land gradient - std. dev.	10.84	10.32	10.97	-0.653	0.336***
	(3.92)	(3.82)	(3.93)	(0.490)	(0.011)
Land gradient - range	52.51	50.26	53.06	-2.793	1.276***
	(15.75)	(15.34)	(15.81)	(2.037)	(0.047)
N communities	1,992	391	1,601		1,992

(1) Community-level means (s.d.) in columns (1)-(3). Column (4) shows treatment (T) mean - control (C) mean and whether this difference is significant. Column (5) shows coefficients from regressions of each covariate on gradient, controlling for all other covariates and district fixed effects. Robust standard errors in columns (4) and (5) are clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level.

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

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

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	(1)	1996	0.07	0.00	0.91	0.09	0.07	0.02***
			(0.08)			(0.07)	(0.08)	(0.00)
	(2)	2001	0.07	0.00	0.80	0.08	0.07	0.02***
			(0.07)			(0.07)	(0.08)	(0.00)
	(3)	Δ_t	0.00			-0.004	0.00	-0.004
			(0.00)			(0.00)	(0.00)	(0.00)
Male Employment Rate	(4)	1996	0.14	0.00	0.99	0.16	0.13	0.03***
			(0.12)			(0.11)	(0.12)	(0.01)
	(5)	2001	0.10	0.00	0.83	0.11	0.10	0.01**
			(0.10)			(0.09)	(0.10)	(0.01)
	(6)	Δ_t	-0.04			-0.05	-0.03	-0.02***
			(0.00)			(0.01)	(0.00)	(0.01)
N Adult Females	(7)	1996	356.07	1	4,553	426.02	338.98	87.04***
			(347.84)			(379.53)	(337.58)	(19.53)
	(8)	2001	421.45	6	3,392	560.88	387.39	173.49***
			(401.50)			(494.99)	(367.37)	(22.32)
	(9)	Δ_t	65.38			134.86	48.41	86.45***
			(11.90)			(31.54)	(12.47)	(16.84)
N Adult Males	(10)	1996	253.15	1	3,135	312.53	238.65	73.89***
			(261.74)			(291.11)	(252.03)	(14.68)
	(11)	2001	310.17	3	2,770	427.37	281.55	145.82***
			(314.09)			(403.18)	(281.00)	(17.42)
	(12)	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 (2), (5) and (6) contain variable means and standard deviations (s.d.).

(2) Mean differences and standard errors are shown for Δ_{T-C} in column (7) and Δ_t in rows (3), (6), (9) and (12).

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

(4) All variables constructed for Africans adults (ages 15-59) only.

Table 3: Assignment to treatment using alternative treatment definitions - OLS results

	Treatment Indicator [1/0]					Year treated	Fraction treated
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gradient*10	-0.040*	-0.040**	-0.040***	-0.040***	-0.050***	-0.120**	-0.020**
	(0.020)	(0.020)	(0.010)	(0.010)	(0.02)	(0.050)	(0.010)
Kilometers to grid*10		-0.050**	-0.020	-0.020	0.017	-0.010	-0.020
		(0.020)	(0.020)	(0.020)	(0.023)	(0.060)	(0.010)
Household density*10		0.020***	0.010**	0.010**	0.012	0.050***	0.000
		(0.000)	(0.010)	(0.010)	(0.012)	(0.020)	(0.000)
Poverty rate		0.034	0.034	0.029	0.073	0.016	0.046
		(0.066)	(0.067)	(0.066)	(0.068)	(0.212)	(0.044)
Adult sex ratio (f/m)		0.350***	0.134	0.124	0.084	-0.080	0.055
		(0.118)	(0.104)	(0.104)	(0.108)	(0.365)	(0.075)
Female-headed hh's		-0.164***	-0.117***	-0.111***	-0.087*	-0.344*	-0.029
		(0.048)	(0.038)	(0.038)	(0.052)	(0.416)	(0.025)
Indian and white adults		-0.693***	-0.576**	-0.571**	53.016	-1.789***	-0.320**
		(0.256)	(0.250)	(0.230)	(60.482)	(0.168)	(0.151)
Kilometers to road*10		0.000	-0.010	-0.010	-0.004	0.200	0.000
		(0.010)	(0.010)	(0.010)	(0.008)	(0.300)	(0.000)
Kilometers to town*10		0.020	0.010	0.010	-0.016	-0.400	0.000
		(0.010)	(0.010)	(0.010)	(0.016)	(0.500)	(0.010)
Men with high school		-0.041	0.090	0.087	-0.237	-0.265	0.273
		(0.451)	(0.399)	(0.387)	(0.469)	(1.253)	(0.221)
Women with high school		0.836**	0.726*	0.783**	1.157**	2.493*	0.219
		(0.419)	(0.395)	(0.375)	(0.469)	(1.280)	(0.209)
Change in water access				0.016	0.109*	-0.265*	0.020
				(0.045)	(0.065)	(0.159)	(0.041)
Change in toilet access				0.178**	0.444	0.931**	0.083*
				(0.086)	(0.272)	(0.382)	(0.047)
District FE	N	N	Y	Y	Y	Y	Y
Sample	All	All	All	All	No elect in 96	All	All
$\bar{y}$	0.196	0.196	0.196	0.196	0.187	0.332	0.096
N	1,992	1,992	1,992	1,992	477	1,992	1,992
R^2	0.009	0.075	0.168	0.169	0.187	0.150	0.199
F-stat on instrument(s)	3.610	5.400	9.060	8.870	9.670	4.920	6.000
Prob>F:	0.060	0.020	0.000	0.000	0.000	0.030	0.010

(1) Dependent variables: columns (1)-(5) is indicator for treatment (1 if the area had a project in between 1996 and 2001), zero otherwise; column (6) 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 (7) is fraction of 1996 households connected between 1996 and 2001. Ten district fixed-effects included in columns (3) to (7). Land gradient in degrees, all distances measured in kilometers.
(2) Robust standard errors clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level.
(3) Sample in column (5) is restricted to set of areas where no households had electricity in 1996.
(4) All controls measured in 1996, except change in access to water and flush toilet measured between 1996 and 2001. Change in water captures change in fraction of households with access to water in the house or less than 200 meters away; change in toilets captures the change in fraction of households with a flush toilet.

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

		OLS		IV	
	$\bar{y}$	No controls	Controls	No controls	Controls
Outcome is Δ_t	(1)	(2)	(3)	(4)	(5)
(1) Lighting with electricity	0.08	0.258***	0.233***	0.661***	0.713***
		(0.031)	(0.031)	(0.233)	(0.232)
(2) Cooking with wood	-0.03	-0.049***	-0.042***	-0.305	-0.283*
		(0.012)	(0.012)	(0.197)	(0.148)
(3) Cooking with electricity	0.03	0.069***	0.059***	0.281**	0.241**
		(0.008)	(0.008)	(0.125)	(0.099)
(4) Water nearby	0.01	-0.03	0.009	-0.449*	-0.287
		(0.028)	(0.023)	(0.251)	(0.231)
(5) Flush toilet	0.09	0.003	0.01**	0.042	0.095
		(0.006)	(0.005)	(0.080)	(0.067)

(1) Cells contain treatment coefficients (robust standard errors clustered at sub-district level) from OLS regressions of dependent variable on treatment dummy and all explanatory variables. Significant at p<0.01***, p<0.05** or p<0.1* level.

(2) Column (1) contains the mean value of each dependent variable, the change in the fraction of households using electricity for lighting or cooking, using wood for cooking, or having access to nearby water or adequate toilet facilities.

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

(4) Excluded instrument is mean community gradient.

(5) Other controls included in each regression: distance to grid, household density, community poverty rate, adult sex ratio (F/M), share of female-headed households, share of Indian/white adults, distance to nearest road, distance to nearest town, share of adult men and women with at least high school, change in fraction of households with water close by, change in proportion of households with flush toilets and ten district fixed-effects. Change in water (toilet) access excluded from controls in rows (4) and (5).

(6) Anderson-Rubin confidence intervals for $\hat{\beta_{T,IV}}$ (not shown here) contain only positive values 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 result.

(7) Each regression contains N= 1,992.

Table 5: Effects of electrification on female employment

Outcome is Δ_t female employment rate	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)
Kilometers 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)
Kilometers 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)
Kilometers 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 high school		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 high school		-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. Significant at p<0.01***, p<0.05** or p<0.1* level.

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

(4) Excluded instrument is mean community land gradient.

(5) All controls (except treatment indicator and change in access to water and toilet services) are measured in 1996. Change in access to water and toilets is measured between 1996 and 2001. Ten district fixed effects included in columns (3),(4),(7),(8).

(6) For the IV results, standard confidence intervals are provided as well as confidence intervals from the Anderson-Rubin test. The AR test is robust to weak instruments and is implemented to be robust to heteroscedasticity.

Table 6: Effects of electrification on male employment

Outcome is Δ_t male employment rate	OLS (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)
Kilometers 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)
Kilometers 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)
Kilometers 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 high school		-1.366	-1.385	-1.407		-1.399	-1.468	-1.491
		(0.113)	(1.153)	(1.258)		(1.258)	(1.241)	(1.343)
Women with high school		1.203	1.074	1.405		0.615	0.712	0.997
		(0.123)	(1.258)	(1.314)		(1.551)	(1.451)	(1.506)
Change in water access				0.278***				0.273***
				(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-.2]	[-.1-0.2]
AR 95% C.I.							[-.05-.25]	[-.05-.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. Significant at p<0.01***, p<0.05** or p<0.1* level.

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

(4) Excluded instrument is mean community land gradient.

(5) All controls (except treatment indicator and change in access to water and toilet services) are measured in 1996. Change in access to water and toilets is measured between 1996 and 2001. Ten district fixed effects included in columns (3),(4),(7),(8).

(6) For the IV results, standard confidence intervals are provided as well as confidence intervals from the Anderson-Rubin test. The AR test is robust to weak instruments and is implemented to be robust to heteroscedasticity.

Table 7: Contribution of measurement error in treatment to female employment result

Outcome is Δ_t female employment	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 includes control areas and:	All treated areas		Treated areas with over 10% change in coverage		Treated areas with over 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 that controls for all covariates. Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level.

(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 in which 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. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression.

Table 8: Placebo experiment and reduced form for employers of women - OLS results

	Placebo Experiment		Growth in female employers	
	Δ_t Female employment	Δ_t Male employment	Δ_t Schools (N)	Δ_tIndian/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	0.028

(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 and white adults in the community between 1996 and 2001. Each regression controls for all covariates.

(2) Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level.

(3) In columns (1) and (2), sample includes only areas that had projects prior to 1996. Full sample in columns (3) and (4) includes all areas treated between 1996 and 2001 or never treated.

(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. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression. In column (4), the level of Indian/white adults in the community is excluded from regression.

(5) Schools data are taken from the 1995 and 2000 Schools Register of Needs. See data appendix for details.

Table 9: Testing for spillovers in female employment by excluding adjacent control areas

	Treatment coefficient		
Δ_t Female employment	OLS	IV	N
	(1)	(2)	(3)
Full Sample	0.001	0.135**	1,992
	(0.005)	(0.062)	
Sample excludes:			
Control areas within 1 kilometer of a treated areas	-0.005	0.104*	1,332
	(0.006)	(0.061)	
Control areas within 5 kilometers of a treated areas	-0.005	0.114	686
	(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. Significant at p<0.01***, p<0.05** or p<0.1* 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 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 education. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression.

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

Quintiles of Predicted Poverty Index	Fraction of sample in quintile (1)	Variance of gradient by quintile (λ_q) (2)	$\Delta\hat{T}\vert_q$ (3)	IV weight (w_q) (4)
Richest quintile	0.24	0.196	0.044	0.138
			(0.044)	
Second richest	0.21	0.203	0.112***	0.319
			(0.040)	
Middle quintile	0.21	0.211	0.127***	0.392
			(0.038)	
Second poorest	0.18	0.210	0.055*	0.145
			(0.033)	
Poorest quintile	0.16	0.199	0.003	0.006
			(0.037)	

(1) Significant at p<0.01***, p<0.05** or p<0.1* level.

(2) Predicted poverty values are 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 then assigned to quintiles, where quintile cut-points are defined by the regression sample i.e. those communities in the steepest half of the gradient distribution.

(2) λ_q is the estimated conditional variance of the gradient dummy (1 =flat, 0 =steep) within each quintile (q): $\hat{E}(P[Z|x,q][1-P(Z|x,q)|q])$.

(3) $\Delta\hat{T}|q = \hat{E}(E((T|z=1,x,q) - \hat{E}(T|z=0,x,q)|q))$ is the estimated difference in treatment probability across top and bottom halves of gradient distribution within each quintile, controlling for covariates. Each estimated coefficient in column (3) is on the interaction of the gradient dummy with each predicted quintile dummy.

(4) w_q is the weight that each quintile contributes to the IV estimate. As described in Kling (2001), it is computed across the columns (1)-(3): for each q, $\omega_q = \frac{[(1)_q*(2)_q*(3)_q]}{\sum_q[(1)_q*(2)_q*(3)_q]}$

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

	Fuel use in home production: Fraction using [X] in 1996			Δ_t Fuel use in home production: Within-quintile difference by gradient			Δ_t Employment by gradient	
Quintile of predicted poverty index	Electric lighting	Electric cooking	Wood cooking	Electric lighting	Electric cooking	Wood cooking	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Richest quintile	0.195	0.100	0.634	0.054*	0.024***	-0.026	-0.003**	-0.015
	(0.282)	(0.171)	(0.300)	(0.032)	(0.010)	(0.017)	(0.008)	(0.010)
Second richest	0.086	0.038	0.762	0.055***	0.026***	-0.037***	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**	0.021***	-0.029***	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	0.003	-0.012	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	-0.002	-0.001	0.006	0.002
	(0.060)	(0.023)	(0.151)	(0.017)	(0.007)	(0.017)	(0.006)	(0.008)

(1) Significant at $p<0.01$***, $p<0.05$** or $p<0.1$* level.

(2) Columns (1)-(3) present the quintile means in 1996, 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.

(3) 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 sub-sample.

(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. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression.

(5) Columns (7) and (8) are akin to reduced-form regressions of the outcome variables on a binary version of the instrument, within quintile.

Table 12: Treatment effect heterogeneity related to child-care

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

(1) Dependent variable is change in proportion of employed African females.

(2) Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level.

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

(4) 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 between 1996 and 2001. The mean (std. dev.) of this variable is 0.87 (.20).

(5) Excluded instruments are average land gradient and the interaction of gradient with the ratio of children to households measured in 2001.

(6) 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. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression.

Table 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)	IV (4)	OLS (5)	IV (6)
Treatment	0.258***	4.00***	0.000	0.133*	-0.010*	0.082
	(0.062)	(1.501)	(0.005)	(0.071)	(0.006)	(0.069)
N	1,992	1,992	1,992	1,992	1,992	1,992
R^2	0.072		0.025		0.137	
Standard C.I.	[.14-.38]	[1.05-6.94]	[-.009-.009]	[-.006-.27]	[-.021-.002]	[.0-.3]
AR confidence interval		[2.15-7.0]		[.05-.45]		[.0-.3]

(1) Dependent variable in columns (1)-(2) is change in log African population, in columns (3)-(6) is change in employment rate of African females or males where the numerator has been adjusted downwards to exclude those African adults who report they have moved to the area in the five years before the relevant Census year.

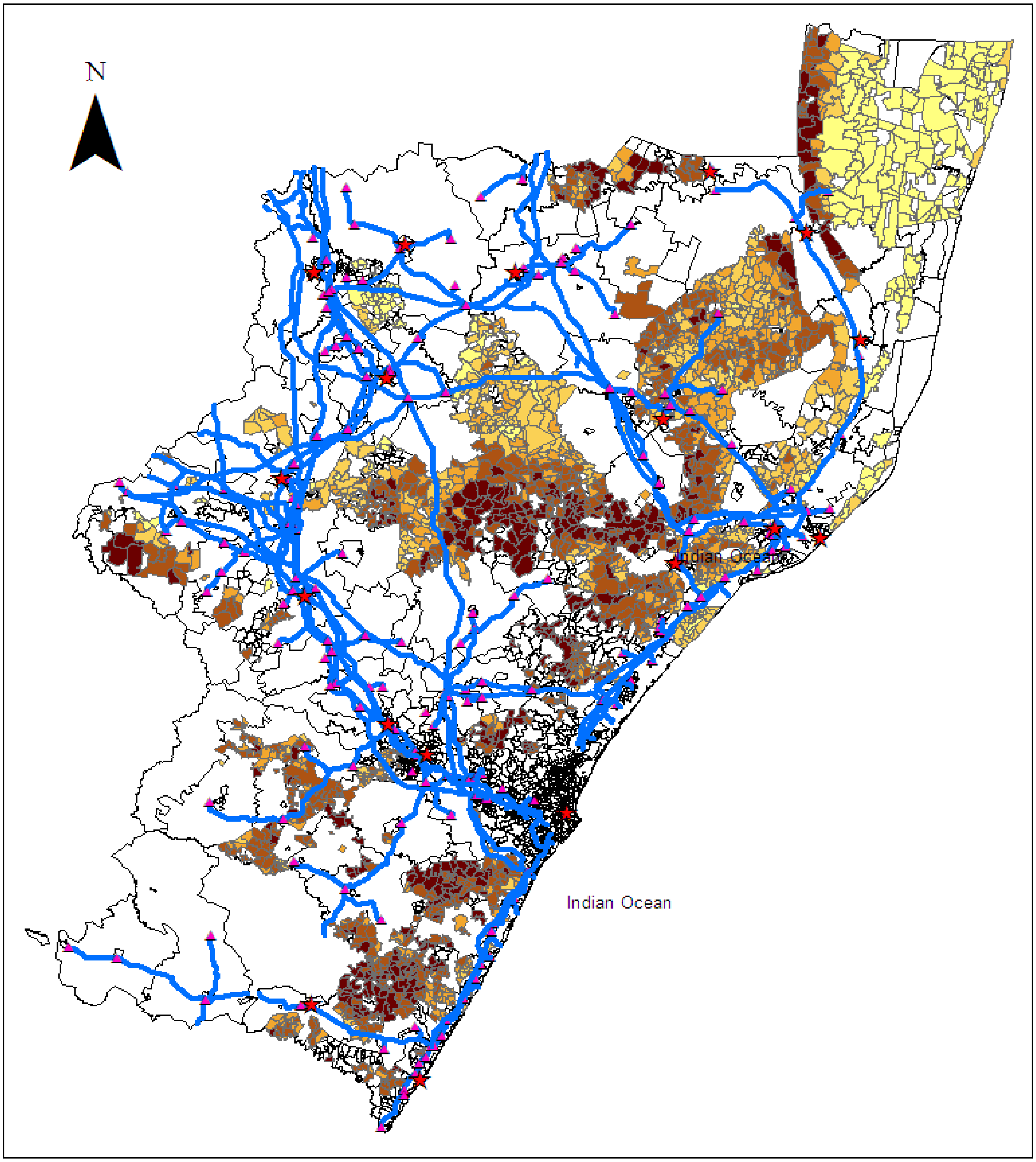
(2) Robust standard errors in parentheses, clustered at sub-district level. Significant at p<0.01***, p<0.05** or p<0.1* level.

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

(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. The change in the fraction of households with water close by is measured between 1996 and 2001, as is the change in fraction of households with flush toilets. Ten district fixed-effects are included in each regression.

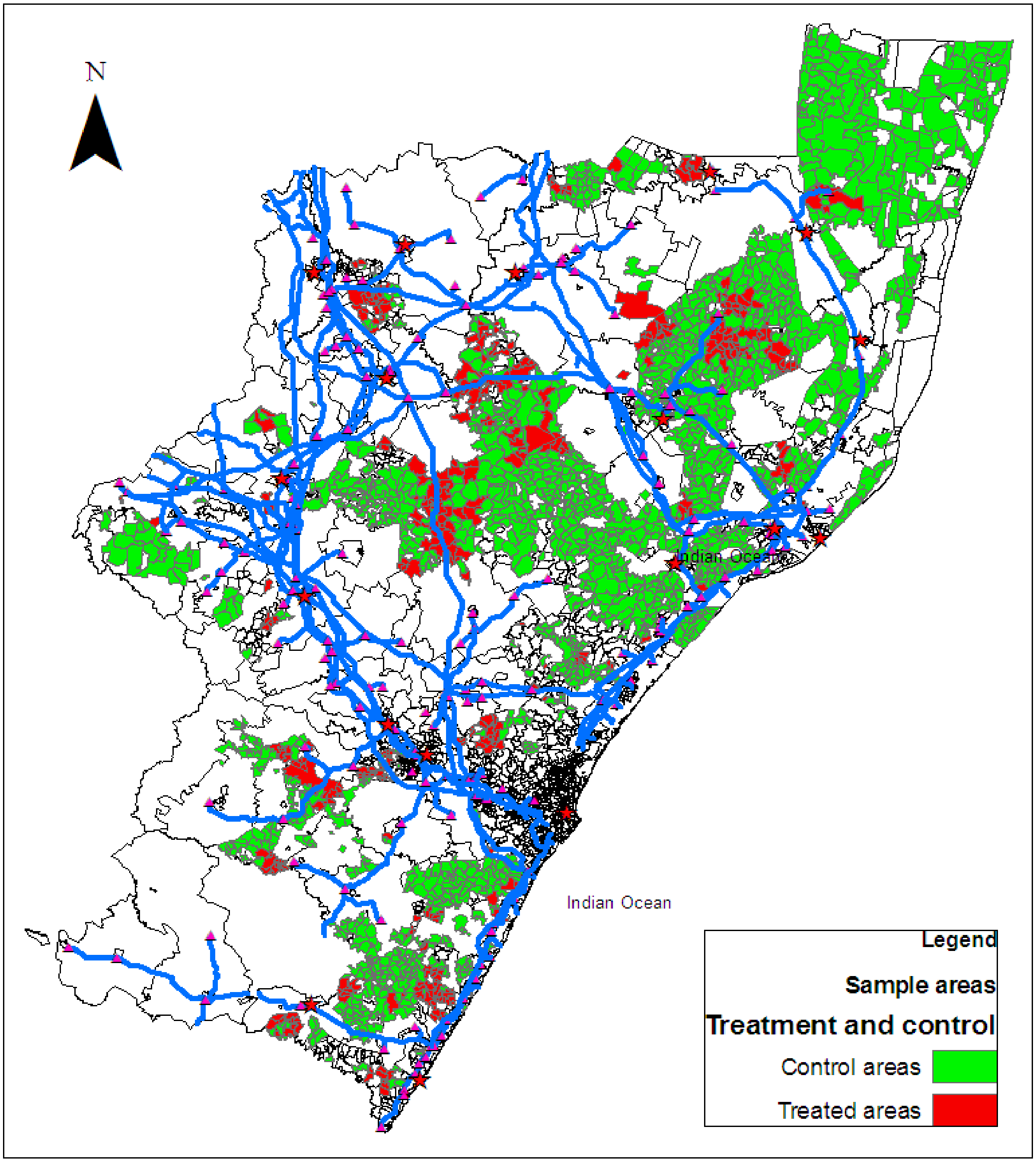
(5) The mean number of adult female (male) in-migrants in 1996 is 15.02 (10.07), in 2001 is 11.56 (8.60).

Figure 1: Spatial distribution of gradient in sample areas: KwaZulu-Natal, South Africa



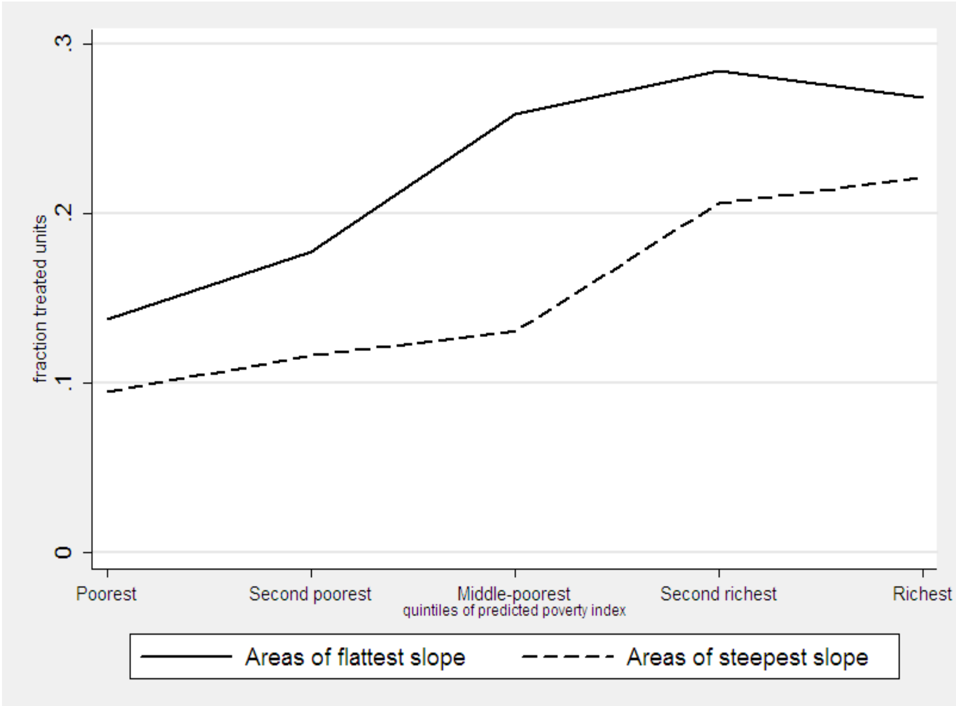
All shaded areas are included in sample. Steeper areas are shaded dark, flatter areas are shaded light (see electronic version for color). Lines depict electricity grid lines in 1996, triangles are electricity substations in 1996 and stars represent towns. N=1,992.

Figure 2: Treated and control areas: KwaZulu-Natal, South Africa



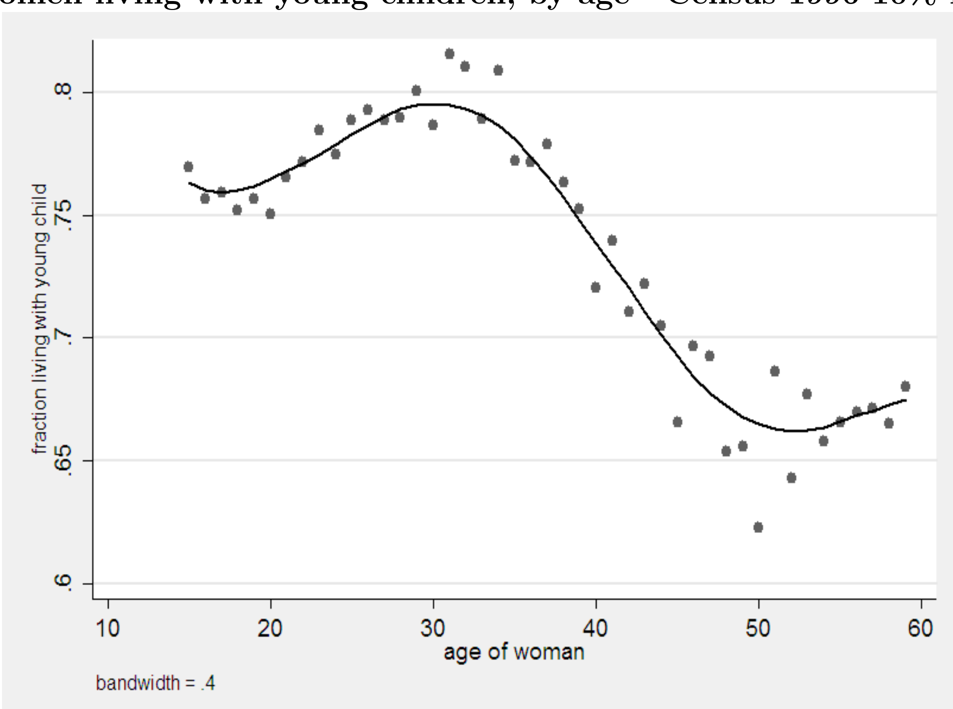
Shaded areas are in the sample: dark-shaded areas are treated with an Eskom project between 1996 and 2001, lighter shaded areas are treated after 2001 or not at all (see electronic version for color). Lines represent electricity grid lines in 1996, triangles are electricity substations in 1996 and stars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

Figure 3: Effect of gradient on treatment probability, by predicted poverty quintile



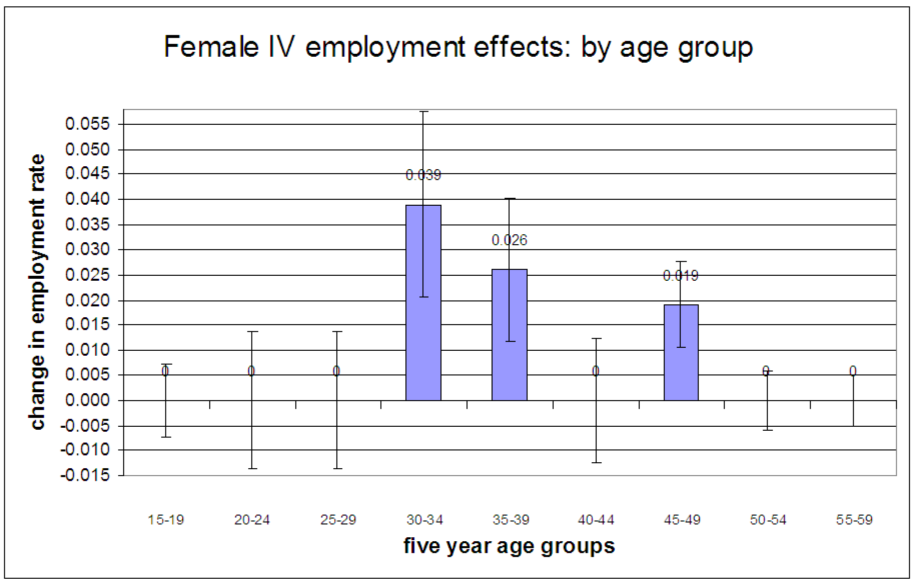
Lines show fraction of each predicted poverty quintile that is treated, by top (steep) and bottom (flat) halves of the gradient distribution. See notes for Table 10 for a description of how poverty index is created. The gap between the two lines indicates at which part of the poverty index the gradient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

Figure 4: Women living with young children, by age - Census 1996 10% micro sample



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

Figure 5: Effects of electrification on female employment, by age group



IV coefficients and standard error bars from IV regressions of change in female employment rate on treatment indicator, controlling for all other variables. Average community land gradient is the excluded variable.