

# Online Appendix

## The Arrival of Fast Internet and Employment in Africa

By JONAS HJORT AND JONAS POULSEN

### ONLINE APPENDIX

#### A1. *World Bank Enterprise Survey*

In Online Appendix Table A2, we use data from the World Bank Enterprise Surveys (WBES) to explore how the arrival of the submarine Internet cables changed the behavior and performance of firms in Ghana, Kenya, Mauritania, Nigeria, Senegal, and Tanzania, as discussed in Section V.

Almost all the firms in the WBES sample are located in relatively large cities and towns, and firms' location is reported only at city/town level. We thus classify a city/town as connected if the backbone network passes through its perimeter.<sup>1</sup> We include country  $\times$  year, location, and either industry or industry  $\times$  year fixed effects in the regressions and thereby approximate (1) as closely as we can with the WBES dataset.

In light of the WBES firms being clustered in space and the approximate connectivity classification required due to the relatively aggregate geographical level at which their location is reported, we view the WBES results as more suggestive than the rest of our analysis.<sup>2</sup>

The WBES results are shown in Online Appendix Table A2. Odd columns include industry fixed effects, while even columns include industry  $\times$  year fixed effects. In the top panel, we restrict the sample to firms located within 10 km distance from the backbone of terrestrial cables, as in the rest of our analysis. However, since this approach leaves us with only 20 clusters in the WBES sample, the bottom panel shows results from the full WBES sample.

As seen in columns 1 and 2, our estimates suggest that access to fast Internet leads African firms in the WBES sample to employ about 14-17 percent more workers per firm.

The results in column 3 indicate that firms are more than twice as likely to provide on-the-job training to their employees when fast Internet becomes available. This result is somewhat sensitive to the sample restriction made in the top panel: in the full sample

<sup>1</sup>Note that this yields the same classification of connected/unconnected cities/towns as "extending" the perimeter by 500 meters when categorizing locations.

<sup>2</sup>In the WBES sample, perhaps because of the approximate, between-city approach used to assign connectivity status, raw baseline differences between the dependent variables in connected and unconnected locations are larger than in our other samples. For example, WBES firms classified as connected employ significantly more workers (33 employees) on average than unconnected firms (14 employees) at baseline. While our difference-in-differences approach controls for level differences across locations, this calls for more caution when interpreting the WBES results.

results in the bottom panel, the estimated increase in on-the-job training is smaller and statistically insignificant, but nevertheless positive and of meaningful magnitude.

We also find a large (but imprecisely estimated) increase in WBES firms' sales per unit of labor costs—a measure of productivity—when fast Internet becomes available, as seen in columns 5 and 6.

In columns 7 and 8, we show evidence of a large increases in direct exports when fast Internet arrives.<sup>3</sup>

In columns 9-12, we document a significant increase in firms' probability of communicating with clients through email and through a website of about 11-15 percent, when fast Internet arrives.

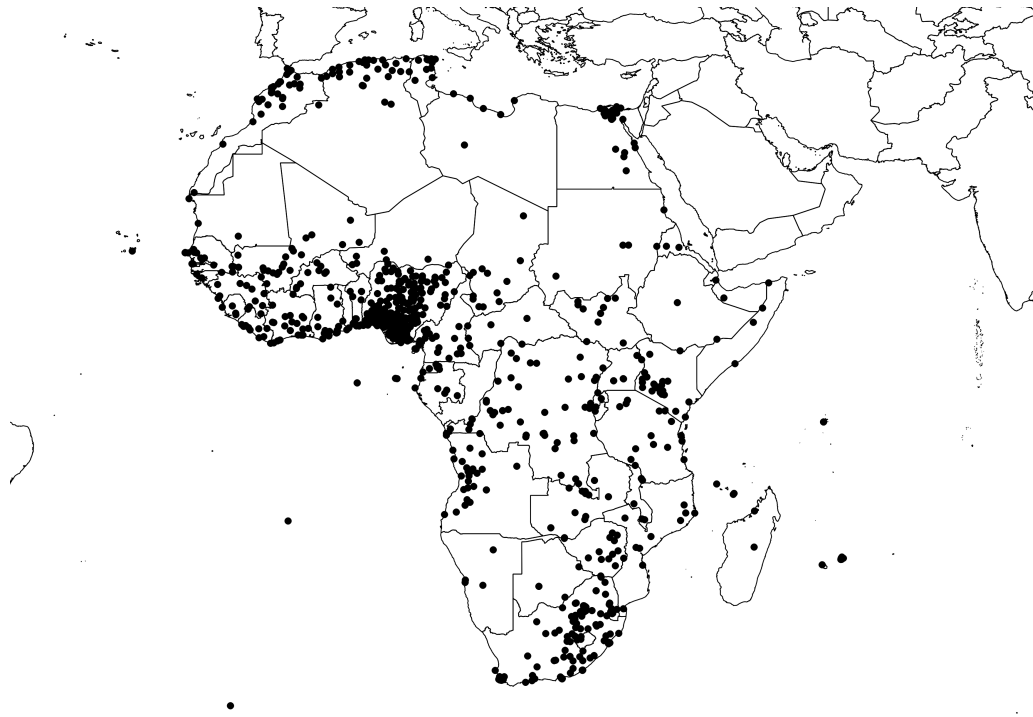
#### *A2. Categorization of South African Firms*

The South Africa Companies and Intellectual Property Commission's administrative dataset of firm registrations categorizes firms only as public, private or NGO. To estimate the effect of fast Internet on sector specific firm creation, we thus need to categorize the firms based on their names. We choose a set of sectors that, as far as possible, correspond to the categories of occupations reported in the DHS and QLFS datasets: services, white collar firms, blue collar firms, retail and sales, technology, manufacturing, finance, agriculture, mining, and tourism and foods.

We start by splitting all firm names on the space between the words. For example, "The South African Mining Company" would render five new variables. The first one would be called "name1" and would have the value "The", the second one would be called "name2" and would have the value "South", and so on.

name2, as specified above, usually contains the important key word that makes it possible to categorize a firm, such as "accounting", "trading", "properties", or "catering". We therefore use name2 to manually match key words to the sectors listed above. For example, we categorize firm X as belonging to the technological sector if name2 contains the words "computer", "tech", "telecommunications", and so on. The mapping of words to sectors is provided in Online Appendix Table A4. Of course, the key words could have ended up in any of the other name variables. We thus run the exact same iteration for all of those, up to name10. With this method, we are able to categorize 67 percent of the firms in the CIPC sample.

<sup>3</sup>The magnitude of the increase is again somewhat smaller in the full sample in the bottom panel.



**FIGURE A1. : LOCATIONS INCLUDED IN DATASET ON INTERNET SPEEDS**

*Note:* This graph plots the locations for which Akamai has information on Internet speed between 2007 and 2014. There are around 920 locations in this sample. We use data for the countries listed in Online Appendix Table [A1](#).

**TABLE A1—: INCLUDED COUNTRIES, SURVEY ROUNDS, AND CABLE CONNECTION**

<b>Country</b>	<b>Survey</b>	<b>First connected</b>
Benin	Afrobarometer (2008, 2011); DHS (2001, 2012)	2010/Q3
D.R. Congo	DHS (2007, 2013)	2012/Q2
Ghana	Afrobarometer (2008, 2012); DHS (2008, 2014)	2010/Q3
Kenya	Afrobarometer (2008, 2011); DHS (2008, 2014)	2009/Q3
Madagascar	Afrobarometer (2008, 2013)	2010/Q3
Mozambique	Afrobarometer (2008, 2012)	2009/Q3
Namibia	DHS (2006, 2013)	2012/Q2
Nigeria	Afrobarometer (2008, 2012); DHS (2008, 2013)	2010/Q3
Senegal	Afrobarometer (2008, 2013)	2010/Q3
South Africa	Afrobarometer (2008, 2011); SA-QLFS (2008/Q1 - 2010/Q2)	2009/Q3
Tanzania	Afrobarometer (2008, 2012); DHS (1999, 2010)	2009/Q3
Togo	DHS (1998, 2013)	2010/Q3

This table shows the survey rounds and connection quarters for countries included in our main analysis. Included countries (1) have surveys conducted both before and after cable arrival; (2) have detailed geographical markers; (3) have employment status for individuals; and (4) are not landlocked. We only make use of *Standard DHS* surveys.

**TABLE A2—: FAST INTERNET, EMPLOYMENT, EXPORTING, ON-THE-JOB TRAINING, SALES AND FIRM COMMUNICATION**

Outcome:	Employment				Sales		Trade		Internet use			
	Emplo- yees (asinh)		On-the -job train- ing (0/1)		Sales per labor cost (asinh)		Direct exports (asinh)		Email commu- nica- tion (0/1)		Website use (0/1)	
Unit of analysis:	Firm											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<b>Panel A:</b>												
Within-10-km restriction												
<b>SubmarineCables × Connected</b>	0.172 (0.073)	0.169 (0.100)	0.145 (0.046)	0.134 (0.044)	0.247 (0.546)	0.372 (0.572)	1.539 (0.684)	1.526 (0.496)	0.131 (0.053)	0.143 (0.042)	0.130 (0.074)	0.135 (0.059)
Observations	4897	4897	5164	5164	3666	3666	4036	4036	5132	5132	5128	5128
Mean of Outcome			0.09	0.09					0.38	0.38	0.12	0.12
<b>Panel B:</b>												
Full sample												
<b>SubmarineCables × Connected</b>	0.157 (0.063)	0.143 (0.060)	0.081 (0.051)	0.075 (0.048)	0.260 (0.439)	0.289 (0.429)	1.015 (0.707)	1.083 (0.603)	0.151 (0.039)	0.153 (0.038)	0.111 (0.049)	0.109 (0.048)
Observations	7230	7230	7543	7543	4935	4935	5589	5589	7494	7494	7488	7488
Mean of Outcome			0.09	0.09					0.38	0.38	0.12	0.12
<b>Country × Time FE</b>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<b>Location FE</b>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<b>Industry FE</b>	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
<b>Industry × Time FE</b>	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Data from World Bank Enterprise Surveys. WBES have produced a dataset accompanying the full WBES dataset that includes cleaned versions of sales variables and the number of employees. The authors account for outliers and deflate sales to 2009 USD. We use of the accompanying dataset for these variables. The WBES includes firms from Ghana, Kenya, Mauritania, Nigeria, Senegal, and Tanzania. 2006-2014. Time FEs are years. Location FEs are cities. The trade outcome is defined (by WBES) as the share of annual sales, and we interact that variable with annual sales to get the value of exports in 2009 USD. The Internet use variables are equal to one if the firm states that it uses email or a website to communicate with clients. Because almost all firms in WBES are located in large cities and sub-city information on their location is not reported, a within-500-meters-of-the-backbone indicator incorrectly classifies many connected WBES firms/cities as unconnected (for example those in Abuja, Accra, Dar es Salaam, and Kano. That these cities appears to be connected to the backbone is indicated by the fact that the maps show backbone cables whose route seemingly was chosen to connect them). In WBES we thus classify firms/cities as connected if the backbone passes through the city limits, as classified in Schneider, A., M. A. Friedl and D. Potere (2009). (This issue is not relevant to our other datasets, where we have much more fine-grained information on individuals and firms' location). Robust standard errors clustered at city level in parentheses.

TABLE A3—: ALTERNATIVE STANDARD ERRORS

Outcome:	Employment (0/1)		
Unit of analysis:	Individual		
Sample:	DHS	Afro- barometer	SA-QLFS
	(1)	(2)	(3)
<b>Panel A: Fast Internet and Employment</b>			
Standard errors accounting for spatial correlation			
<b>SubmarineCables × Connected</b>	0.046 [0.014]	0.077 [0.029]	0.022 [0.005]
Observations	59914	7918	280641
Mean of Outcome	0.68	0.58	0.72
<b>Country × Time FE</b>	Yes	Yes	No
<b>Grid-cell × Connected FE</b>	Yes	Yes	No
<b>Time FE</b>	No	No	Yes
<b>Location FE</b>	No	No	Yes
<b>Panel B: Fast Internet and Employment</b>			
Standard errors clustered on administrative area			
<b>SubmarineCables × Connected</b>	0.032 [0.017]	0.077 [0.045]	0.022 [0.008]
Observations	40798	7918	280641
Mean of Outcome	0.63	0.58	0.72
<b>Country × Time FE</b>	Yes	Yes	No
<b>Grid-cell × Connected FE</b>	Yes	Yes	No
<b>Time FE</b>	No	No	Yes
<b>Location FE</b>	No	No	Yes
Outcome:	Productivity		
Unit of analysis:	Firm		
Sample:	Ethiopian LMMIS		
	(1)		
<b>Panel C: Fast Internet and Firm Productivity</b>			
Bootstrapped Standard Errors			
<b>SubmarineCables × Connected</b>	0.127 [0.068]		
Observations	4321		
<b>Grid-cell × Connected FE</b>	Yes		
<b>Industry × Time FE</b>	Yes		

The DHS sample includes Benin, D.R. Congo, Ghana, Kenya, Namibia, Nigeria, Tanzania, and Togo. The Afrobarometer sample includes Benin, Ghana, Kenya, Madagascar, Mozambique, Nigeria, Senegal, South Africa, and Tanzania. The QLFS survey is from South Africa. Survey years for each DHS and Afrobarometer country are reported in Online Appendix Table A1. QLFS data are 2008/Q1-2010/Q2. Ethiopian LMMIS manufacturing firm census for 2006-2013. Grid-cells are  $0.1 \times 0.1$  decimal degrees, which is roughly  $10 \times 10$  km. Location FEs are enumeration areas in South Africa QLFS. Time is quarters in QLFS and years in DHS, Afrobarometer, and Ethiopian LMMIS. Individuals (locations) are considered connected if they are closer than 0.5 km to the backbone network. This table presents the precision of our main results using Conley standard errors that accounts for spatial auto-correlation in Panel A (?). The cut-off window is 100 km, but the results are virtually unchanged for 50 km, 500 km, and 1,000 km. Panel B clusters the standard errors on a region (DHS and Afrobarometer) and municipality (SA-QLFS, since South Africa does not formally have regions, and there are only nine provinces) level. The estimates and number of observations differ somewhat due to missing information on regions in the DHS data. There are 95 distinct regions in the DHS sample, 154 distinct regions in the Afrobarometer sample, and 341 distinct municipalities in the SA-QLFS sample. Panel C bootstraps the standard error in the last procedure described in Section 6.2.2.

**TABLE A4—: CATEGORIZATION OF SOUTH AFRICAN FIRMS' SECTOR BASED ON KEYWORDS FROM FIRM NAMES**

Services:									
abbatoir	accommodation	acquisitions	advertising	air-conditioning	aircon	airconditioning	airlines	airport	answers
apartments	art	arts	assessing	assessment	assessments	assist	assistance	aviation	baby
bakery	bar	bars	bay	bistro	block	booking	boxing	burial	cafe
care	carriers	carwash	caterers	catering	caterings	chauffeur	child	children	childrens
christian	church	clinic	clinical	clinics	club	coaches	coaching	coffee	coffees
college	commercial	communication	communications	compliance	concerts	conditioning	conference	conferences	consultation
cooking	cosmetics	cottages	council	councillors	courier	couriers	couture	croche	cuisine
cure	dance	daycare	deliveries	design	designs	destination	destinations	dienste	driver
driving	drycleaners	eatery	eat	educare	education	empowerment	entertainment	estate	estates
event	events	exhibitions	export	exporters	exporting	exports	facilitators	fashion	fast
feed	fellowship	fencing	fence	fitness	flats	food	franchising	funeral	funerals
games	gaming	gardening	god	gourmet	guest	guesthouse	gym	habitat	hair
hairstresser	haven	healing	health	healthcare	hire	hiring	homecare	homecleaners	homes
homestead	hospital	hospitality	host	hosting	hotel	hotels	housing	house	immobiliare
implementation	import	importers	imports	information	innovations	inspections	installations	intelligence	islam
kaife	karate	kontraakteurs	laundry	leadership	learning	leasing	leisure	living	lodge
logistic	logistics	logistix	mail	makelaars	mansions	meals	memorials	migration	ministries
ministry	mission	missions	monitoring	motel	networking	nominees	nursery	nursing	orchards
outsource	paintball	park	parking	payment	payments	pet	planning	pools	pre-school
procurement	promotions	properties	property	protection	realty	recruit	recruiting	recruitment	recruits
recycling	rent	rental	rentals	residence	residential	residential	resourcing	restaurant	restaurants
restaurant	retirement	retreat	retreat	rugby	safety	salon	sanctuary	school	security
seminars	service	services	servicing	shaving	shipping	shuttle	shuttles	spa	sport
sports	sportsbar	staffing	storage	strategies	strategy	supervision	surgical	swimming	tavern
tax	taxi	theatre	theatres	tourism	tours	training	transport	transformation	transformers
transportation	transporters	transporting	transporting	travel	travelling	travelling	travels	trucks	trucking
tutoring	undertakers	venue	village	village	villages	villas	wash	waste	wedding
weddings	wellbeing	wellness	wellness	wellness	wellness	wellness	wellness	wellness	wellness
White collar:									
academy	admin	administrasie	administration	administrators	advisors	advisory	akademie	analytics	architects
architects	architectura	architectural	architecture	assurance	attorneys	branding	chemical	chemical	chemicals
consult	consultancies	consultancy	consultant	consultants	consulting	data	decor	dental	designers
directors	editing	entrepreneurs	fashions	forensic	forensics	geoconsultants	ideas	institute	insurance
interior	interiors	konsultante	landscaping	legal	management	managers	managing	marketing	media
optometrist	optometrists	petrochemicals	publication	publishers	publishing	publishings	radiology	research	risk
translation	translation	translation	translation	translation	translation	translation	translation	translation	translation
Blue collar:									
autobody	automobile	baking	blocks	brick	brickforce	bricks	build	builders	building
butchery	car	carpenter	carpentry	clean	cleaners	cleaning	coaters	coating	coatings
concrete	construction	constructions	constructors	contracting	contractors	deco	distribution	distributions	distributor
distributors	drilling	equipment	fabrics	filtration	fishing	flooring	forestry	foundry	freight
fuel	gas	hunters	hunting	installation	installations	irrigation	knitting	konstruksie	laminations
laws	maintenance	mechanics	metal	materials	movors	packing	paint	painters	painting
painting	pavers	pavers	paving	plastering	plumbers	plumbing	recyclers	refrigeration	refurbishing
renovations	repair	repairs	roofing	scaffolding	textiles	towing	truck	woodwork	wholesale
Retail and sales:									
accessories	aesthetics	affairs	alarms	apparel	appliances	auction	auctioneers	auctions	auto
bags	bathrooms	beads	beautique	beauty	bedding	beverages	books	booksellers	bookshop
boutique	brand	brands	bread	canopies	cement	ceramics	cheese	clothing	commerce
commodities	components	condoms	confectionary	confectionery	consumables	cooling	cosmetic	cosmetics	covers
crafts	creations	cupboards	curtain	curtains	dealer	deals	delights	delighting	detailing
diary	discounters	elegance	enterprise	enterprises	enterprises	films	feeds	feeds	footwear
fuels	furnishers	furniture	furnitures	garden	garments	gift	gifting	gifts	glass
goodies	goods	groceries	handelaars	hardware	heating	herbs	hygiene	hygiene	images
ingredients	instruments	jewellers	jewellery	juice	kiosk	kitchen	kitchens	kitchenware	leather
lifestyle	lighting	lights	linen	liquor	liquors	lubricants	machinery	machines	mail
market	markets	mart	materials	meat	meats	medical	medicine	merchandise	merchandising
merchant	mini-supermarket	motor	motorcycles	motors	movies	music	newspaper	noodles	nutrition
oil	oils	optical	opticals	outdoor	outfitter	outfitters	parts	patisserie	pawn
pearls	perfumes	petroleum	pharmaceutica	pharmaceutical	pharmaceuticals	pharmacy	photography	photos	pictures
plants	plants	plates	print	printers	printing	printings	prints	produce	pumps
refrigeration	remedies	resale	retail	retailers	retailing	retailing	sewing	shoes	shop
shopfitters	shopping	signs	snacks	spices	sportswear	stationary	stationers	stationery	store
style	suit	superette	supermarket	superstore	supplements	supplier	suppliers	supplies	supply
sweets	telesales	things	timber	timbers	toiletries	tools	towels	toiletries	trailers
tyre	tyres	upholsters	upholstery	vehicle	vehicles	vehicles	wear	wholesalers	wholesale
wholesaler	wholesalers	windscreen	windcreens	wine	wines	wines	wines	wines	wines
Technology:									
3d	apps	audio	cable	cables	cabling	cellular	computer	computers	computing
digital	electrical	electricals	electrician	electricians	electronics	energy	engineering	engineering	engineers
fi	electricals	infra	internet	internet	mechanical	mobile	multimedia	online	software
solar	tech	technical	technicians	technicians	technik	technique	techniques	technological	technologies
technology	telecommunication	telecommunications	telecoms	telecoms	web	website	websites	websites	websites
Manufacturing:									
brewery	brewing	fabrication	factory	manufacturer	manufacturers	manufactures	manufacturing	plastic	plastics
production	productions	products	refineries	refiners	refinery	refining	refining	refining	refining
Financial:									
accountant	accountants	accounting	asset	audit	auditors	beleggings	bonds	bookkeeping	brokers
capital	cash	clearing	credit	credits	debt	equities	equity	finance	finances
financial	finans	finansiële	fund	funding	futures	holding	holdings	invest	investments
investing	investment	investments	investors	lending	loan	loans	money	mortgage	mortgages
portfolio	portfolios	savings	securities	trade	traders	trading	tradings	trust	wealth
Agriculture:									
agri	agri-business	agribusiness	agricultural	agriculture	boerdery	boerdery	farm	farmers	farming
farms	fisheries	growers	horticultural	horticulture	livestock	poultry	farm	farmers	farming
Mining:									
alu	aluminium	aluminium	coal	copper	diamond	diamonds	gold	iron	mine
mineral	minerals	mines	mining	platinum	steel	steelworks	uranium	uranium	uranium
Tourism and foods:									
bagel	burgers	cakes	chicken	chickens	cookies	fish	fruit	fruits	pizza
resort	resorts	safari	safari's	safaris	tour	vacations	vacations	vacations	vacations

This table displays the key words used when assigning firms observed in the CIPC data to a sector.

**TABLE A5—: EMPLOYMENT AND SKILLED EMPLOYMENT BY EDUCATIONAL ATTAINMENT, BEFORE SUBMARINE CABLE ARRIVAL**

Outcome:	Employed (0/1)			Skilled (0/1)		
	Connected	Un-connected	All	Connected	Un-connected	All
<i>DHS</i>						
Not primary	0.78 [0.41]	0.80 [0.40]	0.80 [0.40]	0.66 [0.47]	0.61 [0.49]	0.62 [0.48]
Primary	0.69 [0.46]	0.72 [0.45]	0.72 [0.45]	0.53 [0.50]	0.56 [0.50]	0.55 [0.50]
Secondary	0.60 [0.49]	0.61 [0.49]	0.61 [0.49]	0.55 [0.50]	0.56 [0.50]	0.56 [0.50]
Higher	0.69 [0.46]	0.70 [0.46]	0.70 [0.46]	0.66 [0.48]	0.69 [0.46]	0.68 [0.47]
<i>Afrobarometer</i>						
Not primary	0.42 [0.49]	0.44 [0.50]	0.43 [0.50]			
Primary	0.52 [0.50]	0.54 [0.50]	0.53 [0.50]			
Secondary	0.60 [0.49]	0.68 [0.47]	0.66 [0.47]			
Higher	0.79 [0.41]	0.79 [0.41]	0.79 [0.41]			
<i>SA-QLFS</i>						
Not primary	0.74 [0.44]	0.73 [0.44]	0.73 [0.44]	0.28 [0.45]	0.28 [0.45]	0.28 [0.45]
Primary	0.73 [0.45]	0.65 [0.48]	0.66 [0.47]	0.44 [0.50]	0.38 [0.49]	0.39 [0.49]
Secondary	0.76 [0.43]	0.71 [0.45]	0.72 [0.45]	0.65 [0.48]	0.60 [0.49]	0.60 [0.49]
Higher	0.91 [0.29]	0.90 [0.30]	0.90 [0.29]	0.88 [0.32]	0.88 [0.33]	0.88 [0.33]

All measures displayed are from the period before submarine cable arrival. Employment rates are from Demographic Health Surveys (DHS), Afrobarometer, and South African Quarterly Labor Force Surveys (QLFS). Occupational skill levels in DHS and QLFS are defined according to ILO ISCO standards. Individuals (locations) are considered connected if they are closer than 0.5 km to the backbone network. Standard deviations are shown in square brackets.