Essays on Development Economics

Zhe Liu

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington
2023

Reading Committee:
Rachel M. Heath, Chair
Alan Griffith
Xu Tan

Program Authorized to Offer Degree:
Economics

University of Washington

Abstract

Essays on Development Economics

Zhe Liu

Chair of the Supervisory Committee:

Associate Professor Rachel M. Heath
Economics

This dissertation consists of three chapters that study labor market frictions, gender norms, and school management practices in developing countries.

In the first chapter, I study how expansion of fast Internet availability affects job outcomes and the extent to which online job information can substitute for social networks. I use a two-way fixed effects identification strategy with continuous treatment at district level, and find that Internet availability has a positive impact on average employment and total income. Jobseekers are more inclined to search for job information online with increased access, while their reliance on social networks remains unchanged. The study also finds that young workers tend to search more through both online and network channels, suggesting that personal connections could complement internet job searching for some individuals. Workers without a primary education are discouraged from searching online and have worse employment outcomes. Constraints on effective uses of Internet job search and Internet activities, such as social networking, could help explain the results. To study the general equilibrium effect of fast internet on the labor market, I use search and matching theory and simulate the Internet access shock. The results show that fast internet can reduce job

search costs for workers and increase matching efficiency for firms, leading to higher wages. The impact on employment depends on the relative importance of these two forces.

In the second chapter, my coauthors Rachel Heath, Alex Phillip, and I examine the relationships between social norms and labor market competitions in India. Studies have shown that social norms have the potential to shape labor market equilibria. We test to what extent labor market conditions can alter social norms. In particular, we test whether men's support for women's work depends on the competition they face from women in their industries. We use labor market data from India to construct a measure of labor market competition that considers the industry percent female of average male worker in a given state and match this to attitudes on women's work from five waves of World Value Survey data spanning from 1990 to 2012. We find that men are more supportive of women's work when the overall female labor force participation is high, however, they are less supportive if more women work in their own industry.

The third chapter is a joint work with Natalia Cantet, Clara Delavallade, Alan Griffith, and Rebecca Thornton, and studies the effects of increasing community participation on school management outcomes. Poor school governance is a major contributing factor to low school results in developing countries. Policies to improve community-based governance are quite common but understudied. We present the results of a cluster-randomized trial of an intervention aimed at improving the functioning of School Management Committees — required in every public school — in rural India. We find large gains in governance activities after both one and two years of the intervention. We show modest improvements in school infrastructure as well as large effects on the number of teachers, and we present suggestive evidence that these improvements can be attributed to the increased committee activities.

Acknowledgements

The successful completion of this dissertation is a testament to the enormous support from many people.

First and foremost, my deepest appreciation goes to my remarkable advisor, Rachel Heath, whose insightful guidance, intellectual brilliance, and belief in my potential have been a constant source of inspiration for both research and life. I vividly recall the time when she hosted a seminar with her baby, illuminating my path as a mom pursuing an academic career.

I am indebted to members of my dissertation committee. Alan Griffith encouraged me to seize opportunities and to apply for workshops and conferences. Xu Tan made the economic theory less intimidating for me. The feedback from Yuya Takahashi and Jacob Vigdor greatly enhanced the quality of my work.

I extend my appreciation to the brilliant minds I met through the JSDE seminars. Emma Riley, who is always resourceful and a fun conference travel partner. Robert Garlic, who provided invaluable comments on my job market paper.

I acknowledge the financial support provided by the Grover and Creta Ensley Fellowship at the University of Washington. I am also grateful to all the people in the economics department who generously contributed their time and expertise to my job search: Melissa Knox, Seth Garz, Fahad Khalil, Jacques Lawarree, Yu-chin Chen, Yael Jacobs, Dennis O'Dea, and Heidi Hannah.

My lovely fellow PhD students, Camilo, Darwin, Diego, Ferdous, Herbert, Jackson, Jasmine, Lia, Monica, Phil, Sajid, Yue, Yida, and Zhou, provided endless encouragement through the highs and lows of this journey.

Lastly, this entire process would not have been possible without the unconditional support of my beloved family. My husband, who always believes in me and is a wonderful *baba* to our little boy. My son, whose tears and laughter distracted me from the anxieties of job hunting. My dad, who understood my decision to return to school, and my mom, who jumped in to help on so many occasions. *"It takes an exceptional village to raise a child and complete a PhD."*

Dedication

To my husband, Alex, and my son, Ian

和我的爸爸妈妈

Contents

List of Figures

List of Tables

Chapter 1

Who benefits from better Internet connectivity? Evidence from the labor market in South Africa

1.1 Introduction

A lack of information is arguably one of the key frictions in labor markets, and growing evidence has shown that information frictions can impede transitions into employment (Caria, Lessing and Hermes, 2019; Abebe et al., 2021; Carranza, Garlick and Orkin, 2020; McCasland and Hardy, forthcoming).Seminal works have done modeling the use of social networks by firms and jobseekers to overcome information frictions in job search (Montgomery, 1991; Calvó-Armengol and Jackson, 2004; Granovetter, 1973; Pellizzari, 2010). In many developing countries, social networks are especially important because it is the main or only information source for many individuals when making labor market decisions (Beaman, Keleher and Magruder, 2018; Caria, Franklin and Witte, 2020).[1] Referrals are also used as key methods for filling vacancies in these countries (Beaman and Magruder, 2012; Abel, Burger and Piraino, 2020; Heath, 2018). In recent years, the growth of Internet adoptions and expansion of job sites have lowered the cost of acquiring and disseminating job related information (Autor, 2001; Kuhn and Skuterud, 2004). Internet-based job search is by now one of the predominant ways of searching for jobs (Kuhn and Mansour, 2014). An important question is: does more access to Internet improve jobseekers labor market outcomes? If so, to what extent can a market mechanism like online job search and hiring, open to all and

[1] Caria, Franklin and Witte (2020) calculated that over 50 percent jobseekers in developing countries heard about current job from a social contact, based on the 2017 International Social Survey Program (ISSP) data by Sapin et al. (2020).

anonymous, substitute for exclusionary personal connections.

Existing studies solely focus on Internet impacts on job outcomes such as employment rate and income. My study aims to contribute by providing evidence on how the Internet may change job search activities. Individual's job search effort response have key implications for aggregate labor market outcomes. Understanding the choice of search channels is also critical for designing policies that can be used to address information frictions.[2] In this paper, I estimate how broadband Internet availability affects job outcomes for workers with different skill levels and age groups in South Africa, and how search methods used by workers and firms respond to faster and more Internet access. To the best of my knowledge, this is the first paper examining whether Internet access has an impact on the choice of job search methods.

I first present a simple model of jobseekers' utility maximization to show that search effort is the key to how employment changes with Internet access. Comparative statics predicts that if more Internet availability brings down the cost of search and increases the marginal productivity of search, a jobseeker will search more and has higher probability of being employed.

I match Internet connection data published in Hjort and Poulsen (2019), with spatially coded panel data of job search activities from South Africa, the National Income Dynamic Studies (NIDS), and compare individuals in locations with different Internet penetration rates, during the gradual roll out of first undersea cables in South Africa. This undersea cable brought much faster speed and traffic capacities. For example, South Africa's average download speed increased from 1,101kbps in January 2008 to 5,616kbps in June 2014.[3] The time required to load key job search websites decreased from 14 seconds to 5 seconds.[4] NIDS is the first and only national household panel survey in South Africa. Over 32,000 individuals across 52 districts from 2008 to 2014 are

[2]For example, training on using LinkedIn (Wheeler et al., 2022), reference letters from previous employers (Abel, Burger and Piraino, 2020), or detailed job search plan (Abel et al., 2019).

[3]Sources: Speedtest Global Index by Ookla, which uses data from millions of Internet speed tests performed on the Speedtest.net platform.

[4]Author's own calculation using `http://www.webpagetest.org`. The main job search website in South Africa `careers24.co.za` is tested.

included in the final data set.

I address the endogeneity issues in two ways. First, I use both location fixed effects and year fixed effects to explore the temporal and spatial variation in the Internet availability across 52 districts in South Africa. This identification is similar in spirit to a difference-in-difference (DID) design at district level with a continuous treatment. Variations in Internet treatment intensity make it possible to evaluate a "does-response" relationship, which policy makers may care more about than the effect of the existence of Internet. Recent literature shows that TWFE estimators may not be robust to heterogeneous treatment across groups and over time (Goodman-Bacon, 2021). Thus, I access the robustness using the estimator proposed by de Chaisemartin and D'Haultfœuille (2020), and find larger but imprecisely estimated impacts of Internet on job outcomes and searches. Second, I analyze the timing of changes in Internet access, and find that the timing does not appear to be systematically related to key observable correlated of employment. Lagged productivity variables such as employment rate, education level, age, and industry distributions do not predict current Internet penetration rates.

I find positive effects of Internet availability on employment and total income. A one-standard-deviation improvement in Internet availability (about 10 percentage points) increases the employment rate for an average jobseeker in the district by 3.6 percent, and increases his or her total income by 8 percent. As to search methods, more Internet availability induces jobseekers to use online job information by about 10 percent more, but more access does not change reliance on personal networks or government agencies. The total number of different search methods ever used by jobseekers declines, driven by less uses of other methods such as contacting other employers or waiting at the side of roads. These estimates are robust to inclusion of a set of time-varying controls for potential productivity factors, as well as allowing for different time trends across areas.

Heterogeneous analysis by age group and education attainment contribute to our understanding of distributional effects of Internet technology change. When more Internet becomes available in the area, both young (between 15 to 24 years old) and older workers increase their employment rates, but only older workers' total income increases. Young workers will use more search methods, and increase their searches through not only online but also personal networks. Compared with skilled workers with beyond primary education, unskilled workers are discouraged from online job search, less likely to be employed, and earn less.

Considering choices of search channels made by young and unskilled workers respond to more Internet differently than older and skilled workers, I provide additional evidence on other constraints preventing them from using this technology for effective job search. Computer ownership and computer literacy are low for less-educated workers. Internet activities like social networking can also help maintain relationships or form new links with a wide network of weak ties. Thus, personal networks could complement the Internet in the job search process for some workers. For example, young workers are more likely to own a computer and know how to use it. Besides using this technology to search for job information online directly, they also use it to enhance personal networks that could be used for sharing job information.

Since the Internet variations are at a district level, cheaper information brought by the Internet are available for both jobseekers and firms. The results on employment and income should reflect the equilibrium outcomes of both labor supply and labor demand. Without employers or firms' data, I cannot say much about how employers' job creation decisions respond to more Internet access empirically. Instead, I use the search and matching theory of unemployment and vacancies, the Diamond-Mortensen-Pissarides(DMP) model (Diamond, 1982; Mortensen, 1986; Pissarides, 1985). I simulate the Internet access shock by changing the parameter value of matching technology or the value of unemployment income, and numerically solve the new general equilibrium. The results show that Internet can change the equilibrium outcomes by decreasing the job search costs for workers, or increasing matching efficiency for firms. The effect on wage is unambiguously positive, but the aggregate effect on employment depends on the relative importance of these two forces.

Understanding the impacts of Internet on labor market in South Africa is salient and has important policy implications. Despite being Africa's most industrialized economy, South Africa still has extremely high levels of unemployment. The recent Covid-19 lockdown has pushed the unemployment rate to a record high above 30% in 2021.[5] In particular, unemployment has been inordinately high for young workers (Figure A1), who may have less access to referral networks and limited information about their employment prospects. The challenge for policymakers is to ensure that all current and future workers can seize the growing economic opportunities that accompany the spread of digital technologies.

This paper adds developing country evidence to a limited literature assessing the linkages between Internet and labor market outcomes. More than 60 per cent of the world's employed population earn their livelihoods in the informal economy, most of them in emerging and developing countries (Bonnet, Vanek and Chen, 2019). For many informal jobs in which workers' effort is important but difficult to induce, formal institutions for firms to share information about workers are absent (Heath, 2018). Given these labor market frictions, it remains questionable if findings for broadband Internet expansion implemented in developed countries are applicable to developing countries. To date, the only direct evidence on the average and distributional effects in developing countries is provided by Hjort and Poulsen (2019) focusing on the Africa continent. They leverage the gradual arrival of sub-marine Internet cables in Africa, and find large positive effects on employment and incomes, particularly for higher-skill occupations, due in part to the technology's impact on firm entry, productivity, and export.

Existing studies in high-income countries show mixed impacts. Kroft and Pope (2014) analyze the expansion of Craigslist in the US, and find that Craigslist significantly lowered classified job advertisements in newspapers, but had no effect on the unemployment rate. Dettling (2017) uses state-wide shares of multifamily residences to instrument for the diffusion of Internet access across the U.S., and finds increases in labor force participation rates of married women, and no corresponding effect for single women or men. Bhuller, Kostol and Vigtel (2019) document that

[5] Source: Statistics South Africa, 2021

broadband expansions in Norway increase online vacancy-postings, lower the average duration of a vacancy, resulting in higher job-finding rates and starting wages, and more stable employment relationships after an unemployment-spell. Akerman, Gaarder and Mogstad (2015) finds the same Internet expansion improves the labor market outcomes and productivity of skilled workers only. The stronger effects of Internet I found in this paper could suggest that limited information may exacerbate other labor markets frictions such as high migration costs or limited public transportation in sprawling cities in less-developed countries (Ardington, Case and Hosegood, 2009; Bryan, Chowdhury and Mobarak, 2014; Franklin, 2018).

My paper complements a growing experimental literature considering the role of limited information in labor market matches in developing countries. Abebe et al. (2021) shows that job application workshop for young jobseekers can help them signal skills better, and generate large and persistent improvements in their labor market outcomes. The effects are larger when combined with formal certificates provided to firms (Carranza, Garlick and Orkin, 2020). Firms may have poor knowledge of candidates, and providing information directly to firms can improve match quality (Banerjee and Chiplunkar, 2020; Abel, Burger and Piraino, 2020). Online platforms such as LinkedIn can help address supply-side information frictions by allowing jobseekers to learn more about job prospects, and also address demand-side frictions by allowing firms to learn more about potential candidates (Wheeler et al., 2022).

My findings on the distributional effects are at odds with the notion that active labor market programs such as training or employment subsidies have larger employment effects for more disadvantage groups(Card, Kluve and Weber, 2018).This could partly because the most disadvantaged group in my study may not have direct access to the Internet technology even if it is made more available in their areas. I provide evidence showing that when Internet becomes available, it is more likely adopted in places where complementary factors such as computer ownership and computer literacy are abundant. In addition, cheaper information are available at a larger scale, and both sides of the labor market respond to this information provision.

This paper also extends the literature on the role of information and communications technology (ICT) in developing countries. ICT such as mobile phones has been attributed with reducing price dispersion across markets and increasing welfare for producers and consumers (Jensen, 2007; Aker and Mbiti, 2010; Goyal, 2010). ICT such as mobile money can help reduce transaction costs and potentially improve informal risk sharing networks (Jack and Suri, 2014). ICT can even influence fertility patterns and bring cultural changes to the society (La Ferrara, Chong and Duryea, 2012). My paper shows that ICT such as Internet can provide cheaper access to job information directly, or reduce communication costs for sharing job information among family and friends, impact the job search methods used by jobseekers, and improve employment outcomes in the labor market.

The rest of the paper proceeds as follows. Section 1.2 provides a simple model analyzing how Internet availability affects jobseekers' search effort and employment. In Section 1.3 I present the data, and in Section 1.4 the estimation strategy. The average and heterogeneous results are in Section 1.5. In Section 1.6, I show additional evidence how Internet access may affect employment and search behavior. Section 1.7 concludes with policy implications.

1.2 Conceptual model

I present a simple model illustrating the relationship between employment, job search, and Internet access. With exogenously provided Internet access, how employment changes depends on the optimal search effort. If search cost decreases while marginal productivity of search increases with more Internet access, jobseekers will exert more effort, and are more likely to find a job.

A jobseeker lives two periods: in the first period, an unemployed individual receives some unemployment benefit b and decide how much effort to spend for job searching s. Cost of job search $\tau(\theta)$ depends on amount of Internet access, and the probability of finding a job depends on both the search effort and amount of Internet access: $p(s, \theta)$. In the second period, if the individual becomes employed, assuming labor supply is inelastic, a fixed income will be given as w. In this set up, internet access amount θ and wage w are given exogenously.[6] The jobseeker chooses job search effort and maximizes the expected lifetime utility as follows:

$$\begin{aligned}\max_{s} \quad & u\left(c_1\right)+\beta E u\left(c_2\right) \\ \text{s.t.} \quad & c_1=b-\tau(\theta) s \\ & c_2=\begin{cases} w & \text{w.p. } p(s, \theta) \\ b & \text{w.p. } 1-p(s, \theta)\end{cases} \\ & 0 \leq p(s, \theta) \leq 1\end{aligned} \tag{1.1}$$

where utility $u(c)$ is assumed to be increasing and strictly concave in consumption.

An interior solution should satisfy the following first order condition:

$$\tau(\theta) u^{\prime}=\beta \frac{\partial p(s, \theta)}{\partial s}[u(w)-u(b)] \tag{1.2}$$

which implies that the individual chooses search effort s optimally such that the marginal utility of giving up consumption equals the expected utility gain from searching for work, which is the difference between employment and unemployment utility in the second period.

For this paper, I am interested in how employment probability may change with the Internet access given exogeneously. The comparative statics is,

$$\frac{d}{d \theta} p(s(\theta), \theta)=\frac{\partial p}{\partial s} s^{\prime}(\theta)+\frac{\partial p}{\partial \theta} \tag{1.3}$$

Assuming the marginal productivity of search and Internet are both positive ($\frac{\partial p}{\partial s}, \frac{\partial p}{\partial \theta}>0$), the effect on employment will depend on $s^{\prime}(\theta)$. In order to see how optimal search effort $s^{\star}(\theta)$ changes with Internet access θ, we can differentiate the first order condition equation 1.2 with respect to θ:

[6] I include a search and matching model allowing wage to be determined endogenously in section 1.6.4. Simulation results show a positive impact of Internet on equilibrium wage.

$$s'(\theta) = \frac{\tau' u' - \beta p_{s\theta}(u^{emp} - u^{unemp})}{\tau u'' + \beta p_{ss}(u^{emp} - u^{unemp})} \tag{1.4}$$

where u^{emp}, u^{unemp} represent the utility being employed and unemployed in period 2 respectively.

Assume $u'' < 0$ and $p_{ss} < 0$, and u$^{emp} > u^{unemp}$ is a necessary condition for the existence of an interior solution, the denominator in equation 1.4 is negative. The sign of the numerator depends on two parts. First, $\tau'(\theta)$, the change in the cost of job search given more Internet access. If we think more Internet means that jobseekers have cheaper access to more job information, the cost of job search should be lower, $\tau'(\theta) < 0$. Second, $p_{s\theta}$ the change in the marginal productivity of search in response to more Internet access. $p_{s\theta} > 0$ if job search by the jobseekers is made more productive with more Internet, eg. Internet technology can help job candidate send out more resumes, or firms can screen candidates and match them with vacancies faster.[7] Then the numerator of equation 1.4 should be negative too. With positive change of optimal search effort($s'(\theta) > 0$), equation 1.3 indicates that employment will increase as well.

Internet can also change the utility of leisure, which will impact the trade off between searching for jobs and staying unemployed. I solve a version of this model including leisure in the jobseeker's utility function in appendix A. The comparative statics predictions are similar.

Using published data from a field experiment that Abel et al. (2019) have done with South Africa youth, I find suggestive evidence that online job search is correlated with higher effort exerted. The original experiment is to test the effects of plan making on job search and employment. Table 1.1 shows the regression results using panel data over two follow-up periods, with only baseline control group observations included. In a period of about 12 weeks, individuals who search jobs online spend 2 hours more, and send out 2.6 more number of applications in total. They are also more likely to receive responses and job offers from the employers.

[7] The marginal productivity of search p_s is not necessarily linear in θ.For example, too much information online can be a distraction from job searching, or ghosted postings can make searches a waste of time. Thus, the marginal productivity of search in response to more Internet access can be negative, $p_{s\theta} < 0$. Then the impact on optimal search effort and employment is unclear.

Table 1.1: Effects of Online Search on Search Behaviors and Employment Outcomes

	(1) Search Hours	(2) Applications	(3) Empl Responses	(4) Job Offers	(5) Employed
Search online	2.091	2.626***	0.535***	0.091***	0.016
	(1.300)	(0.320)	(0.061)	(0.026)	(0.031)
Mean Dep Vars	14.087	3.821	0.543	0.131	0.116
Obs	818	828	828	819	857
R-squared	0.026	0.079	0.048	0.026	0.011

* Notes: Baseline control group observations of Abel et al. (2019) are used. All specifications control for age, gender, education, round, and location fixed effects. Standard errors (in parentheses) are clustered at the individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.3 Data

In the main analysis, I use the South Africa National Income Dynamic Studies (NIDS) for labor market data. NIDS is the first and only national household panel survey in South Africa, and is implemented by the Southern Africa Labour and Development Research Unit (SALDRU) based at the University of Cape Town's School of Economics. The study began in 2008 with a nationally representative sample of over 28,000 individuals in 7,300 households across the country. Stratified random sampling was implemented, whereby 1500-3000 enumerator areas are randomly selected and subsequently 10 households per enumerator area are interviewed. The core survey continued to be repeated with these same household members every two years to three years, with the latest round being conducted in 2017. NIDS provides information about changes in broad themes, including poverty, education, health, household structure, labor market participation and economic activity, migration, and social capital.

I focus on the labor market module in the survey, where working age adults were asked about their labor market participation and economic activity, including employment status, income (wages or the profits of self-employed workers), contract types, and industry. In addition, individuals were asked to check all the job search methods used, including family and friends, online ads, government agency, previous employers, and others.

Table 1.2 provides the summary statistics of working age (15 - 65) individuals used in the analysis. This sample has a balanced representation of urban and rural population. 59 percent of the sample are female. Average worker's age is 33, and 37 percent are between 15 and 24 years old. 52 percent have finished primary education. Cellphone ownership is high (71 percent) compared to computer ownership (5 percent). About one third of the sample report they know how to use a computer. 37 percent of the sample are employed, among which 27 percent have a job paid with regular salary, and 5 percent are self employed. On average, individuals work around 40 hours per week, and earn 3233 ZAR(230 USD) per month. The standard deviation of log income is large, because I put zero for unemployed individuals' income. Network (25 percent) is the most widely used job searching method, while 6 percent of the sample report that they have used online search. Internet is available to 10 percent of the population in an average district with a standard deviation of 14 percentage points.

I use the Internet infrastructure and speed data published in Hjort and Poulsen (2019). Using Mahlknecht's map of submarine cables to measure landing points and times (Mahlknecht 2014), and www.africabandwidthmaps.com and AfTerFibre's (AfTerFibre 2014) maps of terrestrial backbone networks to measure locations' connectivity, Hjort and Poulsen (2019) document whether a city is connected to the Internet quarterly from 2007 to 2014. Average Internet speed for the same locations is also provided by network service company Akamai Technology.

I match this Internet connection data from Hjort and Poulsen (2019) with the NIDS survey data using the geocode and year. While the converge data are available at the city level, individuals in the NIDS outcomes data can only be identified at higher levels of geographies, such as province and district. I aggregate the city-level connection data to district-level by calculating the percentage of cities with connection in one district by year, weighted by its population. 52 districts across 4 waves from 2008 to 2014 are included in the final data set.

Figure 1.1 and 1.2 show the variation in percent of cities connected over time and across dis-

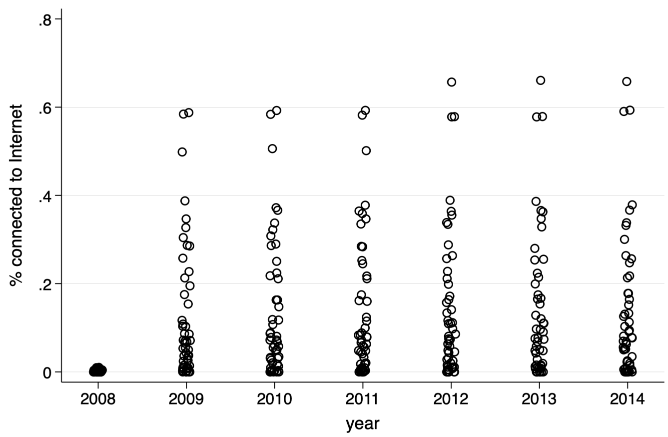
Table 1.2: Sample Summary Statistics

	Obs	Mean	SD
Individual characteristics			
Urban area	38,497	0.47	0.50
Age	38,520	33.49	14.29
Female	38,520	0.59	0.49
Youth(15-24)	38,520	0.37	0.48
Not primary	38,436	0.06	0.23
Primary	38,436	0.39	0.49
Secondary	38,436	0.29	0.45
Higher	38,436	0.26	0.44
Parents with primary education	24,156	0.19	0.39
Own a cellphone	35,311	0.71	0.46
Own a computer	35,301	0.05	0.22
Is computer literate	34,386	0.29	0.46
Household characteristics			
HH owns a cellphone	38,124	0.85	0.36
Spent money on cellphone monthly	29,108	0.74	0.44
HH owns a computer	38,069	0.11	0.31
Spent money on internet monthly	29,144	0.01	0.11
Labor market outcomes			
Employed	37,108	0.37	0.48
Salary job	35,654	0.27	0.44
Self employed	35,651	0.05	0.21
Total income (adjusted)	37,108	2.05	3.68
Salary income(adjusted)	34,696	3.69	6.57
Has permanent duration	9,307	0.54	0.50
Weekly hours	10,516	39.70	17.15
Job search methods			
Network	32,923	0.25	0.43
Online	32,923	0.06	0.24
Government	32,923	0.03	0.17
Others	32,923	0.15	0.36
Internet connection at district level			
% population connected	38,520	0.10	0.14

* Notes: Only workers between age 15 and 65 are included. The income for unemployed workers is adjusted as zero, and inverse hyperbolic sine is used for the log transformation.

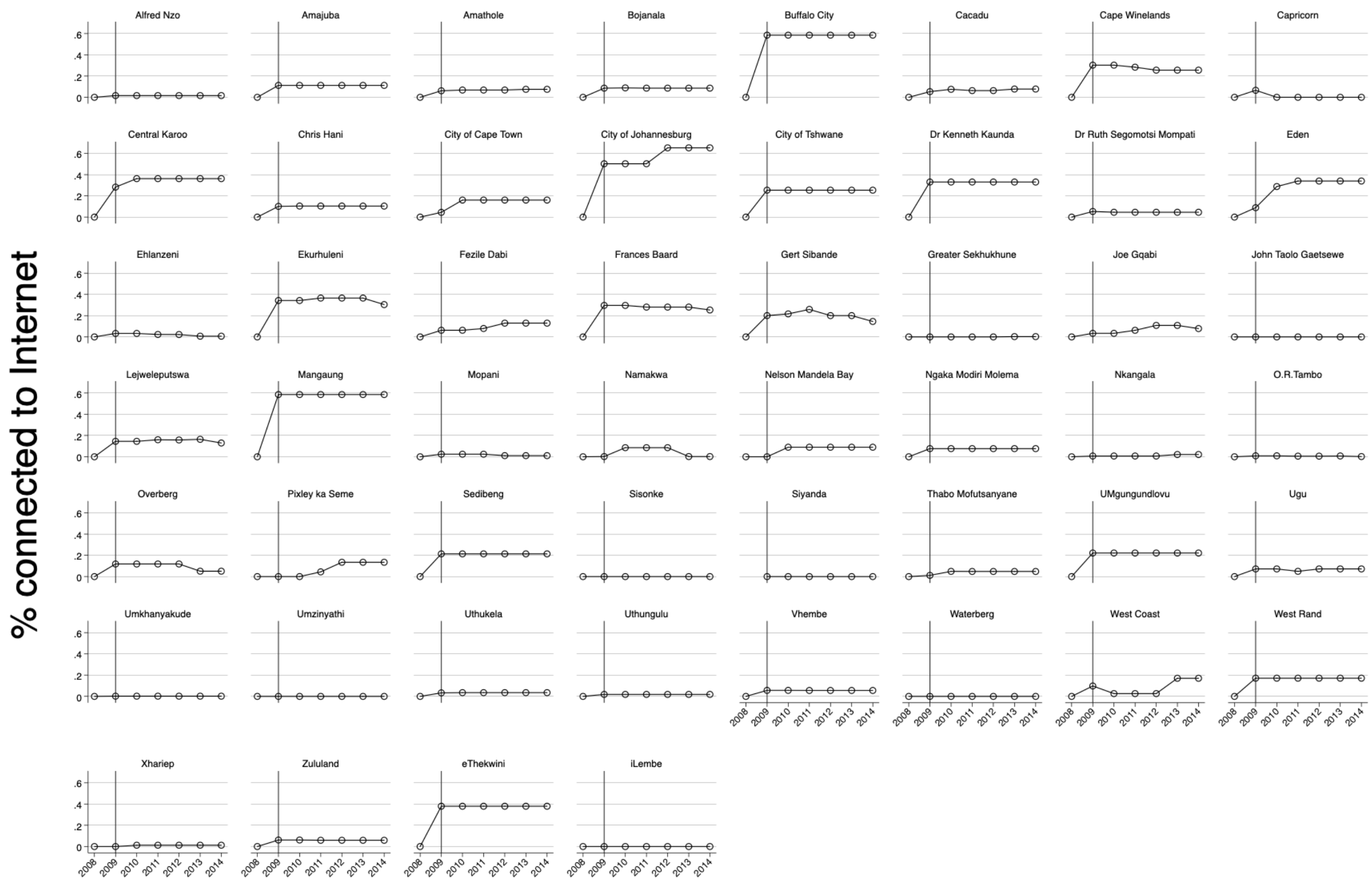
tricts. In 2008, all cities have no fast Internet connection. Over the years, more cities gained access and more districts achieved higher availability rates in 2014. There are also differences in connection timing and access intensity within districts, which generate a continuous measure of availability rates that I exploit as the key variations in my empirical analysis.

Figure 1.1: Comparing Internet availability rates over years



Notes: Each dot represents one single district.In 2008, the percent of populations connected to fast Internet are zero for all districts. The first fast Internet cable was connected in 2009.

Figure 1.2: Comparing Internet availability rates across districts



Notes: Each sub plot represents one district.The vertical line indicates the connection of first fast Internet cable in 2009. Some districts experienced drops in connectivity due to infrastructure deterioration, which could be caused by natural disasters or a lack of maintenance.

1.4 Empirical Strategy

My empirical approach is a two-way fixed effects (TWFE) estimation that controls for location and year fixed effects. I compare individuals across locations with varying degrees of Internet coverage, during the gradual roll out of undersea Internet cable in South Africa. This is motivated by two features of this Internet expansion. First, most of the confounding supply and demand factors are accounted for by the location fixed effects. Second, the timing of the expansion is unlikely to co-vary with key correlates of employment.

1.4.1 Two-way Fixed Effects estimation

I run the following two-way fixed effects estimation as the main specification.

$$Y_{ijt} = \beta PercentConnected_{jt} + X'_{ijt}\alpha + \gamma_t + \theta_j + \epsilon_{ijt} \tag{1.5}$$

where Y_{ijt} is the labor market outcomes for worker i in district j at time t. The set of outcomes of interest are individual-level labor market outcomes, including employment, employment with formal contracts, income, network search, and online search. $PercentConnected_{jt}$ is the percent of population in district j connected to the Internet at time t. This measure allows me to exploit variation within the set of connected districts in their intensity of treatment.

All specifications include both location fixed effects, θ_j, time fixed effects,γ_t, and an idiosyncratic error term, ϵ_{ijt}. X_{ijt} is a vector of individual-specific controls, including age,gender,and education level. Since there could be other unobserved individual-level factors that are endogenous to the choice of search channels, I also include a individual fixed effect in some analyses. In all analyses, standard errors are clustered at the district level.

Within such a set up, as long as there are not omitted idiosyncratic shocks correlated with both Internet connection rate and labor market outcomes, the causal effect of Internet, β, is identified off of comparison between the change in outcomes for locations that gain (more) access to Internet in a given year and the change in outcomes for other locations that without or gain less access at the same time.

Given that I am controlling for fixed effects for districts and years, the core of this design is similar to a difference-in-difference setup at the district level. Districts fixed effects act as controls for the "preperiod" outcomes of workers in the same area that never received Internet, which is the first difference. Treatment and control groups can be defined as workers within a year that had different exposure to Internet access. Comparing the treatment group outcomes from the control group yields the second difference. Because the treatment variable $PercentConnected$ is continuous, I effectively weight these double differences by the difference in Internet connection rates.

1.4.2 Tests for parallel trend assumption and timing of the expansion

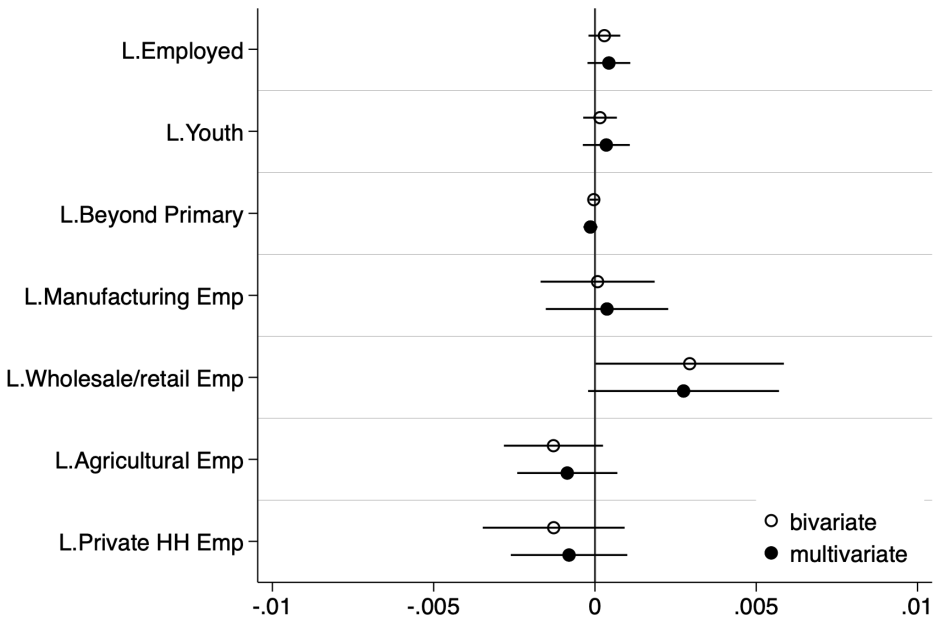
Another threat to identification is that the timing of the expansion might be related to different underlying trends across locations. The parallel trend assumption of difference-in-difference models requires that in the absence of treatment, the difference between the "treatment" and "control" group is constant over time. Two limitations prevent me from producing the standard parallel trend test. First, I only have one time period (2008) prior to the access boost in 2009. Second, since the variation in coverage rate is a continuous variable, I do not have a clear "control" and "treatment" group.

Instead, I test if current Internet allocation is determined by lagged productivity variables as follows:

$$PercentConnected_{jt} = \gamma_t + \theta_j + \lambda c_{j,t-1} + \epsilon_{jt} \tag{1.6}$$

where $PercentConnected_{jt}$ is the Internet availability for location j at time t, γ_t is year fixed effect, and θ_j is location fixed effect, and $c_{j,t-1}$ is district-level variables related to productivity at previous year, including employment rate, education level, percent of young workers, and industry distribution.

Figure 1.3: Internet connectivity and lagged productivity variables



Notes: The bivariate coefficients are from models regressing internet connectivity on a single lagged productivity variable. The multivariate coefficients are from a regression of internet connectivity on all productivity variables.

The coefficients plot of γ in Figure 1.3 shows that most lagged productivity variables do not predict Internet connectivity rates. Though the impact of wholesale/retail sector employment rate is significantly different from zero but small. So I also check that including it as an additional control in the main specifications does not change the estimation results.

1.5 Results

1.5.1 Main effects

In Table 1.3, I show the regression results for specification in 1.5, including district and year fixed effects and demographic controls. I find that one standard deviation increase in Internet connection (about 10 percentage point) increases the probability that an individual is employed by 1.3 percentage point, or 3.5 percent increase off a baseline of 36.6 percent average employment rate (column 1). This result is similar in magnitude to Hjort and Poulsen (2019), where they find the employment increases by 3.1 percent for South Africa in their cross-country sample of Africa countries.[8] Workers can earn an average of 8 percent more in total income when the Internet connectivity in their areas rise by 10 percentage point (column 2).[9]

To see what extent these increases reflect additional economic activity, I use more detailed work-related questions that only employed individuals were asked in the NIDS. Given the truncation by survey design, results in column 3-5 should not be interpreted as casual effects of Internet, but rather should be viewed as an intensive channel of the overall effects. For individuals already working, they will earn more while work less hours with additional Internet (column 3, 4). The estimated effect on having a formal contract is close to zero and insignificant (column 5). This helps rule out the situation that the additional employment comes from formalization of existing informal jobs.

To further explore how individuals job search behavior might change with Internet access, I show results on the search methods in Table 1.4.One standard deviation (about 10 percentage point) increase in Internet availability will induce jobseekers to look for information online by 0.64

[8]The results are not identical, because Hjort and Poulsen (2019) used a different labor force survey, the South Africa Quarterly Labor Force Survey (QLFS), a repeated cross-sectional data. Their Internet connection treatment is binary at a smaller geographic level - enumeration area.

[9]60 percent of the observations are reported not employed and not earning any income, and I put zero as their income.

percentage points more, which is about 10 percent increase from the mean (column 1). Internet's negative impact on network search is small and not statistically significant, suggesting that network channel can be resilient to the Internet access shock (column 2). The impact on use of government agencies for job search is close to zero and insignificant (column 3). The number of different search methods average individuals used declines by 0.91 percentage points or 3.6 percent (column 4). This decline is driven mostly by less use of other search methods such as contacting other employers directly, or waiting at the side of roads. If this total number of methods can be viewed as a proxy for search effort, this result could suggest that Internet access leads to lower search effort.

Table 1.3: Impacts of Internet Connection on Job Outcomes

Outcome	Employed (0/1)	Total income (asinh)	Salary wage (asinh)	Weekly hours (asinh)	Formal contract (0/1)
% connected	0.135**	0.702**	1.610***	-0.199	0.007
	(0.065)	(0.312)	(0.521)	(0.162)	(0.078)
	[0.068]	[0.042]	[0.032]	[0.204]	[0.932]
Mean of outcome	0.366	2.048	3.692	3.533	0.688
Observations	37,034	37,034	10,487	10,487	9,319
R-squared	0.160	0.190	0.162	0.076	0.121

* Notes: Only workers between age 15 and 65 are included. All regressions include both district and year fixed effects, and controls for age, gender and education. Employed equals to 1 if the individual is employed with a salary job or self-employed. Hours and income are summed across each of the individual's jobs if more than one is reported. Total income are calculated using monthly income if salary employed, profit if self-employed, and as zero if unemployed. Inverse hyperbolic sign transformation are done to total income and salary wage. Only employed individuals are asked about wage, working hours, and contract types, so the number of observations for column 3-5 are small. Standard errors (in parentheses) are clustered at the district level. The wild bootstrap p-values [in brackets] are calculated following Cameron, Gelbach and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.4: Impacts of Internet Connection on Search Methods

Outcome	Online (0/1)	Network (0/1)	Government agency (0/1)	Number of search methods
% connected	0.064***	-0.017	0.013	-0.091*
	(0.017)	(0.064)	(0.016)	(0.051)
	[0.009]	[0.803]	[0.436]	[0.062]
Mean of outcome	0.061	0.247	0.030	0.251
Observations	32,856	32,856	32,856	32,856
R-squared	0.085	0.053	0.022	0.031

* Notes: Only workers between age 15 and 65 are included. All regressions include both district and year fixed effects, and controls for age, gender and education. Network, Online and Government variables are equal to 1 if workers have used this method when searching for jobs. "Number of search methods" is the total number of different methods ever used by the individual in the past four weeks. Besides online, network, and government agencies, other search methods include contacting other employers directly, waiting at the side of roads, placing ads, or seeking financial assistant to start own business. Standard errors (in parentheses) are clustered at the district level. The wild bootstrap p-values [in brackets] are calculated following Cameron, Gelbach and Miller (2008). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.5.2 Robustness checks

The first set of robustness checks examines whether the timing of the broadband Internet roll out correlates with time-varying covariates and/or trends. Column 1 shows the results without any controls, and column 2 repeats the main results with time-varying covariates. In column 3, I include linear trends interacted with baseline (year 2008) demographic covariates. In column 4, I allow for municipality-specific linear trends. I also show results including individual fixed effects in column 5. The point estimates are similar across these specifications, except for the estimations on network.

Table 1.5: Main results robustness checks

	TWFE β_{fe}					DID_m
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable						
Employment	0.109*	0.135**	0.102	0.175*	0.101	0.170
	(0.063)	(0.065)	(0.065)	(0.099)	(0.079)	(0.255)
Total Income	0.649**	0.702**	0.609*	0.933	0.481	0.692
	(0.317)	(0.312)	(0.347)	(0.671)	(0.373)	(1.500)
No.of Methods	-0.085*	-0.091*	-0.140**	-0.023	-0.221***	-0.111
	(0.049)	(0.051)	(0.054)	(0.145)	(0.074)	(0.365)
Online	0.067***	0.064***	0.060***	0.079**	0.027	0.019
	(0.019)	(0.017)	(0.018)	(0.036)	(0.019)	(0.068)
Network	-0.029	-0.017	-0.063	0.093	-0.055	0.132
	(0.063)	(0.064)	(0.064)	(0.080)	(0.065)	(0.213)
Observations	32,923	32,856	32,817	32,856	32,923	20,306
Location FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Time-varying covariates		Y	Y	Y		
Trends interacted with						
baseline covariates			Y	Y		
location FE				Y		
Individual FE					Y	

* Notes: Each cell is from a separate regression where the independent variable is $PercentConnected$. Column 1-5 use the standard TWFE estimator, and column 6 uses the robust estimator DID_m proposed by de Chaisemartin and D'Haultfœuille (2020). Less observations are included for the DID_m because comparison between later treated and early treated groups are dropped. Standard errors (in parentheses) are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The second set of robustness checks are related to recent research on TWFE with heterogeneous treatment effects. TWFE regressions are unbiased for an ATE only if the treatment effect are constant between groups and over time. With heterogeneous treatment effects and under a parallel trends assumption, TWFE may estimate a weighted sum of treatment effects across periods and units, with some negative weights.The negative weights could bias the treatment coefficient in TWFE regressions close to zero or negative, even if the treatment effect is positive for every unit

$\times$ period (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2022). Several alternative difference-in-difference (DID) estimators robust to heterogeneous effects have been proposed recently. Most of them apply to binary treatments that follow a staggered design (Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; Borusyak, Jaravel and Spiess, 2022). Only one estimator DID_m proposed by de Chaisemartin and D'Haultfœuille (2020) can be extended for continuous treatments, which applies to my research design. The DID_m estimator is a weighted average, across treatment intensity d and period t, of DIDs comparing the $t-1$ to t outcome evolution of groups whose treatment goes from d to some other value, and of groups with a treatment equal to d at both dates, normalized by the intensity of the treatment change experienced by the switchers.

I first estimate the weights attached to TWFE estimator $\hat{\beta_{fe}}$, and find that 45 percent are positive, 55 percent are negative. The negative weights indicate that $\hat{\beta_{fe}}$ may not be robust to heterogeneous effects, although the negative weights only sum to -0.09. The correlation between the weights attached to $\hat{\beta_{fe}}$ and the year t is equal to 0.08 (t-stats = 9.7), suggesting that the effect of Internet may be different in the early years than in the later years of the panel. Given the heterogeneous treatments, I compute the robust DID_m estimator using *did_multiplegt* in Stata. Table 1.5 column 6 shows that the DID_m estimates share the same sign but with larger effects from the TWFE estimates, except for network search. Although network search results are imprecisely estimated using the standard TWFE. The standard errors of DID_m estimations are larger, probably because less variations are used after dropping comparisons between later and early treated groups as suggested by de Chaisemartin and D'Haultfœuille (2020).

1.5.3 Heterogeneous effects

Given the high unemployment rate among young people in South Africa, I examine the heterogeneous effects of Internet exposure on job outcomes and search channels by workers' age. I use equation 1.7 where I interact $PercentConnected$ with a dummy variable for young workers between 15 and 24 years old. Estimation results are reported in Table 1.6.

$$Y_{ijt} = \alpha + \beta_1(PercentConnected_{jt} \times Youth_i) + \beta_2 PercentConnected_{jt} + \beta_3 Youth_i + X'_{ijt}\delta + \gamma_t + \theta_j + \epsilon_{ijt} \quad (1.7)$$

where the dummy variable $Youth_{it}$ indicates whether individual i is between 15 and 24 years old at time t.

Compared with older workers in areas with more Internet, young workers share similar probability of getting a job (column 1), but earn significantly less (column 3). They will try more number of search methods (column 5), and are more likely to increase searching through personal networks (column 9). These results suggest that young workers spend more effort searching for jobs, but the methods they choose are not as effective as the experienced. Their increasing reliance on personal networks suggests that Internet could make it easier to communicate with family and friends using tools such as emails or social media. I test if the Internet has an impact on the strength of social capital in section 1.6.2.

Family networks are particularly important to the labor outcomes of youth when transitioning from school to work (Kramarz and Skans, 2007). Thus, I include parents' education level and its interaction with Internet access to account for social economic status and network quality. The results on job outcomes are comparable when including these variables (column 1-4). As for search behaviors, Internet will cause workers whose parents have primary education to use less number of search methods, more likely to use online search, and less dependent on network search. These results suggest that existing social network variances can play a role in the young workers' choice of job search methods.

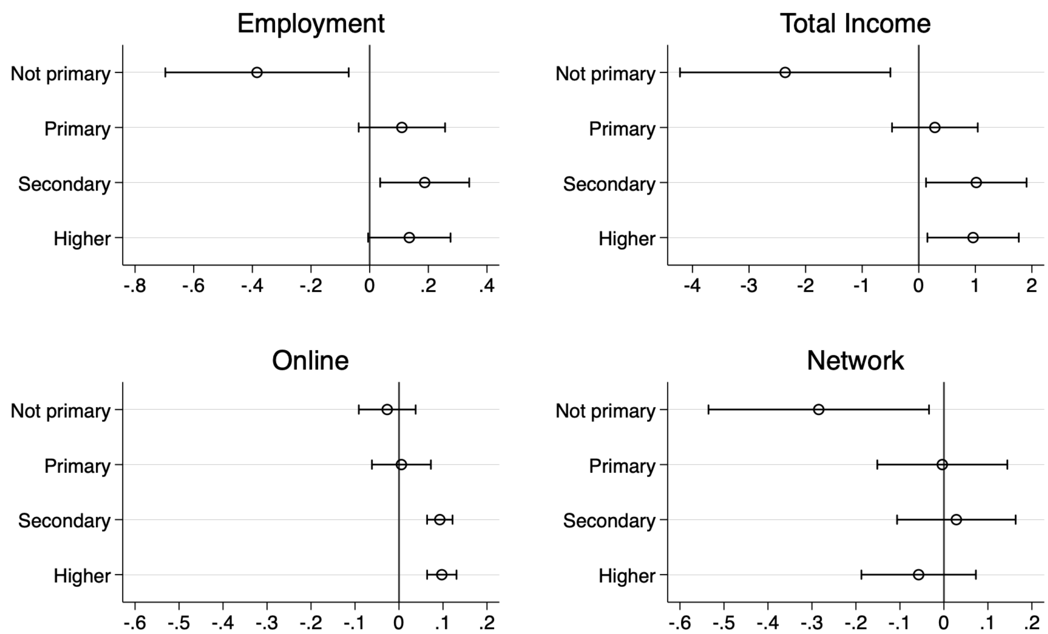
Considering how Internet can be a skill-biased technology as documented in many rich countries (Michaels, Natraj and Van Reenen, 2014; Akerman, Gaarder and Mogstad, 2015), I test if this is true in South Africa by interacting Internet penetration rate with educational attainment dummies for no school, primary, secondary, and tertiary education. The education level is used as a proxy for skill level here.

Results in figure 1.4 indicate that Internet connection increases the employment and income among the more educated workers the most. This finding is similar to Hjort and Poulsen (2019). However, individuals with no education do not benefit from Internet connection. This group of workers search less with both online and network channels, are less likely to find a job, and their total income will decrease. The results on uneducated workers contrast with Hjort and Poulsen (2019), where they find fast Internet reduces unemployment inequality across all education groups.[10]

I also test if parents' education level is a confounding factor in Table A1. Including the interaction terms of Internet and parents' education, main effects of Internet on workers with and without primary education do not change much. The coefficients on parents education interaction term are statistically significant, suggesting that existing social economic differences may play a role in the job outcomes and search methods for jobseekers.

[10]Though Hjort and Poulsen (2019) find the employment outcomes for workers with no education in eight other African countries do not benefit from fast Internet either.

Figure 1.4: Internet effects by education level



Each panel plots the coefficients of Internet penetration rate and highest education level interaction, from regressions of labor outcomes and search channels on Internet connectivity. All models include location and year fixed effects, and control for age and gender. 95% confidence intervals are displayed.

*Notes: Each panel plots the coefficients of Internet connectivity rate and highest education level interaction, from regressions of labor outcomes and search channels on Internet connectivity. All models include location and year fixed effects, and control for age and gender. 95% confidence intervals are displayed.

Table 1.6: Impacts of Internet Connection on Job Outcomes by Age

Outcome	Employed		Income		No.of Methods		Online		Network	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
% connected	0.141*	0.109	1.330***	0.838*	-0.196***	-0.118	0.065***	0.060***	-0.191	-0.132**
	(0.073)	(0.083)	(0.392)	(0.483)	(0.060)	(0.084)	(0.013)	(0.021)	(0.126)	(0.064)
... × youth	0.007	0.053	-1.376**	-1.048	0.332***	0.541***	-0.003	0.011	0.353**	0.412***
	(0.055)	(0.057)	(0.625)	(0.670)	(0.103)	(0.189)	(0.025)	(0.032)	(0.175)	(0.083)
... × w.educated parents		-0.016		-0.144		-0.180**		0.046		-0.095**
		(0.040)		(0.468)		(0.085)		(0.044)		(0.046)
youth	-0.354***	-0.381***	-2.599***	-2.895***	-0.229***	-0.226***	-0.056***	-0.063***	-0.282***	-0.279***
	(0.013)	(0.015)	(0.143)	(0.158)	(0.023)	(0.026)	(0.006)	(0.007)	(0.011)	(0.015)
w.educated parents		-0.012		0.029		-0.035		0.027***		-0.049***
		(0.012)		(0.126)		(0.022)		(0.008)		(0.011)
p-values(youth)	0.017	0.029	0.868	0.739	0.150	0.057	0.052	0.115	0.324	0.001
Mean of outcome	0.366	0.399	2.340	2.681	0.251	0.253	0.061	0.067	0.247	0.265
Observations	37,034	23,993	32,411	21,015	38,436	24,112	32,855	21,300	32,856	21,301
R-squared	0.209	0.199	0.237	0.244	0.037	0.042	0.090	0.106	0.090	0.087

[*] Notes: Only workers between age 15 and 65 are included. All specifications include year, location fixed effects. Control variables include age, gender, education, and baseline demographic covariates interacted with time trends. Standard errors (in parentheses) are clustered at the province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.6 Additional Mechanism Evidence

In this section, I discuss how Internet might affect employment and job search behaviors by skill levels and workers' age.

1.6.1 Access constraints

In Table 1.4, I find more Internet access enhances the use of online information search, and has little impact on the use of social network search. In section 1.5.3, I show choices of search channel made by uneducated or young workers respond to more Internet availability differently than their peers. Even after Internet is made more available in their areas, individuals without primary education will not use online search, and young workers will increase their reliance on personal networks.

One possible explanation is that there are other constraints prevent disadvantaged workers from accessing the Internet for online job search. Figure A2 shows that the high cost of equipment is the most important reason for not having Internet access at home, according to the General Household Survey (GHS) in 2018.

If we consider the computer a tool necessary for online job search, computer ownership can be used as a proxy to test if accessing costs are different for heterogeneous workers.[11]I use both the individual and household survey data from NIDS, and show how Internet affects computer ownership, literacy, and spending in Table 1.7. All regressions include individual fixed effects in additional to location and year fixed effects.

For skilled workers, their probability of owning a computer is 9 percentage point higher than unskilled workers, and they are more likely to be computer literate (Table 1.7,panel A column 1-2). Their households are also more likely to spend money on Internet (column 5). Interestingly, I find cellphone ownership are lower for the skilled workers than the unskilled (column 3, 6). Young workers are obviously more tech-savvy: more likely to own a computer or cellphone, and know how to use a computer (Table 1.7,panel B column 1-3). However, it seems that they are not using this technology to search for job information online directly, but rather to enhance personal networks for sharing job information. Most people probably communicate with families and friends or use social networking websites through a cellphone, so the more widely available cellphone could suggest no significant cost difference in accessing the Internet for network job search. This could explain why we do not see large differences in using network search for the skilled and unskilled in column 9-10 Table A1.

The findings about computer ownership and literacy suggest that technology may not make a difference if there are other constraints. Similar result are found in rural South Africa, where the rollout of mobile phone networks increased employment among women, but only for those who did not have significant family responsibilities (Klonner and Nolen, 2008).

1.6.2 Internet activities - social networking

Job search response could also depend on the various uses of Internet technology. Table A2 Panel B shows that social networking is the most important Internet activity (44.5%), while only about 12% survey respondent uses Internet for job search.[12] It is possible that internet communication may provide a cheap way for people to maintain relationships with people outside of their primary groups, such as classmates, former colleagues, or acquaintances (Gee, Jones and Burke, 2017; Armona, 2021). Social networking might also provide a way through which individuals can form new links by associating with others online who share specific interests. If Granovetter (1973)'s strength of weak ties hypothesis can be applied to online relationship, having a wide network of "weak" ties will provide greater quantity of job information. Then personal networks may complement the Internet in the job search process.

[11] Smartphones which cost less can be a substitute for computers for many functions. However, in a survey of people who used smartphones to apply for a job, 47% had difficulties accessing content that did not display properly, 38% had difficulties entering in a large amount of text, 37% had difficulties submitting required files and supporting documentation, and 23% had difficulties bookmarking saved job applications for later(Smith 2015)

[12] Source: Research ICT Africa (RIA)

Table 1.7: Impacts of Internet Connection on Cell and PC Ownership

	Individual			Household			
	own a computer	computer literate	own a cellphone	own a computer	spent money on internet	own a cellphone	spent money on cellphone
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: by education level							
% connected	-0.005	-0.179***	-0.012	0.023	0.005	0.176**	0.054
	(0.014)	(0.050)	(0.105)	(0.032)	(0.015)	(0.075)	(0.165)
... × beyond primary	0.090***	0.256***	-0.157***	0.107	0.034*	-0.228***	-0.236*
	(0.026)	(0.061)	(0.054)	(0.071)	(0.020)	(0.047)	(0.124)
beyond primary	-0.008	0.080***	0.182***	-0.013*	-0.001	0.025**	0.042**
	(0.005)	(0.015)	(0.016)	(0.007)	(0.004)	(0.011)	(0.019)
p-values(primary)	0.002	0.248	0.073	0.084	0.189	0.322	0.070
Observations	33,829	32,850	33,841	35,715	26,777	35,768	26,724
R-squared	0.530	0.670	0.493	0.583	0.444	0.384	0.446
Panel B: by age							
% connected	0.034	-0.057	-0.155*	0.079	0.030	0.031	-0.108
	(0.021)	(0.044)	(0.085)	(0.052)	(0.023)	(0.064)	(0.113)
... × youth	0.093***	0.209***	0.304***	0.030	-0.010	0.020	0.086
	(0.029)	(0.060)	(0.082)	(0.028)	(0.020)	(0.037)	(0.096)
youth	-0.025***	-0.089***	-0.050***	-0.010	-0.001	0.003	-0.034**
	(0.007)	(0.017)	(0.018)	(0.007)	(0.004)	(0.013)	(0.017)
p-values(youth)	0.000	.059	0.279	0.030	0.407	0.481	0.858
Observations	33,767	32,788	33,779	36,696	27,629	36,750	27,580
R-squared	0.532	0.671	0.492	0.579	0.436	0.380	0.440
Mean of outcome	0.051	0.292	0.714	0.107	0.012	0.851	0.738
Individual FE	Y	Y	Y	Y	Y	Y	Y

* Notes: Only workers between age 15 and 65 are included. All specifications include individual, location and year fixed effects. Standard errors (in parentheses) are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To test if the Internet affects the strength of social capital, I use favor exchanges behaviors with people outside of the household in the past year from the NIDS as dependent variables. Table 1.8 panel A shows that Internet does not impact the favor exchanges activities much differently for more or less educated workers. However, the negative coefficient on $PercentConnected$ on panel B suggests that older workers' social networks might be hurt by the Internet. And the coefficients on $PercentConnected \times youth$ are large, positive, and statistical significant for all favor exchange activities, implying that young workers are more likely to enhance their networks when Internet becomes more available in their areas. This can help explain why we see with more Internet, young people increasingly rely on personal networks for job information previously in Table 1.6. So for young workers without much social capital, they prefer to strength their networks and find job information through their personal networks, rather than using the Internet to search for jobs online directly.

1.6.3 Adoptions by household and firms

In previous analysis, the key variable of interest $PercentConnected$ represents the Internet availability rate in the district where the worker is living, it does not directly reflect the actual Internet access of the households or individual. As a supplement source, I use the General Household Survey (GHS) by Statistics South Africa to show some first stage correlations between Internet availability and adoption.The target population of the GHS survey consists of all private households in all nine provinces of South Africa and residents in workers' hostels. The sample size is about 24 thousand households each year, from 2009 to 2021. This survey includes information about whether households had at least one member who had access to or used the Internet, which can be used as proxy for direct Internet adoption rate.

Households in South Africa are generally more likely to have access to the Internet at work than at home or at Internet cafes or at educational institutions. In 2021, Internet access using mobile devices (66%) is the most common way compared to access at home (6%), at work (16%) and

Table 1.8: Impacts of Internet Connection on Favor Exchanges

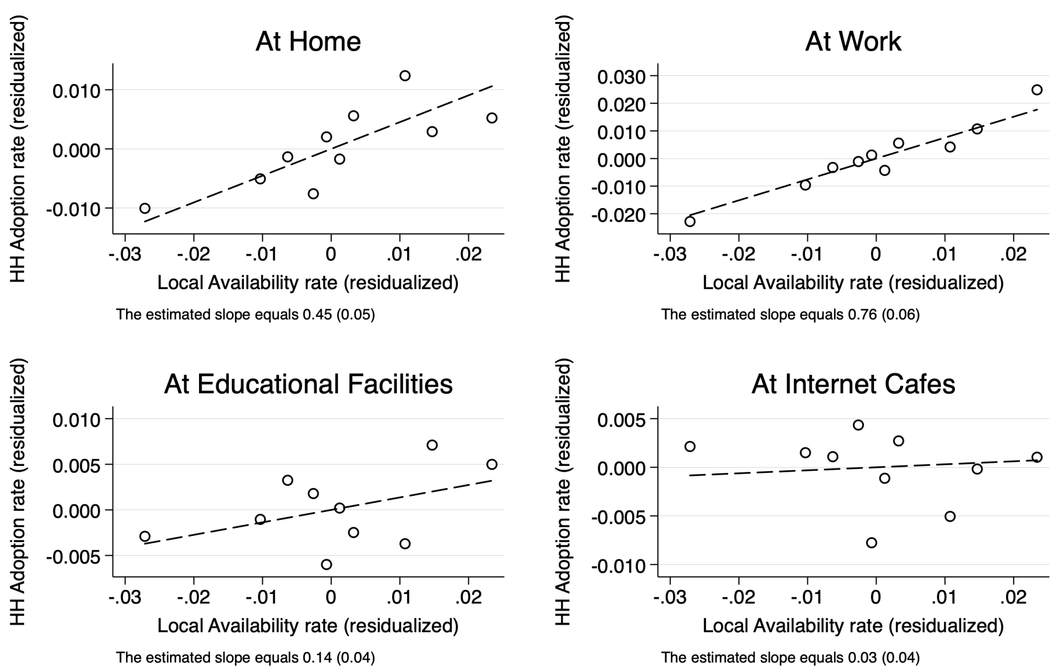
	exchanged favors (1)	gave favors (2)	received favors (3)
Panel A: by education level			
% connected	-0.092	-0.102	-0.003
	(0.085)	(0.089)	(0.047)
... × beyond primary	-0.012	0.033	-0.038
	(0.055)	(0.054)	(0.037)
beyond primary	0.020	-0.012	0.029***
	(0.013)	(0.010)	(0.010)
p-values(primary)	0.122	0.194	0.266
Observations	34,526	34,205	34,525
R-squared	0.369	0.363	0.366
Panel B: by age			
% connected	-0.130*	-0.095	-0.044
	(0.069)	(0.062)	(0.037)
... × youth	0.186***	0.091**	0.101**
	(0.060)	(0.039)	(0.049)
youth	-0.037***	-0.042***	0.001
	(0.014)	(0.012)	(0.012)
p-values(youth)	0.498	0.958	0.337
Observations	34,534	34,213	34,533
R-squared	0.370	0.363	0.366
Mean of outcome	0.138	0.055	0.090
Individual FE	Y	Y	Y

* Notes: Dependent variables equal 1 if individuals have exchanged, gave or received favors with people outside of their household in the past year. All specifications include individual, location and year fixed effects. Standard errors (in parentheses) are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

elsewhere (15%) (Figure A3).

I match the Internet access types data from the GHS with the key variable of interest $PercentConnected$ by province and year, and find a positive impact of broadband Internet availability in the local area on households' Internet adoption rates. Figure 1.5 shows scatterplots of the Internet adoption rate by places of access, against the Internet availability rate in the province, after taking out provincial and year fixed effects. The x-axis reports residuals from a regression of percent of populations connected to broadband on province and year fixed effects, and the y-axis reports residuals from a regression of households having Internet by access types on province and year fixed effects.

Figure 1.5: Internet availability and household adoption rate, by places of access



Note: The scatter plot shows average (residual) adoption at (residual) availability deciles, by places of access

The figure is based on the following regression that uses the sample of households for which we observe whether or not they have Internet access at home, work, nearby Internet cafes, or educational facilities:

$$d_{ijt} = \delta PercentConnected_{jt} + \gamma_t + \theta_j + \nu_{ijt} \tag{1.8}$$

where d_{ijt} equals one if household i in province j at time t had Internet access at home (or at work, at educational facilities, at nearby Internet cafe) and is zero otherwise.

The coefficient on the availability rate δ is about 0.43 with a standard error of 0.07 for Internet access at home. This estimate implies that a 10 percentage point increase in broadband availability will induce 4.5 percent of households to gain Internet access at home. Adoption at work responses the most, while access from Internet cafes do not change much. These findings illustrate that when Internet becomes available, adoption is not universal; instead, it is more likely adopted in places in which complementary factors are abundant, including computer ownership and computer literacy.

Table 1.9: Impacts of Internet Connection on Adoptions by Place of Access

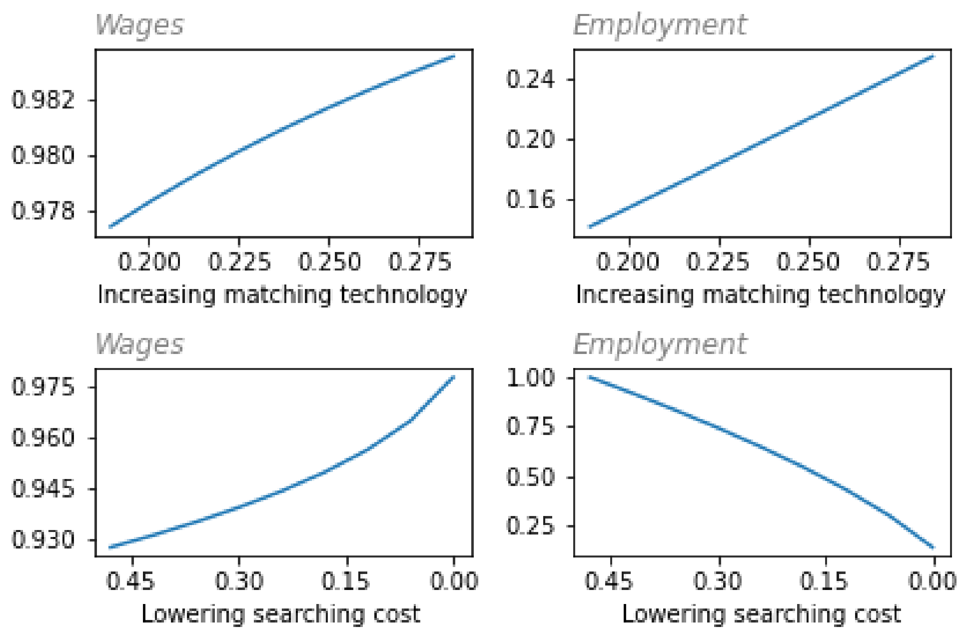
	Anywhere	At home	At work	Educational facilities	Internet cafes
	(1)	(2)	(3)	(4)	(5)
% connected	0.946***	0.427***	0.805***	0.158*	0.086
	(0.173)	(0.073)	(0.113)	(0.090)	(0.135)
Mean of outcome	0.220	0.075	0.133	0.058	0.048
Observations	127,024	126,349	126,349	127,024	127,024
R-squared	0.106	0.046	0.069	0.033	0.029

* Notes: All specifications include location and year fixed effects. Standard errors (in parentheses) are clustered at the province level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.6.4 General equilibrium impacts

Internet can affect both the labor supply and the labor demand, and the results on employment and wages in Table 1.3 and Table A1 should reflect the equilibrium outcomes. Without employers or firms' data, I can not say much about how employers' job creation decisions respond to more Internet access empirically. Instead, I study the expected general equilibrium impacts of Internet using the theory of unemployment and vacancies, the Diamond-Mortensen-Pissarides(DMP) model (Diamond, 1982; Mortensen, 1986; Pissarides, 1985). I simulate the Internet access shock by changing the parameter value of matching technology or the value of unemployment income, and solve the model numerically. Expected equilibrium changes of wage and employment are summarized in Figure 1.6. A brief description of the DMP framework is presented in Appendix B.

Figure 1.6: New equilibrium simulation using DMP model



Notes: By changing the parameter value of matching technology A_t and the value of unemployment income b, I numerically solve the new equilibrium after a Internet access shock. The baseline parameter values are from Hagedorn and Manovskii (2008).

The first mechanism that Internet could impact the labor market is an improvement in matching efficiency. A key process in the DMP-framework is the "matching function", which uses job vacancies and jobseekers as input, and produces a number of firm-worker matches given a matching technology A. Upper panel, Figure 1.6 show that with higher values of A, more hires can be generated from the same number of jobseekers and vacancy, thereby increasing the employment rate. Since jobseekers expect to be matched faster, their outside option improves, which will drive up wages in new employment relationships.

Internet access can also reduce the cost of learning about and applying for jobs. Unemployment income in the DMP model include both actual unemployment transfer and imputed value of time to unemployed workers. Lower searching cost implies higher value of leisure, thereby increasing the value of unemployment income. Everything else equal, this increasing unemployment benefits exerts an upward pressure on the equilibrium wage. This lowers the profits employers receive from filled jobs, leading to a decline in vacancy creation. Lower vacancies imply a lower job finding rate for workers, which leads to an decrease in employment as shown in lower panel, Figure 1.6.

Combining these two mechanisms, the effect on wages is unambiguously positive, but the total effect of the Internet on employment depends on the relative importance of these two.

1.7 Conclusion

This paper provides evidence on how Internet availability affects job market outcomes and job search activity in South Africa. By comparing individuals in areas with various Internet penetration rates, I find that jobseekers in locations with better connectivity have higher employment rates and income, and the impact is driven by a significant increase in employment of experienced and skilled workers. When Internet is made more available, only skilled ones increase their use of online job search. Young workers will search through more methods, while rely more on personal networks.

These findings suggest that not everyone stands to benefit from improved Internet availability automatically. Associated labor market disruptions can be painful and can result in higher inequality. High cost remains the largest barrier for Internet usage.[13] The low-skilled or less-educated almost exclusively use mobile phones to access the Internet. Poor computer literacy could limit the productive use of this technology. Besides improving Internet infrastructure, complementary policies aiming at updating skill and digital literacy are critical for ensuring the overall benefits be shared broadly.

[13] Table A2 Africa ICT access survey, Fig A2

Chapter 2

Who has the right to a job? Labor market competition and men's support for women's work

with Rachel Heath and Alex Philip

2.1 Introduction

There is growing evidence that social norms affect women's labor supply (Alesina, Giuliano and Nunn, 2013; Bursztyn, González and Yanagizawa-Drott, 2020; Fernández and Fogli, 2009; Jayachandran, 2021; Olivetti, Patacchini and Zenou, 2020). The causality appears to go in the other direction as well: an experiment in India that increased women's labor supply by increasing their control over income increased social support for women's work three years after the experiment began (Field et al., 2021).

We ask whether the degree to which women work in industries in which they compete with men affects social support for women's work. We examine the case of India, which is notable for low (and recently decreasing) rates of female labor force participation (Heath and Jayachandran, 2017). We hypothesize that, while overall increases in female labor force participation increase social support for women's work (as in Field et al. (2021)), if women enter sectors in which they compete with men, men will be less likely to support women's work.

To test this hypothesis, we use five rounds of the National Sample Survey (NSS) Employment and Unemployment Modules (ranging from 1987 to 2009) to construct a measure of the competition from women faced by a man in a given state at a certain time: what fraction female is his industry? We match this variable to data on support for women's work from the five rounds of the World Values Survey (spanning from 1990 to 2012), looking in particular at respondents' disagreement with the statement: "When jobs are scarce, men should have more right to a job than women." We estimate plausibly causal effects of exogenous changes in labor market outcomes conditional on state and year fixed effects, paying attention to the recent concerns that time-varying treatment effects can yield biased estimates (de Chaisemartin and D'Haultfœuille, 2020; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021).

We indeed find that, conditional on overall female labor supply, at times when the typical man faces more labor market competition from women, people report less support for women's labor supply. Specifically, a one standard deviation increase (within-state) in competition leads to a 12 percentage point decrease in support for women's work, on a mean of 37%. This coefficient is not statistically different between men and women, fitting with evidence from South Asia that women don't necessarily have any more liberal gender attitudes than do men, in areas such as tolerance of intimate partner violence (Schuler and Islam, 2008) or son preference (Jayachandran, 2017). Meanwhile, the main effect of female labor force participation fits with evidence that increasing female labor supply liberalizes norms around women working: a one standard deviation increase (within-state) in female labor supply leads to a 14 percentage point increase in support for women's work.

We provide evidence that the direction of causal channel behind these results is that labor markets changes affect norms, rather than norms affecting labor supply. First, note that the most obvious reverse causality story – gender norms liberalize, so women feel freer to enter jobs where they work closely with men – would predict a positive relationship between support for women's work and labor market competition from women, which is the opposite of our empirical results. We also test whether lagged support for women's work affects women's labor supply, and find no evidence for a story in which norms change first and then labor supply responds.

Our results relate to a literature on the intra-household determinants of female labor supply

(Heath and Tan, 2020; McKelway, 2020; Field et al., 2021; Lowe and McKelway, 2021). While we estimate wage losses for men when more work in their industries, these jobs are available to their own wives. In a unitary household, men should benefit from the increased household-level income, especially given evidence that norms around housework are sticky enough that women's time in housework does not respond to their increased labor supply (McKelway, 2023). However, it appears instead that men prefer women (even their own wives) not enter their industries, perhaps because women working lowers their own bargaining power.

We also contribute to a literature on the determinants of men's attitudes towards women and their support for women's rights. Exposure to women political leaders (Beaman et al., 2009) or as peers on military teams (Dahl, Kotsadam and Rooth, 2021) liberalizes gender attitudes about women's beliefs. We point out that exposure to women may not always liberalize gender attitudes if such exposure comes with a cost for men. We thus join Fernández (2014) in arguing that men are more supportive of women's rights when such rights do not affect them directly.

The rest of the paper proceeds as follows. In Section 2.2, we present the data sources, a descriptive analysis of gender attitudes, and define the labor competition variable. Section 2.3 outlines the estimation strategy and tests for reverse causality. We show the main results in Section 2.4. In Section 2.5, we examine the robustness of the estimations using a shift-share instrument and alternative two-way fixed effects estimators. We conclude the paper by discussing the policy implications in Section 2.7.

2.2 Data

2.2.1 Data sources

We use two data sources for our analysis: the World Values Survey (WVS) (rounds 1990, 1995, 2001, 2006, 2012) and the Indian National Sample Survey (NSS) Employment and Unemployment

Modules (rounds 1987, 1993, 1999, 2004, and 2009).

The WVS provides data on the beliefs and values of Indian citizens. Sample sizes for each round ranged from 2000 to 4000 respondents, leading to a total sample of 9100 pooled across all rounds, as described in table 2.1.[1]

The NSS is a nation-wide household survey that contains data on labor supply. Sample sizes range from around 450,000 to 670,000 respondents depending on the year. We use the NSS to measure labor force participation as well as create a labor competition measure (details on in section 2.2.3). To determine an individual's economic activity, we rely on the "principle usual activity status" (PUAS) question included in the NSS, which asks about an individual's primary activity for the past 365 days [2]. We categorized industries using their two-digit National Industrial Classification (NIC) code, and collapsed the NSS data by state and year to create state-year specific labor competition variables.

To link the WVS and NSS data, we match each individual at given states in the WVS data to their closet round of NSS with a lag of 2 to 3 years for our analysis: see table A3. That is, we measure whether recent changes in labor market outcomes correspond to current gender attitudes. Given that it likely takes some time for respondents to notice labor market changes and alter their gender attitudes to change in response, we argue that this is a reasonable construction of our key independent variable.

2.2.2 Gender attitude

Our main outcome variable of interest is a binary variable measuring attitudes towards women's work, which was asked in all five waves of the World Values Survey. We specifically focus on the question that asks respondents that if they agree or disagree with the statement that *"When jobs are*

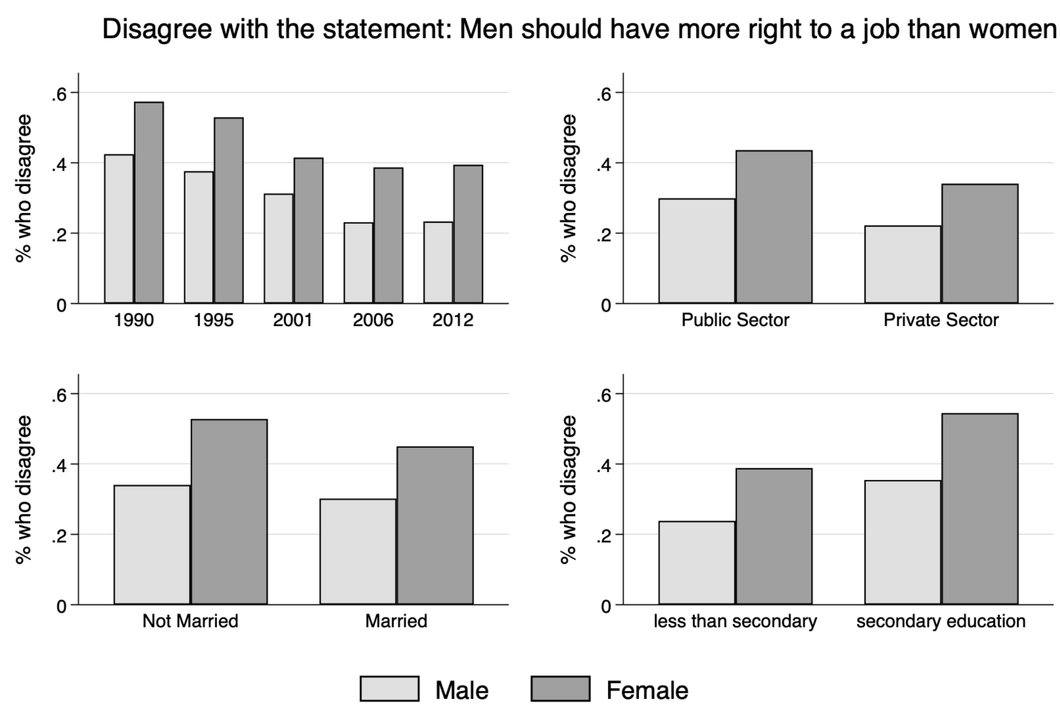
[1] Although the sample used in the WVS study consists of relatively better-off (eg. literate, Hindu) Indians and not nationally representative, it is important to note that these individuals likely play a significant role as opinion leaders in shaping overall social norms.

[2] We define labor force participation as individuals who are either employed or unemployed but seeking employment. The PUAS codes ranging from 11 to 81 (as in Figure A4) are used to identify individuals in the labor force.

scarce, men should have more right to a job than women.”. Here, a *“disagree”* response indicates more egalitarian attitude and greater support for women’s work.

Figure 2.1 shows descriptive analysis of attitudes towards women’s work by gender, survey years, sectors, marital status, and education levels. We find that female respondents generally have a higher percentage of "disagree" responses compared to male respondents. Over the time period we studied, we also observe a noticeable decline in support for women’s work among both men and women. Public sectors workers, individuals who are not married, and those with secondary educations tend to have more liberalized attitudes. Men with less than secondary education have the lowest support, while women who are not married or have completed secondary education show the highest support.

Figure 2.1: Gender attitude towards jobs from WVS, 1990-2012



2.2.3 Define competition from women

To capture the intra-sectoral competition between male and female, we constructed the key measurement of *competition from women* at state-year level, defined as follows:

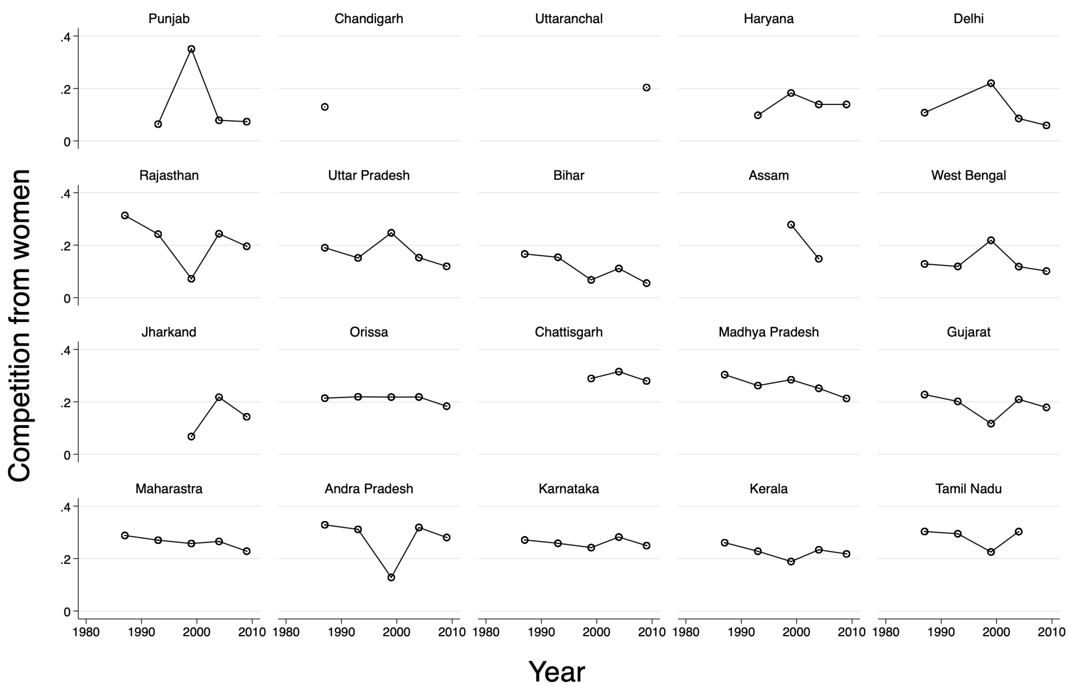
$$Competition_{s,t} = \frac{1}{\sum_{i=1}^{N_{s,t}} \mathbb{1}(male_i)} \sum_{i=1}^{N_{s,t}} \mathbb{1}(male_{i,s,t,k}) \frac{\mathrm{N}_{k,s,t}^{Female}}{\mathrm{N}_{k,s,t}} \tag{2.1}$$

We first computed the concentration of women working in industry k in state s at time t, as the ratio of number of female workers $\mathrm{N}_{k,s,t}^{Female}$ to the total number of workers in that industry $\mathrm{N}_{k,s,t}$. Then we calculated the average concentration of female co-workers for individual male workers at the state-year level, by dividing the sum of these concentrations by the total number of male workers. This variable measures the percentage of female co-workers (within the same industry) that an average male worker would have in a given state. Figure 2.2 shows a plot of this variable by state and year, which are the key variations we explore in our identification.

This variable is correlated with the female labor force participation ($flfp$) at the state-year level[3]; however, it has a distinct interpretation. While the female labor force participation captures the overall female labor supply in a state, the competition variable also takes into account the industry-specific variations in which women decide to join. In other words, the competition variable captures the potential labor market competition that male workers face from female co-workers in the same industry, whereas $flfp$ does not account for this intra-sectoral dynamics. Therefore, the competition variable provides a more nuanced measure of gender-related labor market competition that may affect male and female workers differently. We include $flfp$ as an additional control in all estimation models.

[3]The correlation between competition and female labor force participation is 0.47 after controlling for state and year fixed effects. A scatter plot of the residuals can be found in Figure A5.

Figure 2.2: Competition from women by State and Year



2.2.4 Summary statistics

Table 2.1 presents summary statistics for the individuals included in the final estimation samples merged from the WVS and NSS, broken down by gender.

A total of 9,100 individuals are included, with 42 percent being female. The sample is largely married (78%) and has an average age of 36. Over half of the sample has completed secondary education and are employed, with 8% working in the public sector, 40% in the private sector, and 52% in other less formal sectors. About 37% of the respondents expressed support for women's work, i.e., they disagree with the premise that scarce jobs should go to men. At the state-year level, the average competition from women is 0.2 and the average $flfp$ is 0.27.

Individual characteristics vary significantly between men and women in this sample, so we estimate all our models by gender and include individual specific controls. Women in the sample are younger and more likely to be married. 49% of women has completed secondary education, however, only 26% of them are employed, compared to 74% of men. Conditional on working, the majority of women work in less formal sectors (67%). Average industry-level wage for women are slightly lower and significantly different from men's at the 95% level. Women are also more supportive of women's work (46% of women and 31% of men).

Table 2.1: Summary Statistics

	Total		Men		Women		
	mean	sd	mean	sd	mean	sd	p-value
Individual level							
Age	36.841	11.921	37.157	12.220	36.413	11.490	(0.003)
Female	0.425	0.494	0.000	0.000	1.000	0.000	(.)
Married	0.781	0.414	0.763	0.425	0.805	0.396	(0.000)
Secondary Education	0.569	0.495	0.625	0.484	0.492	0.500	(0.000)
Employed	0.534	0.499	0.745	0.436	0.258	0.437	(0.000)
Support for women working	0.376	0.485	0.311	0.463	0.465	0.499	(0.000)
Conditional on being employed...							
Public Sector	0.083	0.276	0.107	0.309	0.048	0.213	(0.000)
Private Sector	0.397	0.489	0.479	0.500	0.279	0.449	(0.000)
Other Sector	0.520	0.500	0.414	0.493	0.673	0.469	(0.000)
Average industry wage(log)	6.313	0.631	6.327	0.616	6.294	0.650	(0.013)
State-Year level							
Competition from women	0.200	0.074	0.200	0.074	0.200	0.075	(0.797)
$flfp$	0.271	0.125	0.271	0.125	0.272	0.125	(0.545)
Observations	9100		5236		3864		

Notes: Only working age (16-65) individuals are included. P-value for test of the difference in means between Men and Women, where null hypothesis is that they are equal.

2.3 Identification

2.3.1 Two-way fixed effects estimation

We estimate effects of competition on support for women working at the individual level using the following two-way fixed effects (TWFE) model.

$$Support_{ist} = \beta Competition_{st} + \theta flfp_{st} + \gamma_t + \lambda_s + \epsilon_{ist} \tag{2.2}$$

where $Support_{ist}$ is a binary variable that equals 1 if an individual i in state s at time t *disagrees* with the statement that "Men should have more right to a job than women when jobs are scare". $Competition_{st}$ is the percentage of female workers a typical male worker faces in his industry in state s at time t, calculated as in equation 2.1. We also include $flfp_{st}$, the average female labor force participation rate in state s at time t, state fixed effect λ_s, and year fixed effect γ_t. We also estimate a specification with a vector of individual-level controls X_{ist}, including age, education level, and marital status (as well as gender, as described in equation 2.3. We cluster standard errors at the state-year level.

We further explore whether the effects of competition on support for women working differs by the gender of the respondent, estimating equation 2.3:

$$\begin{aligned} Support_{ist} = {} & \beta_1 Competition_{st} + \beta_2 competition_{st} \times female_i \\ & + \theta_1 flfp_{st} + \theta_2 flfp_{st} \times female_i + \pi female_i + \gamma_t + \lambda_s + \epsilon_{ist} \end{aligned} \tag{2.3}$$

where interaction terms of competition (or labor force participation) and gender are included. Here, the main object of interest is β_2, representing the difference in average effect between men and women.

2.3.2 Test for reverse causality

One potential threat to our identification is reverse causality, where individual attitudes towards women working may impact their labor supply decisions. To address this issue, we test whether the lagged support variable can predict labor force participation and competition from women.

$$Y_{ist} = \beta Support_{s,t-1} + \gamma_t + \lambda_s + \epsilon_{ist} \tag{2.4}$$

We estimate equation 2.4, where Y_{ist} represents individual-level labor supply, including whether individual i in state s is in the labor force at time t,and the competition from women that individual i faces in his or her industry in state s at time t. The independent variable $Support_{s,t-1}$ is calculated as the percentage of respondents in the World Value Survey (WVS) who supported women working in state s at time t-1.

We find no evidence of a reverse causality relationship (Table 2.2). The coefficients of lagged support on labor force participation and competition are small and statistically insignificant (column 1, 3). To see whether the effects of lagged support differ by gender, we estimate equation 2.4 with an interaction term of lagged support and gender, and again find small and statistically insignificant impacts (column 2, 4).

2.4 Main results

Table 2.3 regresses a dummy variable for gender attitude towards whether scarce jobs should go to men than women, with a value of 1 for disagreement. We report four sets of estimates, all of which include year and state fixed effects. The first set estimates the average effect of competition on support for women's work, controlling for overall female labor force participation rate as in equation 2.2. The second set explores heterogeneity by gender by adding interaction terms between competition and gender, as in equation 2.3. The third includes individual-specific covariates such

Table 2.2: Lagged support on labor participation and competition

	In the labor force (0/1)		Competition from women	
	(1)	(2)	(3)	(4)
lagged_support	0.011		0.004	
	(0.017)		(0.025)	
lagged_support × Male		-0.005		0.005
		(0.039)		(0.025)
lagged_support × Female		0.022		0.004
		(0.052)		(0.025)
Year FE	Y	Y	Y	Y
Loc FE	Y	Y	Y	Y
Mean of Dep. Var	0.581	0.581	0.211	0.211
Observations	1,258,598	1,258,553	1,258,598	1,258,553
R-squared	0.014	0.347	0.795	0.795

Notes: Data on labor force participation and competition are from the NSS data set. The dependent variable "in the labor force" is a dummy variable defined as 1 if an individual is either employed or unemployed but seeking for jobs. The dependent variable "competition from women" is defined as in equation 2.1.

as age, education level, and marriage status. The fourth set of results includes a dummy variable for whether a state has passed the Hindu Succession Act (HSA)[4], as evidence has shown that this policy could increase women's labor supply, particularly into high-paying jobs (Heath and Tan, 2020).

The results suggest that, conditional on overall female labor supply, when labor market competition increases, people become less supportive of women's work. Specifically, a one standard deviation increase (within-state) in competition from women (about 0.051 – i.e. the composition of a typical man's industry is now 5.1 percent points more female) leads to an 10.7 percentage points decrease in average support for women's work, which represents a 28% decrease relative to a mean of 37% (Column 1). This coefficient is not substantively or statistically different between

[4]Amendments to the HSA explicitly made daughters eligible to be coparceners, which traditionally include only male relatives. This policy was phased into different states in India between 1976 and 2005, with some states enacting the changes earlier than others.

men and women (Column 2), suggesting that both genders may perceive increased competition from women in a similar way. Adding in additional individual controls or HSA policy does not materially affect the estimates (Column 3, 4).

While our main focus is on the competition variable, our finding on the overall female labor force participation variables coincides with previous literature (Field et al., 2021) in suggesting that increasing overall female labor supply liberalizes people's attitudes towards job equality between men and women. A one standard deviation increase in overall female labor force participation rate (about 0.079) will make people 14 percentage points more supportive of women's work, which translates to a substantial 38% increase relative to the mean. This effect holds for both men and women, and is robust to the inclusion of additional covariates and state-specific HSA policy. This finding implies that as women become more involved in the labor market, people become more accepting of women working.

The negative effect of competition from women on support for women's work could suggest that when women's labor market opportunities expand, especially when they work and compete with men in the same sector, men may perceive them as a threat to their own employment prospects and thus may become less supportive. It is also possible that men may believe that working in a sector with higher female representation may threaten their own sense of masculinity (Leavitt et al., 2022).

Moreover, our finding that the negative effect of competition is similar for men and women suggests that women, who have traditionally faced discrimination and more constraints in the labor market, may also view the entrance of other women into their sectors as a threat to their own economic status or perceived role in society.

Overall, our results suggest that while increasing labor market participation for women can lead to positive changes in attitudes towards women's work, it is important to be award of potential backlash from both men and women in response to increased competition from women.

Table 2.3: Main Effects on Support for Women Working

	Support for women working			
	(1)	(2)	(3)	(4)
Competition from women	-2.108*	-2.325*	-2.339*	-2.357*
	(1.231)	(1.270)	(1.269)	(1.256)
Competition from women × Female		0.469	0.446	0.445
		(0.706)	(0.705)	(0.705)
Female LFP	1.801**	1.800**	1.803**	1.813**
	(0.753)	(0.772)	(0.774)	(0.763)
Female LFP × Female		0.038	0.032	0.032
		(0.409)	(0.411)	(0.411)
Female	0.143***	0.039	0.059	0.059
	(0.016)	(0.048)	(0.048)	(0.048)
p-value(net effect of competition on women)		0.156	0.139	0.131
p-value(net effect of flfp on women)		0.022	0.019	0.017
Mean of Dep. Var	0.376	0.376	0.376	0.376
within-state standard deviation of				
competition	0.051	0.051	0.051	0.051
flfp	0.079	0.079	0.079	0.079
Observations	9,100	9,100	9,100	9,100
R-squared	0.096	0.097	0.106	0.106
Year FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Individual controls			Y	Y
HSA				Y

Notes: Dependent variable "Support for women working " equals to 1 if respondents answered "not agree" to the question "When jobs are scarce, men should have more right to a job than women" in the World Value Surveys. All regressions include state and year fixed effects. Column 3 add individual-specific controls including age and education levels. Column 4 also adds a dummy variable indicating whether the state has passed the Hindu Succession Act(HSA). Robust standard errors clustered at state-year level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.5 Robustness check

2.5.1 shift-share version of competition

We argue that our main results, as given in table 2.3, should be interpreted as the effect of an exogenous change in men's labor market competition from women on attitudes about women's work. The fact that we found no effect of lagged support for women's work on labor market outcomes provides evidence against a reverse causality interpretation of our findings (i.e. changing social norms prompt changes in labor supply). To further support our preferred interpretation, we construct a shift-share version of gender competition based on Bartik (1991) to reflect plausibly exogenous exposure to competition from women driven by changes in employment (elsewhere in the country) in a given area's industries that have high lagged values of competition from women.

Specifically, we calculate the predicted competition from women variable in each state and year as follows:

$$\widetilde{competition}_{s,t} = \sum_{i=1}^{N_{s,t}} \overbrace{\mathrm{competition_{k,-s,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N}^{Female}_{k,s,t_0}/\mathrm{N}^{Female}_{k,t_0}}^{\text{Share}} \tag{2.5}$$

where s indexes state, t year, and k industry. The "Shift", $\mathrm{competition_{k,-s,t}}$, is the average competition from women in industry k over all states except for state s in year t. This *delocalized* shift help remove any changes in intra-sectoral competition that might be caused by changes in the underlying characteristics of works in the state. The "Share" is the fraction of female workers in industry k that reside in state s at baseline t_0, which represents the local share of a certain industry. We fixed this share at baseline, and believe it could be less sensitive to endogeneity given high migration costs across states. A similar approach has been used for measuring gender wage gaps (Aizer, 2010) and migration flows (Card, 2001).

With the delocalized shifts, fixed shares, and two-way fixed effects, we argue that the exposure to competition from women is plausibly exogenous. We re-estimated model 2.2 and 2.3 using this shift-share competition variable constructed following 2.5, and reported the results in Table 2.4

Our findings persist when employing the shift-share version of the competition measure. The estimated effect of this competition variable on support for women working is attenuated compared to the main results: a one standard deviation increase in competition (0.014) is associated with a 2.5 percentage points decrease in average support for women working (Column 1). The specifications that account for gender heterogeneity and include additional controls have similar magnitudes of effect, although they are estimated with less precision (Column 2-4).

2.5.2 TWFE estimators with heterogeneous treatment effects

A fast-growing literature has highlighted that two-way fixed effects (TWFE) estimators may be biased if the treatment effects are heterogeneous across groups or over time, leading to the development of alternative estimators (Goodman-Bacon, 2021; Sun and Abraham, 2021; de Chaisemartin and D'Haultfœuille, 2022; Callaway and Sant'Anna, 2021). de Chaisemartin and D'Haultfœuille (2020) propose a new estimator, DID_M, which estimates the average treatment effects across all the group-year cells whose treatment changes from the last to current period. This estimator can be easily extended to applications with a non-binary treatment, which is well-suited for our study.

To use this new estimator, we made two modifications. First, we discretized the continuous competition treatment into 10 percentiles based on its distribution (Figure A6), which allowed us to apply the *did_multiplegt* package in Stata. Second, this new estimator requires a stable group assumption: between each pair of consecutive years, there are states where the competition from women does not change. To meet this assumption, we imposed a threshold of 0.02[5], that is, if the competition in one state changes by less than 0.02 year to year, it is treated as a stable group.

We compared the treatment effect estimated by the standard TWFE estimator (Table 2.5, Panel A) to the new DID_M estimator (Panel B), and found that our estimations are robust. The estimated

[5]This threshold is determined given that average year-to-year change in competition across states is about 0.015.

Table 2.4: Main Effects of Shift-share Competition on Support for Women Working

	Support for women working			
	(1)	(2)	(3)	(4)
$\tilde{Competition}$	-1.813*	-1.707	-1.498	-1.500
	(1.073)	(1.081)	(1.096)	(1.093)
$\tilde{Competition}\times$ Female		-0.298	-0.282	-0.280
		(0.202)	(0.201)	(0.203)
Female LFP	0.551**	0.386	0.380	0.382
	(0.261)	(0.261)	(0.261)	(0.261)
Female LFP $\times$ Female		0.404***	0.379***	0.380***
		(0.132)	(0.132)	(0.132)
Female	0.144***	0.0785**	0.0962***	0.0957***
	(0.016)	(0.036)	(0.036)	(0.036)
Mean of Dep. Var	0.376	0.376	0.376	0.376
within-state standard deviation of				
competition	0.014	0.014	0.014	0.014
flfp	0.079	0.079	0.079	0.079
Observations	9,100	9,100	9,100	9,100
R-squared	0.095	0.097	0.105	0.105
Year FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Control			Y	Y
HSA				Y

Notes: $\tilde{Competition}$ is using delocalized shifts and fixed baseline shares as defined in equation 2.5. Dependent variable "Support for women working " equals to 1 if respondents answered "not agree" to the question "When jobs are scarce, men should have more right to a job than women" in the World Value Surveys. All regressions include state and year fixed effects. Column 3 add individual-specific controls including age and education levels. Column 4 also adds a dummy variable indicating whether the state has passed the Hindu Succession Act(HSA). Robust standard errors clustered at state-year level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

treatment effect of competition using DID_M is -1.96, with a standard error of 1.342, which very close (7% smaller) to the standard TWFE estimator result $\hat{\beta_{fe}}$ in Panel A. We cannot reject the null hypothesis that these two coefficients are statistically different (t-stat = -0.07).

In addition, we estimated the effects of competition using a discretized version of the treatment variable and standard TWFE estimator in Panel C. Our results show that compared to lowest level of competition, higher competition leads to less support for women's work, with negative, smaller, and statistically significant coefficients for dummies on each percentile.

Table 2.5: Comparing different TWFE estimators

	Estimate	Standard error
Panel A: standard TWFE estimator		
$\hat{\beta_{fe}}$	-2.108*	1.231
Panel B: new TWFE estimator		
DID_M	-1.968	1.342
DID_M^{pl}	-0.137	0.056
Panel C: discretized treatment variable		
10th - 20th pctl	-0.232***	0.037
20th - 30th pctl	-0.116***	0.038
30th - 40th pctl	-0.314***	0.044
40th - 50th pctl	-0.277***	0.057
50th - 60th pctl	-0.352***	0.070
60th - 70th pctl	-0.477***	0.086
70th - 80th pctl	-0.589***	0.092
80th - 90th pctl	-0.684***	0.111
$\geq$ *90th pctl*	-0.877***	0.115

Notes: This table reports estimates of the effect of competition from women on support using different estimators. Panel A: $\hat{\beta_{fe}}$ is the $TWFE$ estimation result from Table 2.3(col 1). Panel B: DID_M is the estimator computed following de Chaisemartin and D'Haultfœuille (2020), with $female$ and $flfp$ as controls. To compute the DID_M estimators, competition is categorized into 10 bins based on its distribution. The DID_M estimators require to have stable groups whose treatment does not change between consecutive time periods. To meet this restriction, a threshold of 0.02 is imposed for determining treatment changes. That is, if the competition in one state changes by less than 0.02 year to year, it is treated as a stable group. Panel C report standard $TWFE$ estimation using discretized treatment variable dummies, with the competition level less than the 10th percentile as the reference group.

2.6 Potential mechanism: earning loss for men

As discussed earlier, a possible explanation for the negative impact of competition on gender attitudes could be the perceived threat to individual's employment prospects. This situation could arise if increases in competition from women are driven primarily by supply shifts into the industry rather than positive labor demand shocks, which would lower equilibrium wages in the industry. To investigate this potential mechanism, we analyze wage and sector information from the NSS data sets in this section.

Table A4 presents our analysis of individual wages using the same identification strategy outlined in Section 2.3.1. We find that men's wages decrease as the level of competition from women in their sectors increases. Specifically, a one standard deviation increase (within-state) in competition from women (about 0.051) leads to an 42 percent decrease in men's wage (Column 1). This effect is substantial, statistically significant, and robust after controlling for age and education levels (Column 2). Adding sector controls reduces the effect size by half to about 19 percent, although not statistically significant (Column 3). We do not observe any statistically significant impact on women's wage.

We also find that the negative effect of competition on men's wage is consistent across major sectors. Since adding sector controls shrinks the effect size, we wanted to test whether the results are driven solely by different sector compositions between men and women. To do so, we regressed individual wages on competition interacted with both gender and sectors, and plotted the interaction coefficients of the largest five sectors in Figure A7. The results show that men in the agriculture sector experienced the largest and statistically significant drop in their wage, while the decrease size was comparable with other sectors. Women's wage in these sectors increased but are nosily estimated.

These findings suggest that competition from women in the same sector may lead to a negative impact on men's wage, while women do not experience any significant changes. The potential earning loss for men when an average woman begins working in their sectors is about 175 INR (Indian Rupee) per week [6]. The lack of impacts on women's wages may be due to the fact that

[6] We multiplied the coefficient of competition on the probability of a woman being employed (0.545) by the average wage for a female worker (322.3) to get this estimate.

they already face gender pay gaps and discrimination in the labor market. These results provide suggestive evidence that the negative impact of competition on gender attitudes may be drive by men's concerns over their own job prospects, rather than a general prejudice against women in the workplace.

2.7 Conclusion

Given evidence that women's labor supply affects attitudes towards women working (Field et al., 2021), we test whether the specific industries women enter affect the extent of support for women's work. In particular, we find that support for women's work in a given state is lower at times when the typical man faces more competition from women, as measured by the percent female of his industry: a one standard deviation (within-state) increase in competition leads to an 10.7 percentage points decrease in support for women's work. This effect is similar between men and women. We provide evidence that this result is not driven by reverse causality; there is no evidence that lagged support for women's work affects our measure of labor market competition, and results are very similar if we construct a Bartik-style measure of labor market competition (Bartik, 1991).

These results raise interesting questions for future research, such as the extent to which men's support for their own wife's work differs from their overall support for women's work, and if the degree of intra-household coooperation affect these opinions.

Our results also have implications for policy-makers interested in improving women's labor force participation and economic empowerment. While there is evidence that exposure to female peers or role models affects gender attitudes (Beaman et al., 2009; Dahl, Kotsadam and Rooth, 2021), our results caution that such exposure can backfire if men view women as in competition. This insight thus provides rationale for the existence of single-sex spaces, like schools.

Chapter 3

The effects of increasing community participation on school management outcomes: experimental evidence from India

with Natalia Cantet, Clara Delavallade, Alan Griffith, and Rebecca Thornton

3.1 Introduction

Recent decades have shown vast gains in school enrollment and attendance across many lower- and middle-income countries (United Nations, 2015). However, learning outcomes remain persistently poor in many places (World Bank, 2018). At the same time, links between outcomes and school management have been well documented (Azevedo et al., 2021; Bloom et al., 2015; Lemos, Muralidharan and Scur, 2021). A number of recent studies have highlighted the importance of *school management* or *governance* practices, with a prominent focus on personnel training (Beg, Fitzpatrick and Lucas, 2021; Kraft and Christian, 2022; de Barros et al., 2019) and teacher value-added (Grissom, Egalite and Lindsay, 2021; Chetty, Friedman and Rockoff, 2014)[1]. This research is particularly relevant given the ubiquity of school management programs worldwide [2], and they hold the potential to identify school-level interventions that lead to learning gains at the school level (Fryer, 2017; Grissom, Egalite and Lindsay, 2021).

At the same time, programs to enhance community-level school governance have gained traction in many developing countries. In India, where our study is based, School Management Com-

[1] An earlier set of prominent studies focused on the impacts of various inputs into the education production process, with mixed results (see, e.g., Glewwe et al. (2004); Glewwe, Kremer and Moulin (2009).

[2] Muralidharan and Singh (2020) document World Bank-funded school management programs in 84 countries.

mittees (SMCs) are mandated for each public school by the Right to Free and Compulsory Education Act of 2009. Each committee is made up of parents, educators, and community members, with mandatory representation for women and members of lower castes. Increasing the quantity and the quality of the interactions between the communities and the schools provides the latter with the mechanisms to improve school governance. First, it gives the community the opportunity to better oversee principal and teacher performance and to voice any concerns they may have. Second, it also offers parents and volunteers autonomy to make decisions that will directly affect the schools. Consequently, proponents of school-based management policies argue that it improves the provision and benefits from public services for poor schools (World Bank, 2004). However, little is known about the effects of SMCs (but see, Muralidharan and Singh (2020)).

We seek to fill this gap by providing causal evidence of the effects of an intervention intended to build capacity within School Management Committees. This intervention was implemented by an NGO working in rural Rajasthan and facilitated by local volunteers working under direction of the NGO's staff. The program was implemented for a period of six months during the 2012-13 school year. To evaluate the program, we implemented a cluster-randomized experiment across 229 primary schools in 98 villages.

Our results are as follows. First, we find an increase in SMC activities. Schools assigned to treatment had 17.6% more SMC meetings, a 32.5% increase in School Improvement Plans prepared, and a 37.8% increase in School Improvement Plans completed. Additionally, treated SMCs remain effective in the following school years, as the increase in the number of completed plans remains, on average, 37% higher than in control schools.

Second, we look at effects on school infrastructure and teacher counts. We find a 16.1 percentage point increase (20 percent) in the presence of a kitchen in the school after one year, but no significant effects on other measures of school infrastructure. We find large, significant increases in the number of teachers after both one year (13.7 percent) and two years (18.4 percent) of program implementation.

We also present estimates of effects on enrollment, but we caution that these effects may be at least partly due to simultaneous interventions such as enrollment drives. Additionally, we present instrumental variables and mediation analyses that suggest the enrollment gains may at least be partially due to increased SMC activities[3].

Our findings contribute to a growing literature documenting the effects of community and parental participation programs on school management outcomes. The closest work is Muralidharan and Singh (2020), who evaluate a different intervention aimed at making SMCs more effective, finding no impact on schools or learning outcomes. Also in India, Banerjee et al. (2010) find that offering parents information on school committees and training the community did not have an impact on community involvement or teacher effort in India. Glewwe and Maïga (2011) also do not find effects of management reforms on school outcomes in Madagascar [4].

The paper is organized as follows. In section 2, we describe the institutional background, highlighting the role of SMCs nationwide in Indian school governance. Section 3 describes the intervention, experimental design, and empirical strategy. Section 4 presents reduced-form results of the program on school management, infrastructure, and number of teachers, as well as instrumental variables results on the effects of SMC activities on infrastructure and teachers. Section 5 gives results on enrollment, including reduced-form effects on enrollment and a mediation analysis to explore the role of SMC activities in leading to enrollment gains. Section 6 concludes.

[3] Since the SMC intervention was bundled with other interventions directly aimed at improving learning outcomes, we are hesitant to make claims that this intervention itself led to the learning gains from the bundled intervention that we found in prior work (see Delavallade, Griffith and Thornton (2021))

[4] Additionally, a number of studies have found heterogeneity in the impact of the performance of similar programs in Mexico (Gertler, Patrinos and Rodríguez-Oreggia, 2012), Gambia (Blimpo, Evans and Lahire, 2015), Uganda (Barr et al., 2012), Indonesia (Pradhan et al., 2014) and Niger (Beasley and Huillery, 2015). In contrast, Lassibille et al. (2010) find that strengthening management practices had positive impacts on attendance and learning in Madagascar.

3.2 Background

3.2.1 School Governance

School governance varies worldwide. In developed countries, a decentralized and school-based approach is commonly adopted. For instance, in the US, each state is divided into multiple school districts, with a superintendent or school board overseeing the schools in each district. Additionally, there are Charter Schools that operate under a charter agreement, granting school mangers wide-ranging freedom in school management (De Grauwe, 2005).

The public education administration in India follows a more centralized and hierarchical structure. The Ministry of Education formulates policies and frameworks at the central level, while each state has its own Department of Education led by a Minister of Education. Within states, District Education Offices are responsible for overseeing schools in their respective districts (Muralidharan and Singh, 2020). Starting in 2009, every public school is mandated to establish a School Management Committee (SMC), serving as a bottom-up mechanism for monitoring and improving school functioning.

School-based management has gained traction as a policy approach to improve school governance and management in many developing countries. The World Bank has funded school management improvement programs in 84 countries (Muralidharan and Singh, 2020). For example, Indonesia implemented school-based management in 2003, granting authority to principals, teachers, and community members in making academic decisions (Pradhan et al., 2014). Similar programs have been implemented in countries like Mexico, Columbia, Niger, Madagascar, Sri Lanka (Bando, 2010; Beasley and Huillery, 2015; Glewwe and Maïga, 2011; Lassibille et al., 2010; Aturupane et al., 2022; Rodríguez, Sánchez and Armenta, 2010).

3.2.2 School Management Committees in India

School management committees (SMCs) were first introduced in India as a key element of the Right to Free and Compulsory Education Act (RTE) in 2009, with the aim of improving education quality across the country. Section 21 of the RTE Act mandates the formation of SMCs in all elementary government and government-aided schools.

SMCs are composed of elected representatives, including parents, headmasters, local authorities, and community members. A minimum of half the members must be women and there must be appropriate representation of Scheduled Castes/Scheduled Tribes. SMCs are required to hold meetings at least once a month, with a quorum set at one third of the total members.

SMCs monitor various aspects of the school's operations, including teacher attendance, building repairs, and the Mid-Day Meal (MDM) program[5] . The Mid-Day Meal (MDM) program, which is one of the world's largest school-feeding programs, serving an estimated 104.5 million children in 1.16 million schools in India in 2013-14[6].

Another important responsibility of the SMC is preparing and recommending an annual School Improvement Plan (SIP). This plan includes class-wise enrollment estimates, requirements for additional teachers, and development of necessary physical infrastructure such as separate girls' toilets and kitchen sheds. The SIP also outlines the financial needs for each year. These also mention the officials responsible for executing each specific task and for verifying that the task has been completed. Once the plan is finalized, it is signed by SMC members and submitted to the Block Elementary Education Officer (BEEO) before the end of the respective financial year. These improvement targets are set for the school to achieve incrementally.

Limited baseline data [7] suggests that, during the 2011-12 academic year—prior to the intervention—SMCs met on average 5.6 times per year (see Table A1). On average, they proposed 9.8

[5]As per Section 39 of the National Food Security Act, 2013
[6]Ministry of Human Resource Development, 2015
[7]These calculations are from our limited baseline data (177 of 229 schools in our sample) on SMC activities.

School Improvement Plans, of which 7.46 were completed (see Table A5).

Despite this progress, the effectiveness of these committees has been challenged by limited training provided by the state. For example, official trainings for SMCs only occur once a year and generally involve only three to four SMC members from each school.

SMCs have limited spending powers and face governance inefficiencies that further restrict their authorities. For example, in 2010-2011, these committees were allocated just about 5% of Sarva Shiksha Abhiyan (SSA) funds[8], and even these funds are subject to strict norms set by the government (ASER, 2011). Delays in the disbursement of school grants exacerbate the problem. For example, on average grants reach school bank account only during the second half of the fiscal year in late September. Expenditures are frequently based on orders from district and block officials, rather than being driven by the specific needs fo the schools themselves (PAISA, 2011).

In sum, limited training, insufficient financial autonomy, and government inefficiencies hinder SMCs effectiveness in fulfilling their intended role as a bottom-up system. Our intervention attempted to tackle these challenges by offering more timely and comprehensive training to all SMC members and assisting them in developing and implementing school improvement plans.

3.3 Intervention and Research Design

3.3.1 Intervention

Our study is conducted in the state of Rajasthan, located in northern India, with a population of 68.5 million in 2011. 75% of the population resides in rural areas. Rajasthan has a literacy rate of 67%, which is below national average of 74%[9]. About 60.2% of the children age between 6 to 14 attend public schools, while 35.1% are enrolled in private schools (ASER, 2011).

[8]Sarva Shiksha Abhiyan (SSA) is Government of India's flagship program for achievement of universalization of elementary education in a time bound manner, as mandated by 86th amendment to the Constitution of India making free and compulsory education to children of 6-14 years age group. It was launched in 2001-2002 in partnership with the State Governments.

[9]Data sources: Government of India, Census 2011.

The program we study is part of a larger, multi-faceted, bundled intervention developed and implemented by an Indian NGO, through a cooperation agreement with the Education Ministry. One of its main missions is to increase school enrollment and enhance learning outcomes, with a focus on girls in lower primary school (grades one through five). Initiated in 2011, the program consists of several components that separately target school management, enrollment, and learning.

The school management component of the intervention involves active engagement with school management committees (SMCs) and community volunteers in each village. Local volunteers, trained by NGO staff, worked closely with six SMC members per school throughout the school year[10]. SMC members received training and support to build capacity, foster parent engagement, formulate and implement annual school improvement plans (SIP), and sensitize the broader community to girls' education issues.

In addition to capacity-building for SMCs, the full program comprised two additional interventions designed to promote participation and learning. First, enrollment drives are conducted at the beginning of the school year to enroll and retain girls in school. Second, supplemental teaching are provided in three subjects (English, Hindi, and math) throughout the year to enhance students' learning process. For more detailed information on these two components and their effects at the individual level, refer to Delavallade, Griffith and Thornton (2021).

3.3.2 Experimental Design

The intervention was implemented in 229 schools within 98 villages. Villages were chosen based on the presence of at least one government primary school within four administrative blocks.

Selected villages were randomly assigned to either treatment or control groups (49 treatment and 49 control) by the researchers. Villages were stratified by administrative block, access to electricity, and the number of eligible primary schools within the village.

The random assignment resulted in 117 treatment schools and 112 control schools. Schools

[10] Academic years typically begin in May or June and continue until March

in treatment villages received a suite of programming provided by the NGO, while the control schools received no programming.

3.3.3 Data and Outcome Measures

To measure SMC activities, we collected data on three outcomes: the number of SMC meetings held, the number of School Improvement Plans (SIPs) prepared, and the number of SIPs implemented at each school. Our SMC activity data spans two academic years. In the first year (2012-13), we collected this data monthly from July 2012 to January 2013 for a total of seven observations per school. In the second year (2013-14), we collected this data for most months between May 2013 and March 2014.

Our analysis also focuses on school-level outcomes related to infrastructure and teacher recruitment, including the availability of drinking water, the presence of kitchens, boundary walls, and girls' toilets, as well as the number of teachers at each school. These data are collected by the NGO staff at three distinct time points: baseline in 2011, immediately after the program's completion at the end of 2012, and about one year later in March 2014. Additionally, we compile school-level enrollment figures from the school enrollment roster, which included information on the gender and age of each student.

3.3.4 Summary Statistics across Treatment Arms

In Table 3.1, we present descriptive statistics and balance tests of the pre-intervention school characteristics for the 2011-12 school year. Schools in the control group have, on average, 45.5 enrolled students, out of whom approximately 45% are male. Also, the teacher-pupil ratio equals 0.10. Almost all schools (more than 80 percent) have access to clean drinking water, but less than 40 percent have upper primary levels.

In Column 2 we present the regression coefficient testing the difference in means of baseline variables between the treatment and control schools. We do not find statistically significant differences between treatment and control schools across any measures of school infrastructure, teacher counts, or enrollment.

We have limited data on SMC activities at baseline, available for only 177 of the 229 schools in our sample. For these schools, SMCs in control schools held an average of 5.5 meetings, prepared 9.6 school improvement plans, and completed 7.1 of the plans. We find no statistically significant differences between treatment and control schools in terms of various types of SMC activities (Appendix Table A5).

3.3.5 Empirical Strategy

To measure the impact of the program on school-level SMC activity, we estimate Equation 3.1:

$$Y_{sj} = \beta_0 + \beta_1 T_j + \gamma' X_{sj} + \epsilon_{sj} \tag{3.1}$$

for school s in village j, while T_j is an indicator for village j being assigned to the treatment arm. The main dependent variables, Y_{sj}, captures three measurements of SMC activities: the number of meetings held, the number of school improvement plans prepared, and the number of school improvement plans completed. We include a vector of controls, X_{sj}, which contains baseline enrollment variables (number of girls, number of boys, average student ages), baseline school infrastructure (kitchen, drinking water, boundary wall, girls' toilet), and randomization strata fixed effects. We cluster standard errors by village—the unit of randomization—in all specifications.

Additionally, we estimate treatment effects on school infrastructure and the number of teachers to evaluate impacts on schooling quality and teaching engagement.

To estimate the effect of SMC activities on these outcomes, we employ a two-stage least squares strategy using the random treatment assignment as an instrumental variable for the SMC activities. In this context, Equation 3.2 comprises the first stage, while Equation 3.3 gives the second-stage

Table 3.1: Baseline Characteristics of Schools

	Control (1)	Difference (T-C) (2)	p-value (3)
Panel A: School Type			
Upper Primary School (UPS)	0.393	0.017	0.810
	(0.491)	0.072	
Panel B: Infrastructure			
Drinking Water	0.786	-0.059	0.489
	(0.412)	(-0.085)	
Kitchen	0.786	0.018	0.807
	(0.412)	(0.072)	
Boundary wall	0.643	0.024	0.779
	(0.481)	(0.085)	
Girls Toilet	0.866	-0.003	0.954
	(0.342)	(0.048)	
Electricity	0.312	-0.022	0.774
	(0.466)	(0.076)	
Panel C: Teacher			
No.of Teachers	3.250	-0.139	0.703
	(2.179)	(0.364)	
Panel D: Enrollment			
No. of Students Enrolled	45.509	-1.364	0.720
	(29.905)	(3.792)	
Percent Enrolled - girls	0.550	-0.018	0.392
	(0.177)	(0.021)	
Teacher/Student Ratio	0.092	-0.007	0.540
	(0.088)	(0.011)	
Observations	112	117	229
Joint Test (p-value)			.930

* Notes: 229 baseline sample schools (117 in Control, 112 in Treatment) are included. Column 1 presents the school-level average for control groups. Column 2 presents coefficient of a OLS regression of the indicated variable on treatment, and Column 3 reports the p-values for test of difference in means between Treatment and Control. Standard errors are clustered by village and reported in parentheses.

equation.

$$SMC_{sj} = \pi_0 + \pi_1 T_j + \gamma' X_{sj} + \epsilon_{sj} \tag{3.2}$$

$$Y_{sj} = \beta_0 + \beta_1 \widehat{SMC_{sj}} + \delta' X_{sj} + \epsilon_{sj} \tag{3.3}$$

Since we only have one instrument (T_j) for SMC activities, we estimate versions of Equation 3.3 with SMC_{sj} separately defined as number of meetings, number of school improvement plans prepared, and number of school improvement plans completed.

To evaluate the persistence of the effects, we examine the short-term results for immediately after the program ended in 2012, and the medium-term effects approximately two years after the intervention in 2014.

3.4 Results

This section first presents the effects of program on school-level SMC activities, infrastructure, and teacher recruitment. Then we present the 2SLS estimation results of SMC activities impact on these outcomes.

Impact on SMC Activities

We first show the program's effects on school management committee's activities in Table 3.2. During the seven months of data collection in 2012, treatment schools held an additional 0.796 committee meetings on average, a significant increase of 17.7 percent (Column 2, p-value=0.04). In addition, school committees prepared and completed 34.2 percent and 38.6 percent more improvement plans, respectively (Column 4,6). These results suggest that the program effectively increased the level of organization and planning among SMCs in the short term.

Appendix Figure A1 presents a monthly breakdown of the results. Notably, the effect on SMC meetings remained consistent over time, whereas the number of improvement plans prepared and completed slightly decreased towards the end of 2012.

To evaluate the persistence of the effects, we also examine SMC activities during the 2013-2014 school years, about two years after the intervention[11]. Panel B shows that from May 2013 to March 2014, treatment schools held an average of 0.47 additional committee meetings, and SMCs prepared 17.2 percent more improvement plans. However, these impacts were not statistically significant. SMCs completed 39 percent more improvement plans in 2013 (Column 6, p-value=0.026), indicating that the program had a lasting effect on implementing school improvement plans.

Table 3.2: Effect of Treatment on SMC Activities

	No.of SMC meeting		No.of SIP proposed		No.of SIP completed	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2012 - 13 school year						
Treatment	0.655**	0.796**	1.512	2.330**	1.171	1.853**
	(0.323)	(0.383)	(1.106)	(1.118)	(0.899)	(0.887)
Observations	229	229	229	229	229	229
R-squared	0.024	0.153	0.013	0.140	0.012	0.163
Control group mean	4.277	4.277	6.812	6.812	4.795	4.795
Panel B: 2013- 14 school year						
Treatment	0.510	0.471	1.685	1.572	1.931	2.213**
	(0.308)	(0.298)	(1.739)	(1.271)	(1.295)	(0.979)
Observations	227	227	227	227	227	227
R-squared	0.017	0.212	0.012	0.311	0.026	0.301
Control group mean	4.482	4.482	9.118	9.118	5.727	5.727
Strata FE	N	Y	N	Y	N	Y
Baseline Enrollment Controls	N	Y	N	Y	N	Y
Baseline Infrastructure Controls	N	Y	N	Y	N	Y

* Notes: The dependent variables for Panel A are cumulative SMC activities at the school-level from July 2012 to January 2013, and from May 2013 to March 2014 for Panel B. The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

[11] Two schools dropping out of data collection in 2014 caused a decrease in the number of observations in Panel B to 227.

Impact on School Infrastructure

Table 3.3 presents the treatment effects on school infrastructure, measured in January 2013 (Panel A) and March 2014 (Panel B). We control for baseline enrollment, baseline infrastructure, and strata fixed effects. The impact of treatment on school infrastructure was mixed, with only the installation of kitchen facilities showing a statistically significant effect, with treatment schools increasing the likelihood of having a kitchen installed by 16.1 percentage points over an average of 12.5 percent of control schools having a kitchen 0.125 (Panel A, Column 4). The school kitchen is essential for delivering the Mid-Day Meal (MDM) program[12].

The estimate for school boundary walls suggests a moderate but imprecisely measured effect, with a 3.9 percentage points (or 7 percent) increase (Panel A, Column 6). The program did not improve access to drinking water or gender-specific infrastructure such as girls' toilets, possibly because over 70 percent of the schools installed these facilities at baseline. The positive effects on infrastructures were minor and not statistically significant measured in March 2014 (Panel B), which suggests that the effects on infrastructure improvements may have diminished over time.

Impact on Teacher Recruitment

In addition to investment in school infrastructure, another crucial part of the SMC's responsibility is outlining the needs and recruiting additional teachers. We test the program's impacts on the total number of teachers and their gender breakdowns. As shown in Table 3.4, the program successfully increased the number of total teachers in both the 2012-13 and 2013-14 school years. On average, treatment schools hired 0.51 more teachers one year after the intervention, a 13.7 percent increase compared to control schools (Panel A, Column 2, p-value = 0.029). Hiring male teachers mainly drives this effect: an average of 0.43 male teachers (or 14 percent) were added to the treatment schools (Column 4). The number of female teachers also increased by about 10

[12] Section 39 of the National Food Security Act, 2013 mandates the school management committee (SMC) to monitor the implementation of the MDM program.

Table 3.3: Effect of Treatment on School Infrastructure

	Drinking Water		Kitchen		Boundary Wall		Girls Toilet	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 2012-13 school year								
Treatment	-0.081	-0.058	0.183**	0.161*	0.045	0.039	-0.026	-0.023
	(0.073)	(0.048)	(0.079)	(0.082)	(0.078)	(0.064)	(0.017)	(0.016)
Observations	229	229	229	229	229	229	229	229
R-squared	0.014	0.394	0.049	0.254	0.002	0.324	0.013	0.084
Control group mean	0.902	0.902	0.125	0.125	0.554	0.554	1.000	1.000
Panel B: 2013-14 school year								
Treatment	0.041	0.100	-0.015	0.022	0.020	0.034	0.003	-0.001
	(0.094)	(0.080)	(0.072)	(0.067)	(0.070)	(0.070)	(0.026)	(0.022)
Observations	227	227	227	227	227	227	227	227
R-squared	0.002	0.311	0.000	0.166	0.000	0.305	0.000	0.202
Control group mean	0.745	0.745	0.818	0.818	0.664	0.664	0.955	0.955
Strata FE	N	Y	N	Y	N	Y	N	Y
Baseline Enrollment Controls	N	Y	N	Y	N	Y	N	Y
Baseline Infrastructure Controls	N	Y	N	Y	N	Y	N	Y

* Notes: The dependent variables are measures of school-level infrastructure collected in January 2013 for Panel A, and in March 2014 for Panel B. The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

percent, but the coefficient is not statistically significant (Column 6). The effect was larger in the 2013-14 school year, with treatment schools hiring an additional 0.68 total teachers, of which 0.58 were males (Panel B, Column 2,4).

These results indicate that the program assisted schools in hiring more teachers, especially male teachers, in the short and medium term. The active student enrollment drive, which was a key component of the program (see Delavallade, Griffith and Thornton (2021) for details on the full intervention), may have increased total student enrollment and contributed to this effect. In the subsequent section, we employed a two-stage least squares (2SLS) approach to estimate the marginal effect of SMC activities on teacher recruitment. We also compare the magnitude of this effect to the overall treatment effect to gain a better understanding of the specific contribution of SMCs to these outcomes.

Table 3.4: Effect of Treatment on Teacher Counts

	Total Teachers		Male Teachers		Female Teachers	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2012 - 13 school year						
Treatment	0.225	0.513**	0.372	0.431**	-0.147	0.083
	(0.403)	(0.232)	(0.316)	(0.186)	(0.259)	(0.160)
Observations	229	229	229	229	229	229
R-squared	0.002	0.524	0.008	0.537	0.003	0.373
Control group mean	3.732	3.732	2.902	2.902	0.830	0.830
Panel B: 2013- 14 school year						
Treatment	0.422	0.675***	0.479	0.587***	-0.056	0.088
	(0.408)	(0.210)	(0.323)	(0.188)	(0.242)	(0.154)
Observations	227	227	227	227	227	227
R-squared	0.008	0.546	0.013	0.500	0.000	0.411
Control group mean	3.655	3.655	2.855	2.855	0.800	0.800
Strata FE	N	Y	N	Y	N	Y
Baseline Enrollment Controls	N	Y	N	Y	N	Y
Baseline Infrastructure Controls	N	Y	N	Y	N	Y

* Notes: The dependent variables are measures of total number of teachers collected in January 2013 for Panel A, and in March 2014 for Panel B. The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

Effect of SMC Activities on School Infrastructure and Teachers

The 2SLS results of the instrumental variable analysis in Table 3.5 suggest that SMC activities could directly impact school infrastructure and teacher recruitment. Specifically, completing one school improvement plan (SIP) is most effective in improving kitchen facilities, with 8.7 percentage points more likelihood by the end of 2012-13 school year (Column 2). We find no statistically significant effects of completed SIP on other school infrastructure. Various types of SMC activities, such as meetings and proposed SIP, yield similar results (Appendix Table A6, A7).

Completing one school improvement plan also significantly increases the total number of teachers by 0.28 and the number of male teachers by 0.23 (Panel A, Colum 5,6). These effects are smaller but still significant two years after the program (Panel B). The effect size of SMC activities alone is about half the overall treatment effect size on teacher recruitment (Table 4, Column 2). These findings highlight the importance of involving parents and community members in the SMCs and school improvement planning and implementation.

Table 3.5: Effect of SIP completed on Infrastructure and Teachers Counts (2SLS)

	Infrastructure				Number of Teachers		
	Drinking Water (1)	Kitchen (2)	Boundary Wall (3)	Girls Toilet (4)	Total Teachers (5)	Male Teachers (6)	Female Teachers (7)
Panel A: 2012 - 13 school year							
No.of SIP completed	-0.031	0.087**	0.021	-0.012	0.277*	0.232*	0.045
	(0.029)	(0.041)	(0.035)	(0.009)	(0.152)	(0.140)	(0.076)
Observations	229	229	229	229	229	229	229
Control group mean	0.902	0.125	0.554	1.000	3.732	2.902	0.830
Panel B: 2013- 14 school year							
No.of SIP completed	0.025	0.006	0.009	-0.000	0.168**	0.146**	0.022
	(0.021)	(0.016)	(0.016)	(0.005)	(0.077)	(0.067)	(0.037)
Observations	227	227	227	227	227	227	227
Control group mean	0.745	0.818	0.664	0.955	3.655	2.855	0.800

* Notes: The reported results are from 2SLS regressions using treatment as an instrument variable for SMC activities. In Panel A, the dependent variables are measures of school-level infrastructure and number of teachers, collected at the end of 2012-13 school year in January 2013, and the independent variable is total number of SIP completed during school year (July 2012 - January 2013).In Panel B, the dependent variables are measures of school-level infrastructure and number of teachers, collected at the end of 2013-14 school year in March 2014, and the independent variable is total number of SIP completed between July 2012 and March 2014.The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

3.5 Further Results and Robustness Checks

3.5.1 Effects on Enrollment

We further explore how the program affects school-level enrollment. It is important to note that the SMC intervention is part of a bundled intervention that includes an enrollment drive (see Delavallade, Griffith, and Thornton, 2021). While we caution against drawing causal inferences, these descriptive findings provide valuable insights into the program's effectiveness.

Table 3.6 suggests that the program positively impacted school-level enrollment, particularly for girls. In 2012-13, treatment schools retained an average of 8.1 more total students, of whom 3.25 were girls (Panel A, Column 2, 6). While this is an 8 percent increase, the effect is not statistically significant. In 2013-14, we see larger and statistically significant effects, with an additional 14.26 total students enrolled in treatment schools, representing a 17 percent increase compared to control schools (Panel B, Column 2). Among these additional enrolled students, 7.98 were girls and 6.28 were boys (Panel B, Column 4,6).

Enrollment is the first outcome where we observe a more substantial impact on girls than boys. This difference may be attributed to the targeted enrollment drive that specifically aimed to bring in girls who had dropped out or never enrolled before. These results are consistent with the findings on grade-level enrollment in Delavallade, Griffith and Thornton (2021). The positive impact on boys' enrollment suggests potential spillover benefits of the program.

In addition, we are interested in assessing the contribution of SMC activities to the observed positive impact on enrollment. To do so, we employed the same 2SLS approach and estimated the relationship between SMC activities and enrollment at the school level in Table 3.7. Our results show that completing one additional school improvement plan is associated with an increase of 3.55 total students and 1.98 girls enrolled in the treatment schools in the 2013-14 school year (Panel B, Column 1,3). Proposing a school improvement plan leads to a similar increase in enrollment (Appendix Table A4). Holding one SMC meeting resulted in the largest impact on the total enrollment, an increase of 11 students (Appendix Table A5). These findings suggest regular communication and collaboration between school officials and community members could effectively promote enrollment.

Table 3.6: Effect of Treatment on Enrollment

	Total Enrollment		Enrollment- Boys		Enrollment- Girls	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 2012 - 13 school year						
Treatment	2.542	8.101	3.244	4.845	-0.702	3.256
	(8.368)	(4.916)	(4.832)	(2.982)	(4.478)	(3.041)
Observations	229	229	229	229	229	229
R-squared	0.000	0.760	0.002	0.665	0.000	0.732
Control group mean	94.911	94.911	43.482	43.482	51.429	51.429
Panel B: 2013- 14 school year						
Treatment	9.944	14.261***	2.746	6.281*	7.198*	7.980***
	(7.856)	(4.796)	(4.343)	(3.345)	(4.253)	(2.301)
Observations	227	227	227	227	227	227
R-squared	0.006	0.744	0.002	0.689	0.010	0.697
Control group mean	81.936	81.936	43.682	43.682	38.255	38.255
Strata FE	N	Y	N	Y	N	Y
Baseline Enrollment Controls	N	Y	N	Y	N	Y
Baseline Infrastructure Controls	N	Y	N	Y	N	Y

* Notes: The dependent variables are measures of enrollment collected in January 2013 for Panel A, and in March 2014 for Panel B. The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.7: Effect of SMC activities on School-level Enrollment (2SLS)

	Total Enrollment (1)	Enrollment- Boys (2)	Enrollment- Girls (3)
Panel A: 2012 - 13 school year			
No.of SIP completed	4.373	2.615	1.757
	(2.903)	(1.662)	(1.708)
Observations	229	229	229
Control group mean	94.911	43.482	51.429
Panel B: 2013- 14 school year			
No.of SIP completed	3.550**	1.564	1.986***
	(1.562)	(0.955)	(0.749)
Observations	227	227	227
Control group mean	81.936	43.682	38.255

* Notes: The reported results are from 2SLS regressions using treatment as an instrument variable for SMC activities. In Panel A, the dependent variables are measures of school-level enrollment, collected at the end of 2012-13 school year in January 2013, and the independent variable is total number of SIP completed during school year (July 2012 - January 2013).In Panel B, the dependent variables are measures of school-level enrollment, collected at the end of 2013-14 school year in March 2014, and the independent variable is total number of SIP completed between July 2012 and March 2014.The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

3.5.2 Mediation Analysis

We conducted a mediation analysis to disentangle the direct effect of the treatment as well as the indirect effect mediated by SMC activities, following a set of regression Equations 3.4 outlined by Baron and Kenny (1986).

$$\begin{aligned} y_{sj} &= t_1 + cT_j + \epsilon_{y,j} \\ M_{sj} &= t_2 + \alpha T_j + \epsilon_{M,j} \\ y_{sj} &= t_3 + c^* T_j + \beta M_{sj} + \epsilon^*_{u,j} \end{aligned} \tag{3.4}$$

In the first equation, we regress outcome variables y_{sj} on the randomly assigned treatment variable T_j. The second equation regresses the mediator M_{sj} (SMC activities) on the treatment. The third equation regresses y_{sj} on randomly assigned T_j and observed M_{sj}. Together, the coefficient β represents the mediation effect from M_{sj} to y_{sj}.

The results revealed that the total number of completed SIPs substantially mediated the treatment effects on school-level teacher counts, particularly for female teachers (Appendix Table A10). Approximately 20% of the overall treatment effect on teachers and nearly 99% of the treatment effect on female teachers are mediated through completed SIPs (Column 5, 7). This moderate mediation effect is consistent with our earlier finding that the marginal effects of SMC activities contribute to about half the overall treatment effect on teacher recruitment using the instrumental variable approach as shown in Table 3.5.

Regarding school infrastructure, the mediation effect of SIPs was most pronounced in the short run for the installation of kitchen facilities and in the medium run for the construction of boundary wall, accounting for 17.4% and 73.5% of the total effect respectively (Column 2, 3). These findings underscore the potential of SMC activities to mediate our program's impacts, highlighting SMC's crucial role in driving positive changes to school infrastructure and teacher recruitment.

3.6 Conclusion

In this paper, we present the results of an intervention conducted in rural Rajasthan, India. The aim of the program was to build school management capacity and increase engagement among vulnerable communities.

The program successfully increased the number of school meetings and the number of school improvement plans created and completed in the first year of implementation. The effects of the program on the number of completed plans continues to the following year, which suggests that SMCs in treated schools become more effective in the long run. We also find limited reduced-from effects on school infrastructure as well as a large and significant effect on the number of teachers, an effect driven primarily by the presents of more male teachers in treated schools.

Studies have shown that successful school-based management improves both participation and learning (J-PAL Policy Bulletin, 2017). Therefore, our findings that the program under study led to increased SMC activities has important implications further down the causal chain. Finally, these findings are particularly important since bottom-up school management practices are commonly mandated by governemtns (as in India) and/or receive substantial funding from funding organizations (see Muralidharan and Singh (2020)).

Bibliography

Abebe, Girum, A Stefano Caria, Marcel Fafchamps, Paolo Falco, Simon Franklin, and Simon Quinn. 2021. "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City." *The Review of Economic Studies*, 88(3): 1279–1310.

Abel, Martin, Rulof Burger, and Patrizio Piraino. 2020. "The Value of Reference Letters: Experimental Evidence from South Africa." *American Economic Journal: Applied Economics*, 12(3): 40–71.

Abel, Martin, Rulof Burger, Eliana Carranza, and Patrizio Piraino. 2019. "Bridging the Intention-Behavior Gap? The Effect of Plan-Making Prompts on Job Search and Employment." *American Economic Journal: Applied Economics*, 11(2): 284–301.

Aizer, Anna. 2010. "The Gender Wage Gap and Domestic Violence." *American Economic Review*, 100(4): 1847–1859.

Aker, Jenny C, and Isaac M Mbiti. 2010. "Mobile Phones and Economic Development in Africa." *Journal of Economic Perspectives*, 24(3): 207–232.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015. "The Skill Complementarity of Broadband Internet." *The Quarterly Journal of Economics*, 130(4): 1781–1824.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the origins of gender roles: Women and the plough." *The quarterly journal of economics*, 128(2): 469–530.

Ardington, Cally, Anne Case, and Victoria Hosegood. 2009. "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa." *American Economic Journal: Applied Economics*, 1(1): 22–48.

Armona, Luis. 2021. “Online Social Network Effects in Labor Markets: Evidence From Facebook’s Entry to College Campuses.” *SSRN Electronic Journal*, 75.

ASER. 2011. “Annual Status of Education Report.” ASER Centre.

Aturupane, Harsha, Paul Glewwe, Tomoko Utsumi, Suzanne Wisniewski, and Mari Shojo. 2022. “The impact of Sri lanka’s school-based management programme on teachers’ pedagogical practices and student learning: evidence from a randomised controlled trial.” *Journal of Development Effectiveness*, 14(4): 285–305.

Autor, David H. 2001. “Wiring the Labor Market.” *Journal of Economic Perspectives*, 16.

Azevedo, Joao Pedro, Amer Hasan, Diana Goldemberg, Syedah Aroob Iqbal, and Koen Geven. 2021. “Simulating the Potential Impacts of COVID-19 School Closures on Schooling and Learning Outcomes: A Set of Global Estimates.” *The World Bank Research Observer*.

Bando, Rosangela. 2010. “The Effect of School Based Management on Parent Behavior and the Quality of Education in Mexico.” PhD diss. University of California, Berkeley.

Banerjee, Abhijit, and Gaurav Chiplunkar. 2020. “How Important Are Matching Frictions in the Labor Market? Experimental & Non-Experimental Evidence from a Large Indian Firm.” 52.

Banerjee, Abhijit V, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. “Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India.” *American Economic Journal: Economic Policy*, 2(1): 1–30.

Baron, Reuben M., and David A. Kenny. 1986. “The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations.” *Journal of Personality and Social Psychology*, 51: 1173–1182. Place: US Publisher: American Psychological Association.

Barr, Abigail, Frederick Mugisha, Pieter Serneels, and Andrew Zeitlin. 2012. "Information and collective action in the community monitoring of schools: Field and lab experimental evidence from Uganda."

Bartik, Timothy J. 1991. "Boon or Boondoggle? The Debate Over State and Local Economic Development Policies." *Upjohn Institute for Employment Research*, 17.

Beaman, Lori, and Jeremy Magruder. 2012. "Who Gets the Job Referral? Evidence from a Social Networks Experiment." *American Economic Review*, 102(7): 3574–3593.

Beaman, Lori, Niall Keleher, and Jeremy Magruder. 2018. "Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi." *Journal of Labor Economics*, 36(1): 121–157.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful women: does exposure reduce bias?" *The Quarterly journal of economics*, 124(4): 1497–1540.

Beasley, Elizabeth, and Elise Huillery. 2015. "Willing but Unable: Short-Term Experimental Evidence on Parent Empowerment and School Quality."

Beg, Sabrin, Anne Fitzpatrick, and Adrienne M Lucas. 2021. "Improving Public Sector Service Delivery: The Importance of Management."

Bhuller, Manudeep, Andreas Kostol, and Trond C. Vigtel. 2019. "How Broadband Internet Affects Labor Market Matching." *SSRN Electronic Journal*.

Blimpo, Moussa P., David K. Evans, and Nathalie Lahire. 2015. *Parental Human Capital and Effective School Management: Evidence from the Gambia. Policy Research Working Papers*, The World Bank.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, and John Van Reenen. 2015. "Does Management Matter in Schools?" *The Economic Journal*, 125(584): 647–674. Publisher: [Royal Economic Society, Wiley].

Bonnet, Florence, Joann Vanek, and Martha Chen. 2019. "Women and Men in the Informal Economy – A Statistical Brief." International Labour Office, Geneva.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2022. "Revisiting Event Study Designs: Robust and Efficient Estimation."

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica*, 82(5): 1671–1748.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott. 2020. "Misperceived social norms: Women working outside the home in Saudi Arabia." *American economic review*, 110(10): 2997–3029.

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, 225(2): 200–230.

Calvó-Armengol, Antoni, and Matthew O. Jackson. 2004. "The Effects of Social Networks on Employment and Inequality." *The American Economic Review*, 94(3): 426–454.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3): 414–427.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19.

Card, David, Jochen Kluve, and Andrea Weber. 2018. “What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations.” *Journal of the European Economic Association*, 16(3): 894–931.

Caria, Stefano, Simon Franklin, and Marc Witte. 2020. “Searching with Friends.” *SSRN Electronic Journal*.

Caria, Stefano, Theres Lessing, and Ann Hermes. 2019. “Filling the Gap: How Information Can Help Jobseekers.”

Carranza, Eliana, Robert Garlick, and Kate Orkin. 2020. “Job Search and Matching with Two-Sided Limited Information.”

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. “Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood.” *American Economic Review*, 104(9): 2633–2679.

Dahl, Gordon B, Andreas Kotsadam, and Dan-Olof Rooth. 2021. “Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams.” *The Quarterly journal of economics*, 136(2): 987–1030.

de Barros, Ricardo Paes, Mirela de Carvalho, Samuel Franco, Beatriz Garcia, Ricardo Henriques, and Laura Machado. 2019. “Assessment Of The Impact Of The Jovem De Futuro Program On Learning.”

de Chaisemartin, Clément, and Xavier D’Haultfœuille. 2020. “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects.” *American Economic Review*, 110(9): 2964–2996.

de Chaisemartin, Clément, and Xavier D’Haultfœuille. 2022. “Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey.” NBER Working Paper 29691.

De Grauwe, Anton. 2005. "Improving the Quality of Education through School-Based Management: Learning from International Experiences." *International Review of Education*, 51(4).

Delavallade, Clara, Alan Griffith, and Rebecca Thornton. 2021. "Effects of a Multi-Faceted Education Program on Enrollment, Learning and Gender Equity: Evidence from India." *The World Bank Economic Review*, 35(4): 950–968.

Dettling, Lisa. 2017. "Broadband and the Labor Market: The Impact of Redidential High-speed Internet on Married Women's Labor Force Participation." *ILR Review*, 70(2): 451–482.

Diamond, Peter A. 1982. "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy*, 90(5): 881–894.

Fernández, Raquel. 2014. "Women's rights and development." *Journal of Economic Growth*, 19: 37–80.

Fernández, Raquel, and Alessandra Fogli. 2009. "Culture: An empirical investigation of beliefs, work, and fertility." *American economic journal: Macroeconomics*, 1(1): 146–177.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore. 2021. "On her own account: How strengthening women's financial control impacts labor supply and gender norms." *American Economic Review*, 111(7): 2342–2375.

Franklin, Simon. 2018. "Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies." *The Economic Journal*, 128(614): 2353–2379.

Fryer, Roland G. 2017. "Management and Student Achievement: Evidence from a Randomized Field Experiment." *NBER Working Paper Series*.

Gee, Laura K., Jason Jones, and Moira Burke. 2017. "Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network." *Journal of Labor Economics*, 35(2): 485–518.

Gertler, Paul, Harry Anthony Patrinos, and Eduardo Rodríguez-Oreggia. 2012. "Parental Empowerment In Mexico: Randomized Experiment Of The "Apoyos A La Gestion Escolar (Age)" In Rural Primary Schools In Mexico."

Glewwe, Paul, and Eugenie W. H. Maïga. 2011. "The impacts of school management reforms in Madagascar: do the impacts vary by teacher type?" *Journal of Development Effectiveness*, 3(4): 435–469. Publisher: Routledge.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics*, 1(1): 112–135.

Glewwe, Paul, Michael Kremer, Sylvie Moulin, and Eric Zitzewitz. 2004. "Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya." *Journal of Development Economics*, 74(1): 251–268.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics*, 225(2): 254–277.

Goyal, Aparajita. 2010. "Information, Direct Access to Farmers, and Rural Market Performance in Central India." *American Economic Journal: Applied Economics*, 2(3): 22–45.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology*, 78(6): 1360–1380.

Grissom, Jason, Anna Egalite, and Constance Lindsay. 2021. "How Principals Affect Students and Schools." New York: The Wallace Foundation.

Hagedorn, Marcus, and Iourii Manovskii. 2008. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited." *American Economic Review*, 98(4): 1692–1706.

Heath, Rachel. 2018. "Why Do Firms Hire Using Referrals? Evidence from Bangladeshi Garment Factories." *Journal of Political Economy*, 126(4): 1691–1746.

Heath, Rachel, and Seema Jayachandran. 2017. "The causes and consequences of increased female education and labor force participation in developing countries." In *The Oxford Handbook of Women and the Economy*. 345–367. Oxford University Press.

Heath, Rachel, and Xu Tan. 2020. "Intrahousehold Bargaining, Female Autonomy, and Labor Supply: Theory and Evidence from India." *Journal of the European Economic Association*, 18(4): 1928–1968.

Hjort, Jonas, and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review*, 109(3): 1032–1079.

Jack, William, and Tavneet Suri. 2014. "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution." *American Economic Review*, 104(1): 183–223.

Jayachandran, Seema. 2017. "Fertility decline and missing women." *American Economic Journal: Applied Economics*, 9(1): 118–139.

Jayachandran, Seema. 2021. "Social norms as a barrier to women's employment in developing countries." *IMF Economic Review*, 69(3): 576–595.

Jensen, Robert. 2007. "The Digital Provide: Information (Technology), Market Performance, and Welfare in The." *Source: The Quarterly Journal of Economics*, 122(3): 879–924.

J-PAL Policy Bulletin. 2017. "Roll Call: Getting Children into School." Abdul Latif Jameel Poverty Action Lab., Cambridge, MA.

Klonner, Stefan, and Patrick Nolen. 2008. "Does ICT Bene…t the Poor? Evidence from South Africa." 37.

Kraft, Matthew A., and Alvin Christian. 2022. "Can Teacher Evaluation Systems Produce High-Quality Feedback? An Administrator Training Field Experiment." *American Educational Research Journal*, 59(3): 500–537.

Kramarz, Francis, and Oskar Nordström Skans. 2007. "With a Little Help from My Parents ? Family Networks and Youth Labor Market Entry." 39.

Kroft, Kory, and Devin G. Pope. 2014. "Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist." *Journal of Labor Economics*, 32(2): 259–303.

Kuhn, Peter, and Hani Mansour. 2014. "Is Internet Job Search Still Ineffective?" *The Economic Journal*, 124(581): 1213–1233.

Kuhn, Peter, and Mikal Skuterud. 2004. "Internet Job Search and Unemployment Durations." *American Economic Review*, 94(1): 218–232.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*, 4(4): 1–31.

Lassibille, Gérard, Jee-Peng Tan, Cornelia Jesse, and Trang Van Nguyen. 2010. "Managing for Results in Primary Education in Madagascar: Evaluating the Impact of Selected Workflow Interventions." *The World Bank Economic Review*, 24(2): 303–329. Publisher: Oxford University Press.

Leavitt, Keith, Luke (Lei) Zhu, Anthony Klotz, and Maryam Kouchaki. 2022. "Fragile or robust? Differential effects of gender threats in the workplace among men and women." *Organizational Behavior and Human Decision Processes*, 168: 104112.

Lemos, Renata, Karthik Muralidharan, and Daniela Scur. 2021. "Personnel Management and SchoolProductivity: Evidence from India." *RISE Working Paper Series*.

Lowe, Matt, and Madeline McKelway. 2021. "Coupling labor supply decisions: An experiment in india."

McCasland, Jamie, and Morgan Hardy. forthcoming. "Are Small Firms Labor Constrained? Experimental Evidence from Ghana." *American Economic Journal: Applied Economics*.

McKelway, Madeline. 2020. "Women's employment in India: Intra-household and intra-personal constraints." *Technical Report*.

McKelway, Madeline. 2023. "Experimental Evidence on the Effects of Women's Employment." *Unpublished manuscript*.

Michaels, Guy, Ashwini Natraj, and John Van Reenen. 2014. "Has ICT Polarized Skill Demand? Evidence from Eleven Countries Over Twenty-five Years." *The Review of Economics and Statistics*, 96(1): 60–77.

Montgomery, James D. 1991. "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis." *The American Economic Review*, 81(5): 1408–1418.

Mortensen, Dale T. 1986. "Chapter 15 Job Search and Labor Market Analysis." In *Handbook of Labor Economics*. Vol. 2, 849–919. Elsevier.

Muralidharan, Karthik, and Abhijeet Singh. 2020. "Improving Public Sector Management at Scale? Experimental Evidence on School Governance India." *National Bureau of Economic Research*.

Olivetti, Claudia, Eleonora Patacchini, and Yves Zenou. 2020. "Mothers, peers, and gender-role identity." *Journal of the European Economic Association*, 18(1): 266–301.

PAISA. 2011. "PAISA District Studies:Towards a New Frontier for Governing Elementary Education Finances in India." Center for Policy Research.

Pellizzari, Michele. 2010. "Do Friends and Relatives Really Help in Getting A Good Job?" *Industrial and Labor Relations Review*, 63(3): 494–510.

Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." *The American Economic Review*, 75(4): 676–690.

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, Armida Alisjahbana, and Rima Prama Artha. 2014. "Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia." *American Economic Journal: Applied Economics*, 6(2): 105–126.

Rodríguez, Catherine, Fabio Sánchez, and Armando Armenta. 2010. "Do Interventions at School Level Improve Educational Outcomes? Evidence from a Rural Program in Colombia." *World Development*, 38(3): 415–428.

Sapin, Marlène, Dominique Joye, Christof Wolf, and in collaboration with Johannes Andersen, Yanjie Bian, Ali Carkoglu, Yang-Chi Fu, Ersin Kalaycioglu, Peter V. Marsden, and Tom W. Smith. 2020. "The ISSP 2017 Social Networks and Social Resources Module." *International Journal of Sociology*, 50(1): 1–25.

Schuler, Sidney Ruth, and Farzana Islam. 2008. "Women's acceptance of intimate partner violence within marriage in rural Bangladesh." *Studies in family planning*, 39(1): 49–58.

Smith, Aaron. 2015. "Searching for Work in the Digital Era." Pew Research Center.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, 225(2): 175–199.

United Nations. 2015. "The Millennium Development Goals Report." New York: United Nations.

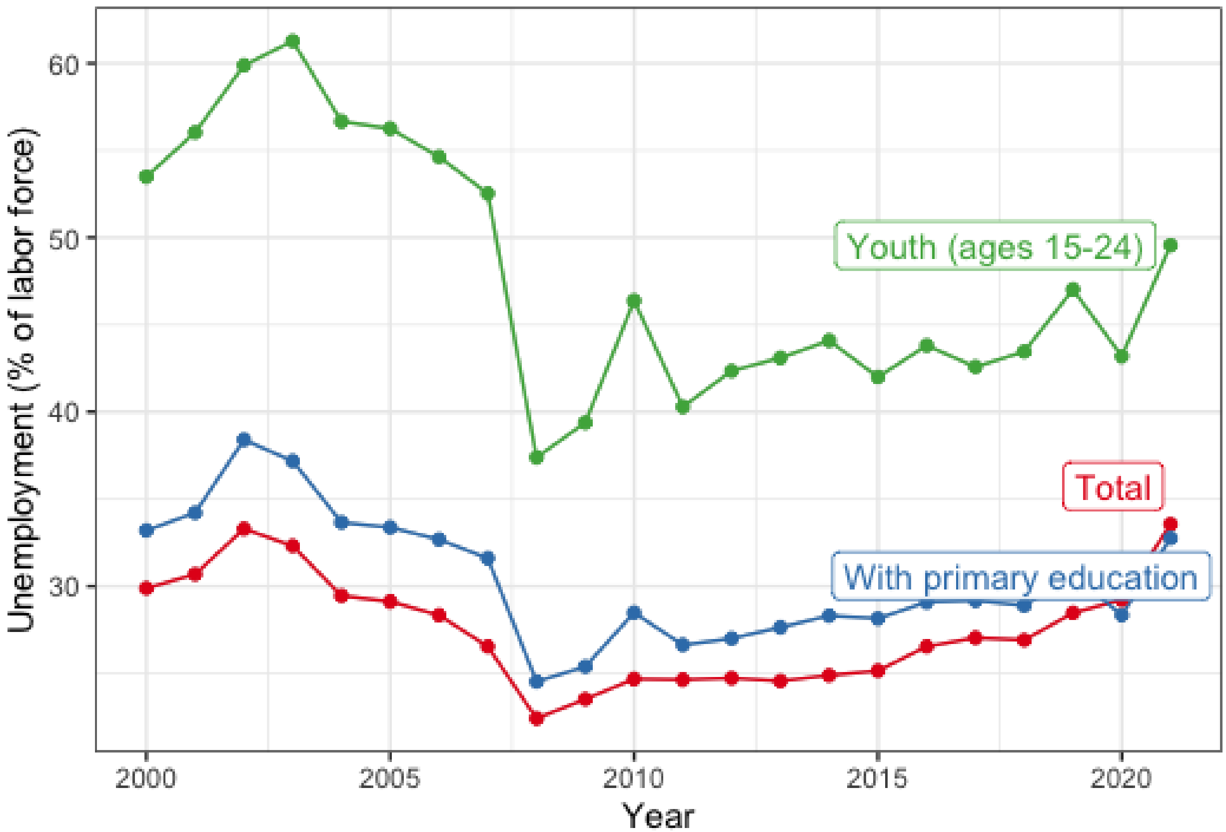
Wheeler, Laurel E, Robert Garlick, Eric Johnson, Patrick Shaw, and Marissa Gargano. 2022. "Linkedin(to) Job Opportunities: Experimental Evidence from Job Readiness Training." *American Economic Journal: Applied Economics*.

World Bank. 2004. "World Development Report 2004: Making Services Work for Poor People." The World Bank, Washington,DC.

World Bank. 2018. "World Development Report 2018: Learning to realize education's promise." The World Bank, Washington,DC.

Appendix

Figure A1: Unemployment Rate in South Africa



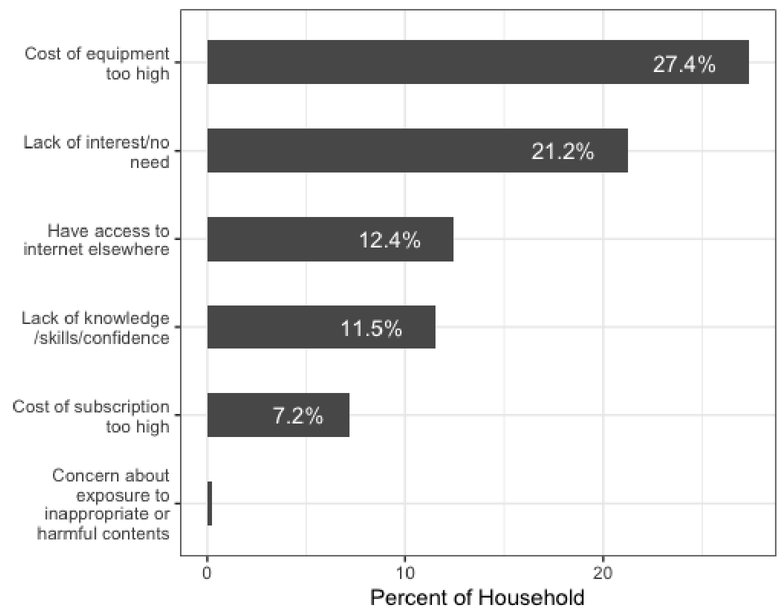
Source: World Development Indicators.

Table A1: Impacts of Internet Connection on Job Outcomes by Education

Outcome	Employed		Income		No.of Methods		Online		Network	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
% connected	0.081	0.054	0.133	-0.268	-0.082	0.015	-0.010	-0.030	-0.016	0.008
	(0.075)	(0.090)	(0.453)	(0.653)	(0.061)	(0.068)	(0.034)	(0.043)	(0.074)	(0.075)
... $\times$ beyond primary	0.074	0.126	0.999	1.374*	-0.044	-0.024	0.118***	0.112**	-0.019	0.016
	(0.071)	(0.084)	(0.624)	(0.686)	(0.075)	(0.078)	(0.038)	(0.045)	(0.049)	(0.049)
... $\times$ w. educated parents		-0.134***		-0.969**		-0.189***		0.122***		-0.283***
		(0.041)		(0.457)		(0.064)		(0.045)		(0.055)
beyond primary	0.163***	0.163***	1.524***	1.666***	0.186***	0.166***	0.090***	0.099***	0.094***	0.079***
	(0.010)	(0.013)	(0.101)	(0.134)	(0.017)	(0.018)	(0.006)	(0.007)	(0.010)	(0.011)
Mean of outcome	0.367	0.399	2.356	2.685	0.258	0.254	0.062	0.067	0.248	0.265
Observations	36,892	24,032	32,278	21,050	37,485	24,151	32,724	21,335	32,725	21,336
R-squared	0.147	0.137	0.168	0.173	0.031	0.034	0.060	0.069	0.053	0.051
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Location FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

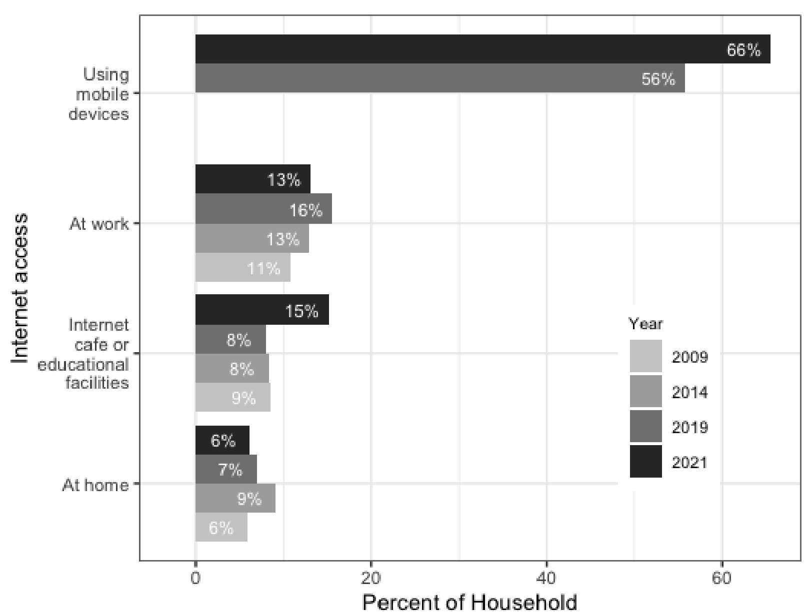
[*] Notes: Only workers between age 15 and 65 are included. All specifications include year, location fixed effects, age and gender control variables. Standard errors (in parentheses) are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A2: Reasons for not having Internet access at home



Source: General Household Survey, 2018, Statistics South Africa.

Figure A3: Households' Internet Access by Place of Access



Source:General Household Survey, Statistics South Africa.

Table A2: South Africa ICT access survey

	2017-2018	2011-2012	2005-2008
Panel A: household attributes			
HH has internet connection	11.6%	16.2%	6.5%
HH with Internet: highest education level			
No school	0.9%		
Primary	1.4%		
Secondary and above	97.6%		
Reasons not having internet			
Cost too high	48.3%		
Not available in the area	5.9%		
Do not need	20.4%		
Do not know how to use it	12.4%		
Others	12.9%		
Panel B: Individual usage			
Used Internet before	68.6%	33.0%	18.6%
Internet usage			
Once a day	50.4%	64.8%	64.4%
Once a week	30.8%	24.6%	24.9%
Once a month	10.3%	9.1%	7.0%
Less than once a month	8.6%	1.5%	3.6%
Most important internet activity			
Social networking	44.5%		
Education	23.5%		
Job search	12.4%		
Work related	11.3%		
Online banking	2.5%		
Others	5.7%		
Limitation for use of the internet (multiple responses)			
Cost	46.6%	62.9%	45.3%
Speed	25.6%	10.1%	8.8%
No interesting content in my language	7.1%		13.0%
Difficult to use	2.8%	73.2%	1.5%
Reason not using internet(single choice)			
Cost	50.6%		
No interest	19.3%		
Do not know how to use it	8.9%		
Not available in my area	3.4%		
Others	17.9%		

Source: Africa ICT access survey.

Table A3: Survey round mapping

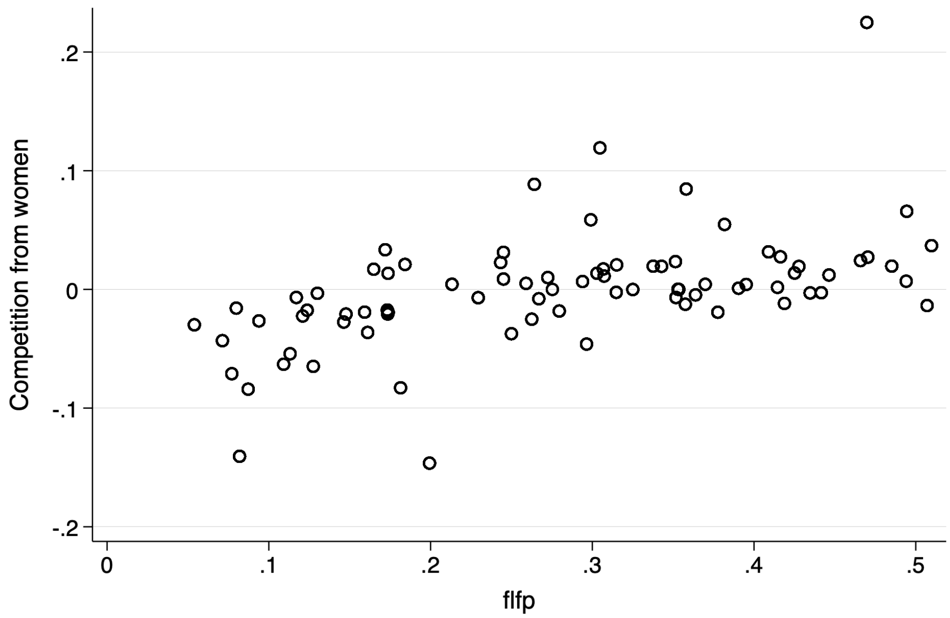
	Test for Reverse Causality	
Main Estimating Equation		
World Values Survey (attitudes)	National Sample Survey (labor outcomes)	Lagged World Values Survey (lagged attitudes)
1990	1987	–
1995	1993	1990
2001	1999	1995
2006	2004	2001
2012	2009	2006

Figure A4: Principle Usual Activity Status (PUAS) Codes in NSS

CATEGORIES

Value	Category
00	NR
11	Worked in h.h. enterprise (self-employed): own account worker
12	Employer
21	Worked as helper in h.h. enterprise (unpaid family worker)
31	Worked as regular salaried/ wage employee
41	Worked as casual wage labour : in public works
51	Worked as casual wage labour : In other types of work
81	Did not work but was seeking and/or available for work
91	Attended educational institution
92	Attended domestic duties only
93	Attended domestic duties and was also engaged in free collection of goods (vegetables, roots, firewood, cattle feed, etc.), sewing, tailoring, weaving, etc. for household use
94	Rentiers, pensioners , remittance recipients, etc.
95	Not able to work due to disability
97	Others (including begging, prostitution, etc.)
99	Children 0 - 4 age -group

Figure A5: Correlation between competition and female labor force participation



Note: The scatter plot shows the residuals controlling for state and year fixed effects.

Figure A6: Histogram of Competition Variable

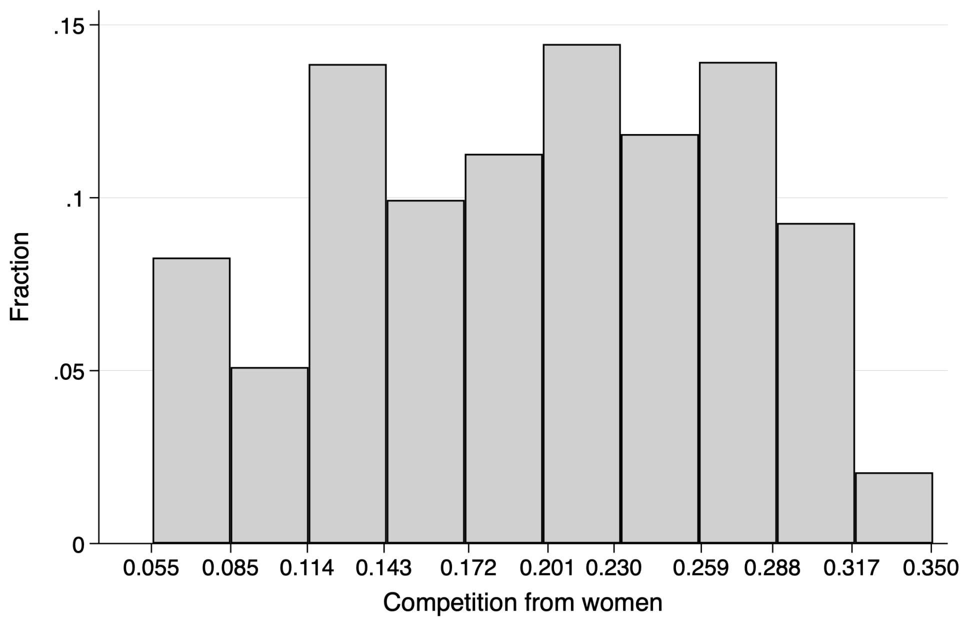


Table A4: Effects of competition on wage

	Wage (asinh)		
	(1)	(2)	(3)
Competition from women × Male	-8.351***	-8.806***	-3.776
	(2.887)	(3.192)	(2.411)
Competition from women × Female	0.797	-0.847	3.578
	(4.107)	(4.452)	(3.533)
Mean of Dep. Variable	3.901	3.234	3.379
Male Mean	4.055	4.055	4.055
Female Mean	3.442	3.442	3.442
Observations	703,065	575,088	539,831
R-squared	0.553	0.474	0.592
Control: individual		Y	Y
Control: sector			Y

Notes: The dependent variable is the inverse hyperbolic sine of wage at individual level. All regressions include state and year fixed effects. Column 2 control for age and education, and column 3 control for sectors additionally. Robust standard errors clustered at state-year level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A7: Effects of Competition on Wage by Sectors

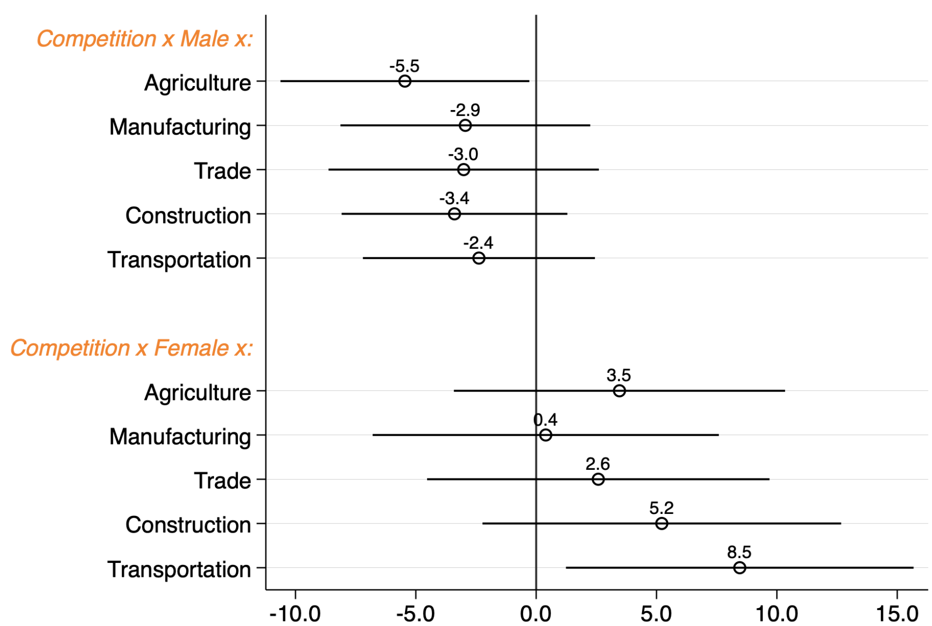


Table A5: Baseline SMC activities

	Control	Difference (T-C)	p-value (test of equality)
	(1)	(2)	(3)
Panel E: SMC activities in 2011			
No.of SMC meeting	5.515	0.202	0.713
	(3.093)	(0.547)	
No.of SIP planned	9.641	0.481	0.792
	(10.925)	(1.822)	
No.of SIP completed	7.146	0.746	0.62
	(9.604)	(1.501)	
Observations	103	74	177

* Notes: Baseline SMC activities information was available for only 177 schools. Column 1 presents the school-level average for control groups. Column 2 presents coefficient of a OLS regression of the indicated variable on treatment, and Column 3 reports the p-values for test of difference in means between Treatment and Control. Standard errors are clustered by village and reported in parentheses.

Table A6: Effect of SMC meetings on Infrastructure and Teachers Counts (2SLS)

	Infrastructure				Number of Teachers		
	Drinking Water (1)	Kitchen (2)	Boundary Wall (3)	Girls Toilet (4)	Total Teachers (5)	Male Teachers (6)	Female Teachers (7)
Panel A: 2012 - 13 school year							
No.of SMC meeting	-0.072	0.202	0.050	-0.029*	0.645*	0.541*	0.104
	(0.054)	(0.125)	(0.074)	(0.017)	(0.337)	(0.298)	(0.183)
Observations	229	229	229	229	229	229	229
Control group mean	0.902	0.125	0.554	1.000	3.732	2.902	0.830
Panel B: 2013- 14 school year							
No.of SMC meeting	0.077	0.017	0.026	-0.001	0.522*	0.453**	0.068
	(0.072)	(0.051)	(0.050)	(0.016)	(0.268)	(0.231)	(0.116)
Observations	227	227	227	227	227	227	227
Control group mean	0.745	0.818	0.664	0.955	3.655	2.855	0.800

* Notes: The reported results are from 2SLS regressions using treatment as an instrument variable for SMC activities. In Panel A, the dependent variables are measures of school-level infrastructure and number of teachers, collected at the end of 2012-13 school year in January 2013, and the independent variable is total number of SMC meetings during school year (July 2012 - January 2013).In Panel B, the dependent variables are measures of school-level infrastructure and number of teachers, collected at the end of 2013-14 school year in March 2014, and the independent variable is total number of SMC meetings between July 2012 and March 2014.The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Effect of SIP Proposed on Infrastructure and Teachers Counts (2SLS)

	Infrastructure				Number of Teachers		
	Drinking Water (1)	Kitchen (2)	Boundary Wall (3)	Girls Toilet (4)	Total Teachers (5)	Male Teachers (6)	Female Teachers (7)
Panel A: 2012 - 13 school year							
No.of SIP proposed	-0.025	0.069**	0.017	-0.010	0.220*	0.185*	0.035
	(0.024)	(0.034)	(0.027)	(0.008)	(0.122)	(0.108)	(0.062)
Observations	229	229	229	229	229	229	229
Control group mean	0.902	0.125	0.554	1.000	3.732	2.902	0.830
Panel B: 2013- 14 school year							
No.of SIP proposed	0.026	0.006	0.009	-0.000	0.177*	0.154*	0.023
	(0.023)	(0.017)	(0.017)	(0.005)	(0.096)	(0.082)	(0.040)
Observations	227	227	227	227	227	227	227
Control group mean	0.745	0.818	0.664	0.955	3.655	2.855	0.800

* Notes: The reported results are from 2SLS regressions using treatment as an instrument variable for SMC activities. In Panel A, the dependent variables are measures of school-level infrastructure and number of teachers, collected at the end of 2012-13 school year in January 2013, and the independent variable is total number of SIP planned during school year (July 2012 - January 2013).In Panel B, the dependent variables are measures of school-level infrastructure and number of teachers, collected at the end of 2013-14 school year in March 2014, and the independent variable is total number of SIP planned between July 2012 and March 2014.The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Effect of SMC meetings on School-level Enrollment (2SLS)

	Total Enrollment (1)	Enrollment- Boys (2)	Enrollment- Girls (3)
Panel A: 2012 - 13 school year			
No.of SMC meeting	10.180	6.089*	4.091
	(6.282)	(3.455)	(3.919)
Observations	229	229	229
Control group mean	94.911	43.482	51.429
Panel B: 2013- 14 school year			
No.of SMC meeting	11.019**	4.853	6.166**
	(5.571)	(3.104)	(2.923)
Observations	227	227	227
Control group mean	81.936	43.682	38.255

[*] Notes: The reported results are from 2SLS regressions using treatment as an instrument variable for SMC activities. In Panel A, the dependent variables are measures of school-level enrollment, collected at the end of 2012-13 school year in January 2013, and the independent variable is total number of SMC meetings during school year (July 2012 - January 2013).In Panel B, the dependent variables are measures of school-level enrollment, collected at the end of 2013-14 school year in March 2014, and the independent variable is total number of SMC meetings between July 2012 and March 2014.The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Effect of SIP proposed on School-level Enrollment (2SLS)

	Total Enrollment (1)	Enrollment- Boys (2)	Enrollment- Girls (3)
Panel A: 2012 - 13 school year			
No.of SIP proposed	3.477	2.079*	1.397
	(2.189)	(1.259)	(1.318)
Observations	229	229	229
Control group mean	94.911	43.482	51.429
Panel B: 2013- 14 school year			
No.of SIP proposed	3.733**	1.644	2.089**
	(1.842)	(1.062)	(0.927)
Observations	227	227	227
Control group mean	81.936	43.682	38.255

* Notes: The reported results are from 2SLS regressions using treatment as an instrument variable for SMC activities. In Panel A, the dependent variables are measures of school-level enrollment, collected at the end of 2012-13 school year in January 2013, and the independent variable is total number of SIP proposed during school year (July 2012 - January 2013).In Panel B, the dependent variables are measures of school-level enrollment, collected at the end of 2013-14 school year in March 2014, and the independent variable is total number of SIP proposed between July 2012 and March 2014.The baseline enrollment controls include the number of girls, number of boys, and average age of students enrolled at baseline (in grades 3-5, in 2011). School infrastructure controls are indicator variables for whether the school had a kitchen, a border wall, running water, and girls' toilet at baseline. Robust standard errors in parentheses, clustered by village. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Mediation Analysis

	Infrastructure				Number of Teachers		
	Drinking Water (1)	Kitchen (2)	Boundary Wall (3)	Girls Toilet (4)	Total (5)	Male (6)	Female (7)
Panel A: 2012 - 13 school year							
direct effect of Treatment	-0.059	0.133	0.043	-0.024	0.410	0.409	0.000
	(0.043)	(0.056)	(0.065)	(0.017)	(0.258)	(0.224)	(0.160)
indirect effect through SIP compeleted	0.001	0.028	-0.004	0.001	0.103	0.021	0.082
	(0.007)	(0.015)	(0.011)	(0.003)	(0.061)	(0.038)	(0.044)
total effect	-0.058	0.161	0.039	-0.023	0.513	0.431	0.083
	(0.043)	(0.056)	(0.064)	(0.017)	(0.258)	(0.221)	(0.161)
proportion of total effect mediated	-0.017	0.174	-0.103	-0.043	0.201	0.049	0.988
Panel B: 2013 - 14 school year							
direct effect of Treatment	0.104	0.025	0.009	-0.008	0.554	0.525	0.028
	(0.058)	(0.059)	(0.064)	(0.030)	(0.258)	(0.243)	(0.161)
indirect effect through SIP compeleted	-0.004	-0.003	0.025	0.007	0.121	0.061	0.060
	(0.013)	(0.013)	(0.016)	(0.007)	(0.069)	(0.058)	(0.041)
total effect	0.100	0.022	0.034	-0.001	0.675	0.587	0.088
	(0.056)	(0.057)	(0.063)	(0.029)	(0.254)	(0.238)	(0.158)
proportion of total effect mediated	-0.040	-0.136	0.735	-7.000	0.179	0.104	0.682

* Notes: The reported results are from mediation analysis of treatment impacts on school infrastructures and teacher numbers, using the number of SIP completed as the mediator. In Panel A, the dependent variables are measures of school-level infrastructure and number of teachers, collected at the end of 2012-13 school year in January 2013. In Panel B, the dependent variables are collected at the end of 2013-14 school year in March 2014. All regressions include strata fixed effects, baseline enrollment controls, and baseline school infrastructure controls. Bootstrap standard errors in parentheses.

Chapter A

A model of jobseeker's utility maximization with leisure

I include leisure in the utility function for jobseekers in this version of the conceptual model. A jobseeker lives two periods with a supply of Internet access θ. In the first period, an unemployed individual receives some unemployment benefit b, and needs to allocates his time (normalized to 1) between job searching s and leisure l. The probability of finding a job depends on the search effort and amount of Internet access: $p(s,\theta)$. In the second period, if the individual becomes employed, the wage and labor supplied will be given as w and h. The jobseeker chooses job search effort and maximizes the expected lifetime utility as follows:

$$\begin{aligned} \max_{s} \quad & u\left(c_1, \ell_1, \theta\right) + \beta E u\left(c_2, \ell_2, \theta\right) \\ \text{s.t.} \quad & c_1 = b \\ & \ell_1 = 1 - s \\ & c_2 = \begin{cases} wh & \text{w.p. } p(s,\theta) \\ b & \text{w.p. } 1 - p(s,\theta) \end{cases} \\ & \ell_2 = \begin{cases} 1-h & \text{w.p. } p(s,\theta) \\ 1 & \text{w.p. } 1 - p(s,\theta) \end{cases} \\ & 0 \leq s, p(s,\theta) \leq 1 \end{aligned} \tag{A1}$$

An interior solution should satisfy the following first order condition:

$$\frac{\partial u\left(b, 1-s, \theta\right)}{\partial l_1} = \beta \frac{\partial p(s,\theta)}{\partial s}\left[u\left(wh, 1-h, \theta\right) - u\left(b, 1, \theta\right)\right] \tag{A2}$$

which implies that the individual chooses search effort s optimally such that the marginal utility of giving up leisure equals the expected utility gain from searching for work, which is the difference between employment and unemployment utility in the second period.

For this paper, I am interested in how employment probability may change with the Internet access, which is provided exogenously. That is,

$$\frac{d}{d\theta}p(s(\theta),\theta) = \frac{\partial p}{\partial s}s'(\theta) + \frac{\partial p}{\partial \theta} \tag{A3}$$

Assuming the marginal productivity of search and Internet are both positive ($\frac{\partial p}{\partial s}, \frac{\partial p}{\partial \theta} > 0$), the effect on employment will depend on $s'(\theta)$. In order to see how optimal search effort $s^\star(\theta)$ changes with Internet access θ, we can differentiate the first order condition equation A2 with respect to θ:

$$s'(\theta) = \frac{\beta p_{s\theta}\left(u^{emp} - u^{unemp}\right) + \beta p_s \frac{\partial}{\partial \theta}\left(u^{emp} - u^{unemp}\right) - u_{\ell\theta}}{-u^1_{\ell\ell} - \beta p_{ss}\left(u^{emp} - u^{unemp}\right)} \tag{A4}$$

where u^1, u^{emp}, u^{unemp} represent the utility in period 1, being employed and unemployed in period 2 respectively.

Since $u^{emp} > u^{unemp}$ is a necessary condition for the existence of an interior solution, the denominator in equation A4 is positive. The sign of the numerator depends on three parts. First, $p_{s\theta}$, the change in the marginal productivity of search in response to more Internet access. Second, $\frac{\partial}{\partial \theta}\left(u^{emp} - u^{unemp}\right)$, the difference between employment and unemployment utility in response to more Internet access. Third, $u_{\ell\theta}$, the change in marginal utility from leisure in response to more Internet access.

Chapter B

DMP framework

I summarize the standard equilibrium search and matching model briefly in this appendix.

The hiring process is governed by a matching function that produces worker-employer pairs using job vacancies and jobseekers as inputs,

$$H_t = A_t v_t^{\alpha} u_t^{1-\alpha} \tag{A1}$$

where u_t is the number of jobseekers, v_t is the number of vacant jobs, and A_t is the efficiency of the search and matching process.

The probability of finding a job match for the unemployed worker is given by $A_t(v_t/u_t)^{\alpha} = A_t(\theta_t)^{\alpha}$, where θ_t represents the labor market tightness.

All workers face the same constant unemployment risk λ. At steady states, the flow into unemployment $\lambda(1-u)$ should equal the flow out of unemployment $A\theta^{\alpha}u$. Unemployment can be solved in terms of two transition rates,

$$u = \frac{\lambda}{\lambda + A(\theta)^{\alpha}} \tag{A2}$$

Workers maximize the net present value of income and randomly search for vacant jobs while unemployed. The flow value of being unemployed is $rU = b + A(\theta)^{\alpha}(W - U)$, and the flow value of working is $rW = w + \lambda(U - W)$. Firms receive a flow value of profits for active jobs according to $rJ = p - w - \lambda J$, and the flow value of vacancy is $rV = -c + A(\theta)^{\alpha-1}(J - V)$. In profit-maximizing equilibrium, the expected value of a vacancy is driven to zero by free entry of new vacancies. We can derive the job creation condition as,

$$p - w - \frac{(r+\lambda)c}{A(\theta)^{\alpha}} = 0 \tag{A3}$$

The wage is assumed to be derived from a Nash bargaining solution: the w that maximizes the weighted product of the worker's and the firm's net return from the job match.

$$w = \arg\max \left(W - U\right)^{\beta} \left(J - V\right)^{1-\beta}, \tag{A4}$$

where β can interpreted as a relative measure of labor's bargaining strength, and it is between 0 and 1. First order condition gives the wage setting condition as,

$$w = (1-\beta)b + \beta p(1 + c\theta) \tag{A5}$$

Equilibrium is a unique set of (u, θ, w) that satisfies the flow equilibrium condition A2, the job creation condition A3, and the wage equation A5. By changing the parameter value of matching technology A_t and the value of unemployment income b, I numerically solve the new equilibrium after a Internet access shock. The simulated results are shown in Figure 1.6.