Do Information and Communication Technologies Empower Female Workers? Evidence from Vietnam *

Heiwai Tang[†]
Hong Kong University

Yao Wang[‡]
The Ohio State University

Victoria Wenxin Xie[§]
Santa Clara University

Natalie Chun[¶]
Cambly Inc.

July 2025

Abstract

We examine the impact of information and communication technology (ICT) adoption on female employment outcomes in Vietnam, using data at both the firm and household levels. Our identification strategy leverages geographic variation in the engineering costs of connecting to Vietnam's ICT backbone, along with the temporal shock of the 2007 major network capacity expansion. We find that increased ICT adoption raises firms' female labor share and reduces gender gaps in monthly work days, formal employment, and time spent on housework. We also document a plausible mechanism: ICT adoption shifts labor demand away from routine and manual tasks toward social, customer-facing, and service-oriented tasks—areas in which women are more likely to hold a comparative advantage.

JEL Classifications: J16, O33, L86, H54

Keywords: Information Communication Technologies (ICT), Infrastructure, Gender Inequality

*Acknowledgments: We thank the Asian Development Bank for financial support, and Alex Hollingsworth Julie Cullen, Audrey Guo, Elisabetta Gentile, and seminar participants at the AEDE Faculty Brownbag for their helpful comments.

[†]Email: tangecon@hku.hk.
[‡]Email: wang.16488@osu.edu.
[§]Email: wxie@scu.edu.
[¶]Email: natalie.l.chun@gmail.com

1 Introduction

Since the 1990s, advances in information and communication technologies (ICT) have significantly reduced the costs of acquiring information and coordinating activities within and across firms. Over the same period, income inequality within countries has risen sharply, even as the gap between nations has narrowed due to the rapid growth of several emerging economies (Baldwin, 2013; Piketty's, 2014). In particular, gender inequality in labor markets remains especially persistent, with women continuing to face barriers to entry, advancement, and equal pay across sectors and skill levels (Goldin, 2014; Blau and Kahn, 2017; International Monetary Fund, 2018; Hsieh et al., 2019; World Economic Forum, 2023). Rising inequality has been linked to political and social unrest in nondemocratic states (Campante and Chor, 2012) and to the surge of populism and polarization in Western democracies (Autor et al., 2020). Yet, despite the profound transformations ICT has brought to labor markets, its role in shaping gender disparities remains underexplored.

This paper contributes to the debate on how ICT adoption shapes female labor market outcomes, with a particular focus on broadband internet.[1] In contrast to much of the existing literature that focuses on advanced economies, we investigate this question in the context of a large developing country, Vietnam. Our analysis leverages two rich data sources: Vietnam's annual enterprise census, a panel dataset that provides detailed information on firm-level ICT adoption and female employment; and nationally representative household surveys that offer granular data on individual labor market outcomes, time use, and household-level ICT adoption. These comprehensive datasets allow us to directly observe ICT adoption at both the firm and household levels, enabling us to assess the effects of national ICT infrastructure expansion on female employment.

Identifying the causal impact of ICT expansion on labor market outcomes is challenging, as ICT infrastructure is not randomly placed, and firms and households may endogenously sort into areas with better access. To address this, our paper leverages two dimensions of exogenous

[1]Throughout this paper, we use the term "ICT adoption" to refer specifically to the adoption and use of broadband internet infrastructure.

variation that influence firms' and households' ICT access. First, we exploit temporal variation arising from the nationwide expansion of ICT infrastructure in 2007. In early 2000s, Vietnam experienced a dramatic expansion in ICT infrastructure, marked by a major upgrade in 2007 to its domestic backbone and submarine cable systems. By comparing the same firms and similar households before and after the ICT capacity expansion, we can effectively remove the effects of time-invariant observables and unobservables. Second, we exploit geographic variation in the engineering costs of accessing the ICT backbone. Conditional on ICT expansion, not all firms and households were equally affected; treatment intensity depends on the geographic difficulty of connecting to the nearest ICT backbone node—the core infrastructure that transmits high-capacity internet across regions. To capture this, we construct a proxy for ICT accessibility for each firm or household based on terrain, soil quality, and hydrological features along the path to their closest backbone node. Specifically, the presence of permanent water bodies (such as lakes and rivers), unstable soil, and rugged terrain increases the complexity and cost of infrastructure deployment, reducing ICT accessibility.[2]

We then interact this spatial variation in engineering suitability with the temporal expansion of ICT capacity after 2007 to construct an instrument for changes in firm-level or household-level ICT adoption. Our approach, similar to a generalized difference-in-differences strategy (Nunn and Qian, 2011), compares changes in ICT adoption (and female labor outcomes) across firms (or households) located in areas more or less geographically suited for infrastructure deployment, before and after the expansion. Since we directly observe ICT adoption at both the firm and household levels, we use the spatial and temporal variation as instruments for ICT adoption, rather than as treatment indicators as in a generalized difference-in-differences (DiD) framework.

Our identification strategy relies on two key assumptions: First, relevance: firms and households facing fewer geographic barriers benefited more from the ICT expansion (Banerji and Chowdhury, 2013); Second, the exclusion restriction: the instrument affects labor market outcomes only through its effect on ICT adoption. That is, there were no concurrent shocks around

[2]For example, these features can significantly extend the actual fiber-optic cable length required, as the cable must be routed around natural obstacles rather than along a direct, straight-line path.

2007 that are correlated with geographic suitability for connecting to backbone nodes. We address this challenge by directly controlling firm fixed effects, year fixed effects, and time- and firm- varying factors that might bias our estimates. We also account for differential time trends by backbone node, by distance to backbone nodes, as well as local population growth and urbanization trends.

We find that increasing firm-level adoption of ICT significantly increases the share of female labor. We begin by validating our instrumental variable. Descriptive statistics reveal a sharp rise in ICT use following Vietnam's major national ICT capacity expansion in 2007. Our first-stage results confirm that a firm's exposure to this expansion critically depends on its access to ICT infrastructure, which is shaped by local terrain, hydrological and soil conditions relative to the nearest ICT backbone node. In the second stage, we identify a robust and statistically significant positive effect of ICT adoption on female employment shares. A 10% increase in the number of Internet-connected personal computers (PCs) per worker raises the female labor share by 3.6 percentage points—an 11% increase over the sample's baseline average of 34%. Likewise, a 10% increase in LAN-connected PCs per worker raises the female labor share by 4.7 percentage points. We also provide supporting evidence from a reduced-form specification that shows no significant pretrends. Additionally, These findings are robust to a range of sensitivity checks, including alternative instrument constructions and different standard error clustering methods.

We complement the firm-level analysis with data from the Vietnam Household Living Standards Survey (VHLSS) to provide further insights into the informal economy and a richer set of labor market outcomes. The VHLSS records whether a household has internet access, which we use as a proxy for household-level ICT adoption. We adopt a similar identification strategy as in the firm-level analysis, but account for the repeated cross-sectional nature of the VHLSS by controlling for commune fixed effects—the most granular geographic unit available—and a rich set of household and individual characteristics. Our results show that household ICT adoption has significantly different effects on women compared to men across a range of labor market and time-use outcomes. Relative to men, ICT access increases women's likelihood of

participating in the formal sector, monthly work days, while reducing the time they spend on housework.

Lastly, to understand why ICT adoption disproportionately benefits women, we hypothesize that technological change shifts production toward tasks in which women hold a comparative advantage (Autor et al., 2003; Autor and Dorn, 2013; Deming, 2017; Acemoglu and Restrepo, 2020; Hötte et al., 2023). ICT substitutes for routine and manual labor while complementing non-routine, social, and service-oriented tasks—areas where women are more likely to specialize, as documented gender differences in social and cognitive skills (Rueckert and Naybar, 2008; Schulte-Rüther et al., 2008; Cortes and Pan, 2018). Using data from Vietnam, we examine occupational sorting patterns and find suggestive evidence consistent with this mechanism: women are overrepresented in sectors more intensive in these complementary tasks.

We provide empirical support for this mechanism using both firm-level and household-level data. At the firm level, we find that the positive effect of ICT adoption on female labor share is significantly stronger in industries with lower routine and manual task intensity and higher service, social, and customer-facing task intensity. At the individual level, we show that women in internet-connected households are more likely than men to shift into occupations characterized by social and service-oriented tasks, and less likely to remain in routine or manual roles.

This paper contributes to the growing literature on the impact of ICT on labor market outcomes (Autor et al., 1998; Crandall et al., 2007; Forman et al., 2012; Michaels et al., 2014; Akerman et al., 2015), which has predominantly focused on developed countries. In the context of developing economies, Hjort and Poulsen (2019) show that internet access has large positive effects on employment rates in Africa. Likewise, Hjort and Tian (2021) provide a comprehensive review highlighting that internet connectivity can yield substantial labor market and productivity gains, although the effects vary by context and the form of access. Recent work by Azam et al. (2024) finds that mobile internet expansion improves employment and earnings outcomes in India, particularly among more ICT-skilled female workers. Our study extends this literature by focusing on Vietnam, one of the fastest growing developing economies, and by examining

gender-specific labor market outcomes, an underexplored dimension in this literature. A key methodological innovation of this paper is the use of exogenous geographic features–such as terrain, soil quality, and water bodies along the path to ICT backbone nodes—as proxies for the engineering costs of fiber-optic deployment, which serve as instruments for firms' and households' ICT access. Our identification strategy is conceptually similar to that of Ellis et al. (2025), who exploit variation in terrain and the spatial distribution of mobile phone towers to predict for mobile signal strength.[3]

Second, our study contributes to the growing literature on technological change, task composition, and gender inequality (Autor et al., 2003; Aksoy et al., 2021; Beaudry and Lewis, 2014; Black and Spitz-Oener, 2010; Cortés et al., 2024; Greenwood et al., 2016; Kwon, 2022; Ngai and Petrongolo, 2017; Olivetti and Petrongolo, 2016; Vidart, 2024).[4] Two studies are particularly closely related to ours. Juhn et al. (2014) link trade-induced technological change to improved female outcomes in Mexico, especially in blue-collar jobs, as computerization reduced physical task demands. While they focus on machinery upgrades, we study broadband ICT expansion in Vietnam and use direct firm-level measures of ICT adoption, allowing us to examine impacts in both formal and informal sectors. Chiplunkar and Goldberg (2022) find that 3G mobile internet increased female labor force participation in developing countries, particularly wage jobs in service sector jobs—a result consistent with ours. Similarly, Chung et al. (2025) find that mobile phone access boosted women's employment and shifted gender norms in Afghanistan. However, unlike their focus on mobile internet or broadly defined trade-induced computerization, we study broadband infrastructure, using a distinct identification strategy.

Third, our paper contributes to the literature on labor market returns to social skills. Prior research shows that non-cognitive skills are increasingly valued (Heckman et al., 2006; Borghans et al., 2008; Lindqvist and Vestman, 2011). Deming (2017) models rising returns to social skills as jobs become less routine—an effect reinforced by ICT-driven shifts toward flexible work design (Autor et al., 2003; Bresnahan et al., 2002; Bartel et al., 2007). We build on this strand of work and

[3]Other studies, including Miner (2015) and Geraci et al. (2022), also exploit variation in backbone network connectivity, but rely on distance to the backbone as the instrument rather than using engineering-based cost proxies as we do.

[4]See Goldin (2006) for a broader discussion of the economic, sociological, and political drivers of female labor market trends in 20th-century U.S.

examine gender-specific effects of ICT expansion in Vietnam, where women are shown to have comparative advantages in cognitive and social-emotional tasks (Welch, 2000; Schulte-Rüther et al., 2008; Rueckert and Naybar, 2008; Rendall, 2017). The patterns we discover in our paper suggest that greater flexibility in task design, as highlighted by Deming (2017), may be a key mechanism linking ICT expansion to female employment gains in developing countries.

The paper proceeds as follows. Section 2 outlines the institutional background. Section 3 describes the data. Section 4 presents our identification strategy, the construction of the instrumental variable, and firm-level results. Section 5 extends the analysis using household survey data. Section 6 investigates potential mechanisms. Section 7 concludes.

2 Background

Internet Expansion in Vietnam As a relatively late adopter of the Internet, the Vietnamese government recognized that expanding low-cost access to ICT would be critical to achieving its broader economic reform goals under the *Doi Moi* (open door) policy initiated in the late 1980s. The country's first steps toward connectivity began in 1992, with the development of a national computing network through collaboration between Hanoi's Institute of Information Technology (IOIT) and the Australian National University.[5] Following provisional regulations, Internet services became available to commercial users in the late 1990s. The core infrastructure supporting Internet provision was the national optical fiber backbone, constructed by Vietnam Telecom National (VTN), a subsidiary of the Vietnam Post and Telecommunications Group (VNPT). This backbone served as the primary high-capacity transmission system, enabling long-distance data connectivity across provinces and forming the foundation for the expansion of broadband access nationwide. Panel B of Figure 2 displays the early terrestrial backbone and node sites, based on Transmission Network Data from the International Telecommunication Union (ITU). The national network, built using a ring topology, connected Hanoi to Ho Chi Minh City through fiber optic systems, with extensions to the North, Northeast, Central, and

[5]The foundation of Vietnam's Internet infrastructure was laid with the creation of the Vietnam Academic Research and Education Network (VARENet) in 1994, which supported academic communication and research exchanges.

Southern regions.

Despite these efforts, Internet penetration and bandwidth remained limited in the early 2000s due to high costs and capacity constraints in the optical transmission infrastructure. As of 2005, Vietnam had only 210,000 fixed broadband subscriptions, and its total international bandwidth hovered around 3 Gbps (Tuan, 2011).

A turning point occurred in 2007. This year marked major expansions in Vietnam's ICT capacity, driven by two transformative technological upgrades. First, the country's main submarine cable, SEA-ME-WE 3—which directly connected Vietnam to over 30 countries across East Asia, Africa, Europe, and Oceania—underwent a critical 10G capacity upgrade. This was the first major enhancement since the cable's initial deployment in 1999 and significantly improved international connectivity in terms of both speed and reliability. Second, Vietnam's domestic terrestrial backbone was substantially upgraded in the same year. According to the Vietnam Post and Telecommunications Group, this expansion enabled much broader residential and commercial access to ADSL broadband. Complementing these infrastructure improvements, Vietnam's accession to the World Trade Organization (WTO) in 2007 lowered the cost of importing ICT equipment and encouraged foreign investment through joint ventures, fostering competition and facilitating the adoption of global standards.

Taken together, these developments marked a watershed moment for Vietnam's digital transformation. They accelerated Internet speed, broadened access, and dramatically reduced costs. Within a decade, Vietnam had transformed from one of Southeast Asia's technological laggards into a regional leader in digital connectivity. Reflecting this transformation, both domestic and international bandwidth grew sharply after 2007. As illustrated in Figure 1, international bandwidth surged from 3,615 Mbps in 2005 to over 400 Gbps by 2013. By 2011, Vietnam had become a highly competitive ICT market, with 90 Internet providers and five cellular data companies (Tuan, 2011). By 2015, as shown in Appendix Table A.1, average bandwidth had exceeded 24,000 bits per second per person, and entry-level broadband plans were available for as little as $3 per month, making Internet access broadly affordable.

Gender Differences in the Labor Market Vietnam's labor market, while exhibiting relatively small gender gaps in labor force participation, displays persistent disparities in employment type and earnings. In 2013, 85% of males and 77% of females aged 15–64 participated in the labor force (Appendix Table A.2). Although gender wage gaps narrowed between 2007 and 2013—particularly among less-educated workers—men continued to earn approximately 15–30% more than women on average (Appendix Table A.3). As in many developing countries, informal employment remains widespread. Women are also less likely to hold formal wage jobs: only 30% of female workers are formally employed, compared to 40% of males.

Against this backdrop, ICT adoption may help reduce gender disparities in labor market outcomes. If firm-level ICT use enhances the productivity of tasks in which women have a comparative advantage, it could translate into higher wages and more stable employment for female workers.

3 Data

Firm-level Data The firm-level data used in our analysis come from the Annual Enterprise Census of Vietnam, conducted by the General Statistics Office (GSO) and covering the period from 2004 to 2009. These annual censuses include all registered enterprises across both manufacturing and non-manufacturing sectors, encompassing state-owned, collectively owned, private, and foreign-invested firms. The dataset provides rich firm-level information, including detailed balance sheets, total employment, disaggregated employment by skill and gender, capital, revenues, expenses, profits, and liabilities.

A distinguishing feature of the Vietnamese enterprise census is its comprehensive coverage of firm-level ICT usage. In particular, it reports whether firms use personal computers, have internet access, and operate a Local Area Network (LAN).[6] Panel A of Table 1 summarizes the key firm-level variables related to ICT adoption since 2004 to 2007. Data for 2006 are unavailable, as the survey did not collect ICT adoption information that year. The two primary variables

[6] A Local Area Network (LAN) is a network of interconnected computers and devices within a limited geographic area—such as an office or building—that allows users to share data, applications, and resources like printers or internet connections.

are the number of Internet-connected personal computers (PCs) per worker and the number of LAN-connected PCs per worker. Both indicators show a steady upward trend over time. The average number of Internet-connected PCs per worker increased from 0.05 in 2004 to 0.16 in 2009, while the average number of LAN-connected PCs per worker rose from 0.06 to 0.13 over the same period.[7] Notably, according to GSO, in 2004, internet penetration among firms was approximately 49%, rising to 85% by 2009.

We also examine the raw correlation between ICT adoption and firms' female labor share. Figure 3 presents binscatter plots illustrating these relationships. The left panel shows a positive association (correlation = 0.08) between the female labor share and the number of internet-connected PCs per worker. Similarly, the right panel displays a positive correlation (correlation = 0.07) between the number of LAN-connected PCs per worker and the female labor share.

Although the Vietnam Enterprise Surveys do not provide the exact coordinates of firms, we can identify their province (*tinh*), district (*huyen*), and commune (*xa*) using concorded geographic identifiers. We use the centroids of each commune as a proxy for the geographic location of the firms. There are 11,162 communes (*xa*) in Vietnam, which represent the lowest administrative level in the country, nested within provinces (*tinh*) and districts (*huyen*). The average population of a commune is approximately 6,500, making it roughly comparable to a U.S. Census tract in terms of population, but much smaller in geographic size due to Vietnam's higher population density.[8] Panel A of Figure 2 displays the communes of the firms in *blue*, as reported in the Vietnamese Enterprise Surveys.

Household-level Data To complement our firm-level analysis and shed light on broader labor market outcomes—particularly within the informal economy—we also use individual-level data from the Vietnam Household Living Standards Survey (VHLSS), which spans the years 2002 to 2008, the most recent wave that includes household ICT adoption information. The VHLSS is a nationally representative, multi-topic household survey administered by the

[7] The GSO did not report ICT-related data in the 2006 census round.

[8] To extract the commune coordinates, we developed a Python program that matches the three levels of administrative names with Google Maps' Geocoding Platform. For a small number of mismatched communes, we manually corrected the coordinates.

General Statistics Office of Vietnam to monitor living conditions and socio-economic indicators. These surveys, conducted every two years, include workers across all industries and report detailed demographic characteristics such as gender, age, education, marital status, and commune-level geographic location. This information enables us to geocode each household and measure its proximity to the nearest ICT backbone node.

The labor force module provides detailed information on employment outcomes, including labor force participation, number of workdays per month, employment sector, formality of employment (formal wage work vs. self-employment), whether the individual holds a second job, and engagement in housework—an especially relevant measure in the context of gender differences. Due to differences in employment classification in the 2002 wave, we exclude that year when analyzing formal employment as the outcome variable to ensure consistency. Although the survey includes a question on whether individuals worked for a salary in the past year, the definition of this variable is inconsistent between 2008 and other survey years. Additionally, the survey reports only total household income and earnings, without disaggregating by gender. As a result, we omit these variables from our analysis.[9]

Importantly, the VHLSS records whether a household has a computer with internet access, which we use as a proxy for household-level ICT adoption. Although the survey also reports the frequency of internet use among connected households, we exclude this variable from our analysis due to a high rate of missing values. Panel B of Table 1 documents the rise in household internet connectivity from 2002 to 2008, with the share increasing from just 0.3 percent to 4 percent. [10]

Terrestrial Backbone Nodes Location We obtain Vietnam's historical terrestrial backbone network maps from the International Telecommunication Union (ITU) and the Vietnam Telecom

[9]We follow the procedures in McCaig and Pavcnik (2018) and McCaig and Pavcnik (2021) to clean individual and household identifiers. Similar data have been used in McCaig and Pavcnik (2018) to examine the impact of trade liberalization on labor reallocation.

[10]This figure is notably lower than estimates reported by the International Telecommunication Union (ITU) World Telecommunication/ICT Indicators Database, as our measure is limited to internet access via a home computer, which remained relatively uncommon for Vietnamese households in 2008. In contrast, the ITU metric includes access through a broader range of devices, including mobile phones, personal digital assistants, gaming consoles, and digital televisions (International Telecommunication Union (ITU), 2025).

National (VTN). Panel B of Figure 2 displays the original map, which consists of fiber optic transmission systems connecting Hanoi and Ho Chi Minh City, along with regional networks serving the North, Northeast, Central, and Southern regions. Existing inter-provincial backbone lines are shown in dark red, with main backbone nodes—telecommunication stations—marked by black dots outlined in red.

Since backbone nodes—telecommunication stations—aggregate and distribute ICT infrastructure, they serve as the primary physical access points for high-speed connectivity to firms, households, and public services in surrounding areas. We digitize their geographic locations, which are later used to construct instrumental variables for firm- and household-level ICT adoption. Panel A of Figure 2 illustrates the geographic distribution of Vietnam's terrestrial backbone nodes (red dots), overlaid with commune-level administrative boundaries (blue points). While there is a notable concentration of nodes around major urban and industrial hubs—such as Hanoi in the north, Da Nang in the central region, and Ho Chi Minh City in the south—the backbone network is relatively evenly distributed along the country's north–south axis. The majority of communes are located in close proximity to at least one backbone node.

Table 2 presents additional summary statistics on the relative location of firms to the backbone nodes. The Annual Enterprise Census from 2004–2009 includes data for 96,492 unique firm IDs, distributed across 64 provinces, 665 districts, and 6,989 communes. Approximately 87% of firms are located within 20 kilometers of the nearest terrestrial backbone node site. In our main specification, we exclude firms located in communes more than 20 kilometers away from the backbone to ensure we are comparing firms in similar locations. We also conduct robustness checks by testing sensitivity to alternative exclusion thresholds.

Geographic Data To construct our instrument for firms' ICT access, we exploit terrain, hydrological, and soil characteristics along the shortest path between each firm and its nearest ICT backbone node. We draw on the following satellite data sources to capture localized geographic features: (1) elevation from NASA's SRTM Digital Elevation Model (Farr et al., 2007), (2) terrain ruggedness from Nunn and Puga (2012), (3) soil quality indicators from SoilGrids (Hengl et al., 2017), and (4) hydrological features from the JRC Global Surface Water Mapping Layers, V1.3

(Pekel et al., 2016).

Appendix Table A.4 summarizes the variables that may affect the engineering feasibility of ICT infrastructure deployment. For instance, higher elevation or steep slopes increase construction costs, while rugged terrain complicates cable installation and tower placement. Soil characteristics such as texture and bulk density affect the stability of underground installations, and nearby water bodies may pose physical barriers to infrastructure development.

4 Firm-level ICT Adoption and Female Employment Share

4.1 Empirical Strategy

Instrumental Variable Construction In this section, we examine how firm-level ICT adoption affects the share of female employment within firms. Identifying causal effects is challenging, as the placement of ICT infrastructure is not random. Instead, the location of backbone nodes is often shaped by policymakers' strategic decisions and may correlate with local socio-economic conditions that also influence female employment outcomes, raising concerns about omitted variable bias. Additionally, reverse causality is possible: firms with a higher share of female workers may undergo structural changes that increase their demand for ICT.

To address these concerns, we exploit a major policy-driven expansion in Vietnam's ICT infrastructure that substantially improved Internet access. In 2007, Vietnam's core domestic terrestrial backbone and submarine cable systems underwent significant capacity upgrades, leading to a sharp increase in Internet transmission speeds (see details in Section 2). However, this period also overlapped with several other transformative economic developments, including Vietnam's accession to the World Trade Organization (WTO) and subsequent waves of globalization and urbanization. To isolate the effects of ICT improvements from these concurrent trends, we construct a firm-level treatment intensity measure based on geographic variation in the cost of connecting to ICT infrastructure—specifically, the engineering difficulty of linking to the nearest backbone node.

Firms connect to ICT backbone nodes via fiber optic networks, and the construction costs of

these connections depend on the terrain, soil and hydrological conditions between the firm and the nearest node. A growing body of engineering research highlights that terrain and soil characteristics play a critical role in shaping installation costs of infrastructure (Biggs and Mahony, 2004). For example, the Fiber Broadband Association and Cartesian's 2023 report estimates that underground fiber deployment costs approximately $10 per foot in soft soil but doubles to $20 per foot in rocky terrain due to increased excavation difficulty and labor requirements (Fiber Broadband Association and Cartesian, 2023). The report also highlights that additional local challenges—such as crossing waterways or dealing with unstable underground conditions—can further raise deployment costs. Geospatial analyses similarly show that geographic variation, particularly in terrain, significantly influences infrastructure installation expenses. As a result, firms located in areas with flat terrain, stable soils, or generally favorable construction conditions faced lower costs and faster connections to backbone nodes. When Vietnam's domestic backbone network underwent major capacity upgrades in 2007, these firms experienced more immediate and pronounced improvements in ICT access. Accordingly, they are considered to have received higher treatment intensity in our empirical framework.

Motivated by insights from engineering research, we use the terrain, soil, and hydrological characteristics between each firm and its nearest backbone node as proxies for the cost of establishing ICT connections. Specifically, for each firm, we identify the closest backbone node based on Euclidean distance. We then draw a straight line between the firm and this node and create a 3 km buffer around this path. We term this buffer *connectivity buffer* for future reference. Within this buffer, we compute the mean, standard deviation, minimum, and maximum of all included pixels for key geographic characteristics—such as terrain ruggedness, elevation, soil properties, and hydrological features (see Table A.4 for details). Finally, We interact the geographic characteristics of each firm's connectivity buffer with a post-2007 indicator to construct our instrumental variable for firm-level ICT adoption.

Main Specification To examine the effect of firm-level ICT adoption on the share of female employment, we first estimate the following specification:

$$Y_{it} = \alpha_0 + \beta_1 ICT_{it} + \gamma_i + \theta_t + \beta_2 X_{it} + \beta_3 \log(DistNode_i) \times Post07_t + \beta_4 \delta_{n(i)} \times Post07_t + \epsilon_{it} \quad (1)$$

where Y_{it} is the dependent variable, representing the share of female workers in total employment for firm i in year t. The main independent variable of interest, ICT_{it}, captures firm-level ICT adoption. We use two alternative measures for robustness: (1) the log number of PCs connected to the Internet per worker and (2) the log number of PCs connected to local area networks (LAN) per worker. The specification includes firm fixed effects, γ_i, to control for time-invariant firm-level characteristics, and year fixed effects, θ_t, to control for national time trends common to all firms. X_{it} is a vector of time-varying firm-level controls, including firm age, the square of firm age, a foreign ownership dummy, a state ownership dummy, total sales, fixed assets, and total employment. Robust standard errors are clustered at the firm level to account for serial correlation in the error term ϵ_{it}.[11]

To address endogeneity concerns, we instrument ICT adoption using the interaction between the geographic characteristics of the connectivity buffer and the timing of the 2007 policy expansion, as described above. In addition, we include two key controls. First, we interact the firm's closest node ID, denoted as $\delta_{n(i)}$, with a post-2007 indicator, $Post07_t$, to allow for node-specific differential trends after the ICT infrastructure upgrade. Second, we interact the log distance between firm i and its nearest backbone node, $\log(DistNode_i)$, with $Post07_t$, to control for differential post-2007 trends that vary by distance to backbone nodes. To form better comparison, we exclude locations that are more than 20 km away from any backbone node site.

Figure 4 illustrates the variation exploited in our identification strategy. In the figure, the blue triangle represents a backbone node, while the green and red circles denote two firms that both identify this node as their nearest node. Although the firms are equidistant from the same node, they differ in the geographic characteristics along the path to the node. Specifically, we compare the change in ICT adoption for Firm A (green circle) before and after 2007 with the

[11] The main results are robust to alternative clustering strategies, at the commune level, commune-by-year level, and two-way clustering at both the commune-by-year and firm levels, which we discuss in the robustness checks in Section 4.

corresponding change for Firm B (red circle) over the same period. The difference in these changes is driven by the distinct geographic conditions within their respective connectivity buffers (green for Firm A and red for Firm B). We then examine how these firm-level changes in ICT adoption affect their female employment shares.

The exclusion restriction of our IV strategy is that, conditional on being at a similar distance from the same nearest backbone node, the terrain, soil and hydrologic characteristics along the path between a firm and that node influence the firm's female labor share solely through their effect on ICT adoption. A potential concern with this exclusion restriction is that, despite relying on highly localized variation (Figure 4), geographic characteristics between firms and ICT nodes might also affect changes in a firm's female labor share through channels unrelated to internet access—such as differences in urbanization or population growth along the connectivity path. To further address this concern, we control for the average population density within the *connectivity buffer*, $\log(pop)_{it}$, and include each firm's distance to the nearest urban center ($DistCBD_i$), interacted with year fixed effects, $Year_t$, allowing for flexible control of trends by proximity to urban areas. Consequently, we estimate the following IV regression as our baseline specification, using the geographic characteristics of the *connectivity buffer* as the instruments for firm-level ICT adoption, ICT_{it}:

$$Y_{it} = \alpha_0 + \beta_1 ICT_{it} + \gamma_i + \theta_t + \beta_2 X_{it} + \beta_3 \log(DistNode_i) \times Post07_t + \beta_4 \delta_{n(i)} \times Post07_t + \beta_5 \log(DistCBD_i) \times Year_t + \beta_6 \log(pop)_{it} + \epsilon_{it} \quad (2)$$

4.2 Results

First Stage Results Given the large number of potential geographic characteristics that may influence firms' access to ICT infrastructure, we employ a post-double-selection LASSO procedure to identify the optimal set of instrumental variables, following Belloni et al. (2012). Specifically, we select the terrain, soil, and hydrological features of the connectivity buffer that best predict firms' ICT adoption. After applying the LASSO procedure, we estimate the first-stage

regression using the selected variables. Table 3 presents the results from this estimation:

$$\begin{aligned} ICT_{it} = \alpha + \beta_1 selectIV_i \times Post07_t + \gamma_i + \theta_t + \beta_2 X_{it} \\ + \beta_3 \log(DistNode_i) \times Post07_t + \beta_4 \delta_{n(i)} \times Post07_t \\ + \beta_5 \log(DistCBD_i) \times Year_t + \beta_6 \log(pop)_{it} + \epsilon_{it} \end{aligned} \quad (3)$$

where $selectIV_i$ denotes the set of instruments selected using the LASSO procedure, and all other variables are defined as in equation (2). For ease of presentation, we report only the names of the selected instruments, but the estimated coefficients, β_1, correspond to their interaction terms with the post-policy indicator, $selectIV_i \times Post2007_t$.

Table 3 presents the estimation results. Out of 32 potential geographic characteristics, only 5–6 are selected by the LASSO procedure. In Column 1, where the dependent variable is the log of PCs with internet access per worker, we find that higher sand content within the buffer zone facilitates ICT adoption—consistent with engineering evidence linking sandy soils to lower fiber installation costs. The results also reveal a significant positive effect of coarse fragment content. While coarse fragments (e.g., gravel or pebbles) can complicate trenching and increase deployment costs, moderate levels—such as those found in gravelly loams—enhance drainage and reduce the risk of waterlogging and cable corrosion, which is particularly important in Vietnam's flood-prone deltas. The overall positive effect suggests that the benefits of improved drainage and soil stability outweigh the potential costs of trenching. Finally, we find that higher soil bulk density negatively affects ICT adoption, indicating that denser soils hinder infrastructure deployment.

The LASSO procedure's selection of the "minimum" rather than the "average" of soil characteristics within the buffer indicates that the least favorable soil condition plays a more critical role in determining engineering difficulties, highlighting a "bottleneck" effect where the most challenging area limits overall feasibility.

Hydrological factors also play an important role in ICT deployment. ICT adoption is nega-

tively associated with long-term water occurrence, which is defined as how persistently water has appeared at a location since 1984. This suggests that locations with consistently high water presence, such as wetlands or lake basins, face greater physical access constraints and are therefore less suitable for ICT infrastructure deployment. Finally, we find that ICT adoption is positively associated with spatial variability in water occurrence. This indicates that areas with more spatially heterogeneous hydrological conditions are more likely to experience broadband expansion, possibly reflecting flexible installation strategies and adaptive infrastructure planning in these regions.

Interestingly, elevation and terrain ruggedness were not selected by the LASSO procedure, likely because the inclusion of backbone node fixed effects limits identifying variation within localized areas, where terrain differences are minimal. Column 2 reports results using the log of personal computers connected to a LAN per capita as the dependent variable. While the LASSO procedure selects fewer instruments in this specification, the estimated patterns remain largely consistent with those in column 1.

Results on Female Employment Shares Panel A of Table 4 presents the estimated effect of ICT adoption on the log number of Internet-connected PCs per worker. Each entry reports the coefficient on β_1 from a specific specification of regressions (1) or (2). Column 1 presents the OLS estimates from equation (1). This specification controls for a range of firm-level characteristics—including log sales, log assets, log employment, firm age, foreign and state ownership—as well as year and firm fixed effects. It also accounts for node-specific post-2007 trends and post-2007 trends that vary with distance to backbone nodes. We find a positive and statistically significant correlation between ICT adoption and the female labor share: a 10 percent increase in the number of Internet-connected PCs per worker is associated with a 0.62 percentage point increase in a firm's female labor share. However, the OLS estimates may be biased due to firms' nonrandom access to ICT infrastructure and the presence of time-varying, firm-level omitted variables. For instance, changes in productivity or managerial capability could influence both a firm's propensity to adopt ICT and its demand for female workers.

To address potential endogeneity concerns, in Column 2, we begin by estimating specifi-

cation (1), using the geographic conditions of the *connectivity buffer* between each firm and its nearest ICT backbone node as instruments, as outlined in Section 4. In column 3, we introduce an additional control for population growth within the *connectivity buffer* using GHS-POP, Global Human Settlement Layer data (Schiavina et al., 2023).[12] Since population data is only available for 2005 and 2010, we calculate the annual population growth rate for each *connectivity buffer* and infer the annual population levels for the intervening years. To further account for urbanization effects, in column 4, we include further include the interactions between year fixed effects and the firm's distance to the nearest urban center, thereby controlling for potential confounding effects from differential urbanization patterns or population growth.

Column 4 is our baseline specification. Across all IV regressions (Columns 2-4), the Kleibergen-Paap Wald F statistics exceed 10, indicating strong instrument relevance (Kleibergen and Paap, 2006). The Kleibergen-Paap LM test strongly rejects the null of underidentification, confirming that the instruments are relevant for the endogenous density variable. In our baseline model (Column 4), the Anderson–Rubin Wald test yields a p-value of 0.063, suggesting that the IV estimates are statistically robust.

Our results are consistent across specifications from column 2 to column 4. The IV estimates suggest a substantially larger positive effect of ICT adoption on the female labor share compared to the OLS estimate in column 1. Specifically, a 10% increase in the number of internet-connected PCs per worker leads to a 3.3-3.6 percentage point increase in the share of female workers at the firm. This suggests that OLS likely underestimates the true causal effect, possibly because ICT infrastructure was more likely to be deployed in areas with characteristics negatively correlated with female labor share, such as male-dominated labor markets.

Panel B of Table 4 reports OLS and IV results using an alternative measure of ICT adoption: the log number of LAN-connected PCs per worker. The estimates consistently indicate a positive—and slightly larger—effect of ICT adoption on the female labor share. Specifically, a 10% increase in LAN-connected PCs per worker is associated with a 4.3–4.7 percentage point increase in the female share of employment.

[12]GHS-POP is a globally harmonized population dataset produced by the European Commission's Joint Research Centre. It provides gridded population estimates at high spatial resolution by combining census data with satellite-derived built-up areas.

Parallel Trends Our instrumental variable strategy builds on the intuition of a generalized difference-in-differences (DiD) approach (Nunn and Qian, 2011). Specifically, the instrument captures differential changes in female labor share in the pre-expansion period, relative to the post-expansion periods between locations that were more or less easily connected to the ICT backbone network. While we observe firm-level ICT adoption directly and therefore do not implement the DiD design, we use this variation as an instrument for ICT adoption. The key identifying assumption—analogous to the parallel trends assumption in DiD—is that, prior to the 2007 policy change, firms in locations with differing geographic suitability for ICT connectivity should not exhibit systematically different trends in the outcome of interest: the female labor share.

To assess the validity of this assumption,we estimate the following reduced-form specification, which takes the form of an event study regression within a generalized DiD framework:

$$\begin{aligned} Y_{it} = \alpha + \sum_{t=2004}^{2009} \beta_{1t} selectIV_i \times Year_t + \gamma_i + \theta_t + \beta_2 X_{it} \\ + \beta_3 \log(DistNode_i) \times Post07_t + \beta_4 \delta_{n(i)} \times Post07_t \\ + \beta_5 \log(DistCBD_i) \times Year_t + \beta_6 \log(pop)_{it} + \epsilon_{it} \end{aligned} \quad (4)$$

All variables are defined as in the baseline specification in equation (2). The only difference in equation (4) is that we replace the endogenous ICT adoption variable ICT_{it} with our LASSO-selected geographic instruments, interacted with year fixed effects, $selectIV_i \times Year_t$. The resulting coefficients β_{1t} capture the year-by-year correlation between the geographic predictors of ICT connectivity and female labor share. Given the presence of multiple LASSO-selected instruments, we also conduct joint significance tests of the full set of β_{1t} coefficients of all instruments for each year.

Panel A of Table 5 reports estimates from equation (4). Column 1 presents results using the LASSO-selected geographic instruments with the log number of Internet-connected PCs per worker as the ICT adoption measure, while Column 2 uses the selected instruments with the

log number of LAN-connected PCs per worker. Coefficients for 2006 are omitted due to data unavailability, as discussed in Section 3, and coefficients for 2007 are excluded because 2007 is used as the reference year.

Most interaction terms between the LASSO-selected geographic instruments and year fixed effects are statistically insignificant in the pre-2007 period, suggesting no strong evidence of pre-existing differential trends in female labor share based on the geographic characteristics of the *connectivity buffer*. Panel B reinforces this finding: joint significance tests for the pre-expansion years yield p-values of 0.100 and 0.169 in Column 1, and 0.417 and 0.102 in Column 2, for 2004 and 2005 respectively. These results indicate that the selected instruments do not systematically predict female labor share trends prior to the 2007 expansion. This supports the validity of the parallel trends assumption in a generalized DiD framework and lends credibility to our instrumental variable strategy.

Robustness Check We conduct several robustness checks to assess the sensitivity of our main results. First, we modify the construction of the instrument by excluding the 250-meter area surrounding each firm from the connectivity buffer. This adjustment helps ensure that the instrument is not mechanically driven by firm-level geography. As shown in Columns 3 and 4 of Appendix Table A.5, the resulting estimates remain significantly positive and consistent with the baseline results in column 4 of Table 4.

Second, our baseline instrument uses a 3 km-wide *connectivity buffer* along the path between each firm and its nearest backbone node. This choice reflects two considerations: (1) the actual connection path may deviate from the straight-line distance due to terrain features, and (2) the spatial resolution of our soil and hydrological data (250 m) constrains granularity. To test robustness, we employ a narrower 0.5 km buffer in column 5 and 6 of Appendix Table A.5. Our results remain consistent with the baseline results in column 4 of Table 4.

For our main results in Table 4, we restrict the sample to firms located within 20 km of the nearest domestic backbone node site. This restriction ensures that we are comparing firms with similar pre-2007 trends in female labor share. In Appendix Table A.6, we assess the robustness of our findings to alternative distance thresholds. Columns 1–2 report the baseline results using

the 20 km cutoff, Columns 3–4 present results using a 25 km cutoff, and Columns 5–6 use a 15 km cutoff.

Finally, Appendix Table A.7 shows that our findings are robust to alternative clustering strategies, including clustering at the commune level, commune-by-year level, and two-way clustering at both the commune-by-year and firm levels.

5 Individual-level ICT Adoption and Female Labor Market Outcomes

Empirical Strategy Although the Enterprise Census provides direct measures of firm-level ICT adoption, it lacks disaggregated information on key labor market outcomes by gender and, importantly, does not include activity in the informal sector. This is a significant limitation, particularly in Vietnam—where, as in many developing countries—the informal economy accounts for the majority of total employment (McCaig and Pavcnik, 2018), as noted in Section 3. Moreover, women are less likely to participate in formal wage employment, with rates around 30%, compared to 40% for men.

To address these gaps, we complement our firm-level analysis with household-level data to examine a broader set of outcomes, including employment status, and the likelihood of working in the informal sector. Specifically, we use data from the Vietnam Household Living Standards Survey (VHLSS) for the years 2002, 2004, 2006, and 2008, which asks households whether they have internet access—a measure we use to proxy household-level ICT adoption. As with the firm-level data, we geo-code each household by commune and measure their proximity to ICT backbone nodes.

We adopt the following specification—closely paralleling the firm-level specification in equation (2)—to examine the effects of ICT adoption on female labor market outcomes at the household level.

$$Y_{it} = \alpha_0 + \beta_0 ICT_{it} + \beta_1 ICT_{it} \times Female_i + \beta_2 Female_i + \gamma_{c(i)} + \theta_t + \beta_3 X_{it} + \beta_4 \log(DistNode_i) \times Post07_t + \beta_5 \delta_{n(i)} \times Post07_t + \beta_6 \log(DistCBD_i) \times Year_t + \beta_7 \log(pop)_{it} + \epsilon_{it} \quad (5)$$

The dependent variable Y_{it} represents individual i's labor market outcome in year t, including an indicator for whether they are employed in the formal sector, monthly work days, whether they hold a second job, whether they engage in housework, and total hours spent on housework. The main independent variable, ICT_{it}, is a binary indicator equal to one if individual i's household had internet access in year t, serving as a proxy for household-level ICT adoption. To examine gender-differentiated effects of ICT expansion, we interact ICT_{it} with a female indicator, $Female_i$, which equals one if the individual is female. The coefficient of this interaction term, β_1, is our parameter of interest, which captures the differential effect of ICT on women relative to men. We also include the main effect of $Female_i$ to capture gender differences unrelated to ICT. As in our earlier analysis, we restrict the sample to households located within 20 km from the nearest backbone node.

Unlike the panel structure of the firm-level data, the household data are repeated cross-sectional.[13] As such, we cannot include household fixed effects. Instead, we include commune fixed effects ($\gamma_{c(i)}$)—the most granular geographic unit available—along with year fixed effects (θ_t) to control for unobserved heterogeneity across communes and over time.

The control vector X_{it} includes individual-level characteristics such as age, age squared, education level (and its square), and marital status, as well as household-level characteristics including household size (and its square), household composition measures such as the female share, college-educated share, adult share, and the household's average age and education, along with their squared terms. Consistent with the firm-level specification in equation (2), we include interactions between $\log(\text{DistNode}_i)$ and a post-2007 dummy ($Post07_t$), and between node-specific fixed effects ($\delta_{n(i)}$) and $Post07_t$. We also control for the average population density

[13] Although a household panel exists between the 2002–2004, 2004–2006, and 2006–2008 waves of the VHLSS (McCaig, 2021), we are unable to construct a sufficiently long panel spanning 2007—the year of the policy shock—with adequate variation in our ICT measures of interest.

within the connectivity buffer ($\log(\text{pop}_{it})$), and include the distance from each household to the nearest urban center (DistCBD$_i$), interacted with year fixed effects, to flexibly account for trends by urban proximity.

We instrument both ICT_{it} and its interaction with gender ($ICT_{it} \times Female_i$) using the same approach as in the firm-level analysis, as detailed in Section 4. Specifically, we use geographic characteristics of the *connectivity buffer* interacted with a post-2007 indicator. To enhance instrument relevance and mitigate overfitting, again, we employ a post-double-selection LASSO procedure to select the optimal subset of geographic variables to serve as instruments.

Results Table 6 presents the IV estimates for each labor market outcome listed in the column headers. Relative to the firm-level analysis, the instruments are notably weaker in predicting household-level ICT access: in all columns, the Kleibergen-Paap Wald F statistics fall below the conventional threshold of 10, raising concerns about weak instruments. To address this, we report Anderson-Rubin Wald test results, which are robust to weak instrument bias. These tests reject the null hypothesis at the 5% level in columns 1, 4, and 5, indicating statistically significant effects for those outcomes despite the weak first stage. Moreover, the Kleibergen-Paap LM underidentification tests reject the null across all columns, confirming that the models are identified. However, the results–particularly those in columns 2 and 3—should be interpreted with caution given the limited first-stage strength.

Across all five outcomes, we find limited evidence that ICT adoption directly benefits men: the main effect is statistically insignificant in almost all columns. However, we find that household-level ICT adoption leads to economically and statistically significant improvements in women's labor market outcomes relative to men. At the 10% significance level, ICT adoption increases the likelihood of formal employment for women by 14 percentage points more than it does for their male counterparts (column 1). This is a substantial gain, given the advantages of formal employment in terms of job stability, earnings, and legal protections in developing countries (Heckman and Pagés, 2003; Besley and Burgess, 2004). ICT adoption also helps narrow gender gaps in other dimensions: ICT access increases women's monthly workdays by 1.88 relative to men, statistically significant at the 10% level (column 2). However, we do not find a significant

gender differential effects of ICT in holding a second job (column 3).

In terms of time allocation, ICT adoption reduces women's likelihood of engaging in any housework by 23 percentage points relative to men (column 4). Conditional on engaging in housework, it also reduces the number of hours women spend per day on such tasks by 2.3 hours (column 5). Both effects are statistically and economically significant, indicating that ICT access not only narrows gender disparities in labor market outcomes but also reduces women's time allocation in home production.

6 Mechanism

Why does ICT expansion increase the female labor share and disproportionately improve women's labor market outcomes? While multiple channels may be at play, we focus on one key mechanism: ICT adoption changes production processes in ways that raise the relative productivity of tasks in which women hold a comparative advantage. This potential mechanism builds on two strands of literature: (1) how ICT reshapes firms' task and skill demands, and (2) gender differences in task-based comparative advantage.

We focus on five categories of task intensity that capture the nature of work activities across occupations: routine, manual, social, customer-facing, and service-oriented tasks. Routine tasks involve repetitive, rule-based activities such as data entry or bookkeeping. Manual tasks require physical effort, including lifting, operating machinery, or construction work. Social tasks rely on interpersonal interaction and communication, such as teaching, coordinating, or negotiating. Customer-facing tasks involve direct engagement with clients or patrons, including roles in sales, retail, or hospitality. Service-oriented tasks emphasize care and support, often combining social and customer interactions, and are common in healthcare, education, and personal services.

First, ICT tends to substitute for routine and manual labor while complementing non-routine, information-intensive, and socially interactive tasks. Prior studies show that ICT automates routine cognitive tasks such as clerical work and repetitive production, leading to reduced demand for these occupations, especially in manufacturing and data entry (Autor et al.,

2003; Acemoglu and Restrepo, 2020). Similarly, physical manual tasks are increasingly displaced by technology (Hötte et al., 2023). In contrast, ICT complements social and service-oriented roles. Deming (2017) documents rising demand for jobs requiring interpersonal interaction and teamwork, such as management and customer service. ICT also enhances the effectiveness of human interaction in customer-facing and low-skill service occupations, including retail, food service, and personal care (Autor and Dorn, 2013).

Second, women are often found to have a comparative advantage in cognitive ("brain") tasks, particularly those involving social, service, and customer-oriented skills, over routine and manual ("brawn") jobs (Welch, 2000; Rendall, 2018). Neuroscience and psychology studies suggest that women, on average, outperform men in emotional recognition and social sensitivity (Rueckert and Naybar, 2008; Schulte-Rüther et al., 2008; Cortes and Pan, 2018). This aligns with occupational sorting patterns: women are disproportionately employed in healthcare, education, and customer service—sectors characterized by high demand for non-routine, social, and emotional skills (Blau and Kahn, 2017). In contrast, women are underrepresented in routine manual sectors more vulnerable to automation (Autor et al., 2003).

Taken together, ICT adoption increases the relative demand for non-routine, social, and service tasks—precisely the areas where women hold a comparative advantage. In particular, we expect to see the following two patterns if this mechanism is at work: First, a firm's adoption of ICT increases the share of female employment more in industries with lower routine and manual task intensity and higher service, customer-facing, and social task intensity—areas in which women have a comparative advantage. Second, households' ICT adoption increases women's likelihood—relative to men—of working in occupations with lower routine and manual task intensity and higher service, customer-facing, and social task intensity.

To examine these patterns, we link occupational-level task intensity measures with Vietnam's household and firm survey data. These task measures, based on U.S. O*NET data and established task-based frameworks, capture the extent to which different occupations—and, by extension, industries—rely on routine, manual, social, customer-facing, and service-oriented tasks (U.S. Department of Labor, Employment and Training Administration, 2025). For the

household analysis, we merge these measures using workers' reported occupations. For the firm-level analysis, due to limited occupational detail, we construct industry-level task intensities by combining O*NET task scores with occupation shares from Indonesia's 2010 National Labor Force Survey, which offers a comparable economic and industrial structure to Vietnam. Further details on task construction and classification are provided in the Appendix Section B.

We then test these mechanisms using both firm-level and household-level data. First, at the firm level, we examine whether the effect of ICT adoption on female labor share varies across industries with different task compositions. Second, at the individual level, we assess whether household ICT adoption disproportionately shifts women's occupational choices—relative to men's—toward tasks more aligned with their comparative advantage.

Gender sorting by occupational task intensity In this section, we provide suggestive evidence that female workers have comparative advantages in social, customer-facing, and service-oriented tasks using Vietnam data. Figure 5 illustrates the correlation between the industry-level average female labor share and industry-level task intensity. Each panel presents a scatterplot and fitted linear trend line, where the y-axis denotes the female labor share and the x-axis represents a specific type of task intensity. Each dot corresponds to an industry.[14]

We find that female labor share is positively correlated with service, customer, and social task intensity (Panels E, D, and C), with correlations of 0.237, 0.203, and 0.129, respectively. In contrast, the relationship is weakly negative for routine tasks (Panel A, correlation = –0.049) and strongly negative for manual tasks (Panel B, correlation = –0.344). In essence, industries that rely more heavily on interpersonal and service-related skills tend to employ more women, whereas industries that require physically demanding, manual labor tend to employ fewer. These patterns are consistent with the hypothesis that women have a comparative advantage in tasks involving communication, social interaction, and service provision.

[14] As described in Section B, we construct industry-level task intensity measures for Vietnamese firms by mapping U.S. O*NET occupational task scores to Vietnamese industries. While occupation-level data would be preferable, the Vietnam Enterprise Surveys report only industry codes. Under the assumption of no factor intensity reversal, we follow established procedures to aggregate task intensity to the industry level, as detailed in Section B.

Firm Level Evidence Given the observed patterns of gender sorting by task intensity across sectors, we next examine how ICT expansion affects the female labor share at the firm level. As ICT adoption tends to complement areas where women have a comparative advantage—we expect heterogeneous effects across industries with varying levels of task intensity. To provide suggestive evidence for this mechanism, we estimate a modified specification that allows for interaction between firm-level ICT adoption and industry-level task intensity:

$$\begin{aligned} Y_{it} = \alpha_0 + \beta_1 ICT_{it} + \gamma_1 T_j + \gamma_2 ICT_{it} \times T_j + \gamma_i + \theta_t + \beta_2 X_{it} \\ + \beta_3 \log(DistNode_i) \times Post07_t + \beta_4 \delta_{n(i)} \times Post07_t \\ + \beta_5 \log(DistCBD_i) \times Year_t + \beta_6 \log(pop)_{it} + \beta_7 Sector_i \times Year_t + \epsilon_{it} \end{aligned} \quad (6)$$

The dependent variable Y_{it} denotes the female labor share of firm i in year t. Relative to the baseline specification in equation (2), we augment the model by including an interaction between firm-level ICT adoption proxied by log number of PCs with internet per worker, ICT_{it}, and industry-level task intensity, T_j, derived from O*NET data as detailed in Section 3. In the first stage, we instrument for both $ICT_{it} \times Post07_t$ and $ICT_{it} \times Post07_t \times T_j$ using the corresponding interactions of our geographic instruments–characteristics of the connectivity buffer–with $Post07_t$ and $Post07_t \times T_j$. The specification includes the same set of controls as in equation (2). In addition, we control for sector-specific time trends, $Sector_i \times Year_t$, to account for differential trends in female employment and ICT adoption across sectors that are unrelated to our mechanism. Here, sectors refer to manufacturing and non-manufacturing, a more aggregated classification than the 3-digit industry used to define task intensity. This control is particularly important because task intensity is highly correlated with sectors, and we do not have valid instruments for task intensity. Without controlling for this, the estimated interaction effects could be confounded by sectors that are simultaneously becoming more female-intensive and adopting ICT for reasons unrelated to task-based comparative advantage.

Table 7 presents the 2SLS estimates of the effect of ICT adoption on the female labor share,

allowing the effect to vary by industry-level task intensity. Column 1 shows that our baseline result remains robust when we include sector-by-year fixed effects. The estimated coefficient on ICT adoption ($\beta_1 = 0.406$) is statistically significant at the 1% level and slightly larger than the baseline estimate without these additional controls.

Columns 2–6 examine heterogeneity in the impact of ICT adoption by interacting the ICT measure with five alternative task intensity indicators—routine, manual, social, customer-facing, and service-oriented—constructed from O*NET data as described in Section 3. Each specification reports the estimated coefficients on β_1 (the main effect of ICT_{it}) and γ_2 (the interaction between ICT_{it} and T_j), capturing how the effect of ICT adoption varies by task intensity.

Across all specifications, we find consistent evidence that ICT adoption leads to larger increases in the female labor share in industries that rely more heavily on tasks where women are likely to hold a comparative advantage. Specifically, column 2 and 3 shows that the interactions between ICT adoption and routine and manual task intensity are both negative and significant, indicating that ICT adoption has a smaller impact on female labor share in more routine-intensive industries. In contrast, columns 4 to 6 show positive and statistically significant interactions between ICT and social, customer, and service task intensities. These results suggest that ICT adoption shifts production in ways that increase the relative productivity of non-manual, interpersonal, and service-oriented tasks—areas where women tend to have a comparative advantage.

Individual Level Evidence Due to limitations in the firm-level data, task intensity can only be measured at the 3-digit industry level, preventing us from capturing more granular variations at the occupation level. To address this, we supplement our analysis with data VHLSS, which reports individuals' occupations and allows us to examine occupation-level variation in task intensity. We harmonize occupational codes across VHLSS survey waves and map them to the 1990 U.S. Census occupational codes developed by Dorn (2009), which we then link to task intensity measures from Autor et al. (2003) and Deming (2017) which are based on the O*NET database. Using these individual-level task intensity measures as dependent variables, we estimate specification (5) to examine the gender-differentiated effects of ICT expansion at

the individual level.

Table 8 presents results from individual-level regressions linking household ICT access to occupational task intensity, allowing for gender-specific effects. The interaction terms with the female indicator reveal significant gender differences: negative for routine and manual tasks and positive for social, customer, and service tasks respectively (all significant at 1% or 5%), indicating that women in internet-connected households are less likely to engage in routine and manual tasks and more likely to work in socially, customer-facing, or service-oriented occupations, compared to men. These results reinforce the firm-level evidence and suggest that ICT adoption facilitates female employment in occupations emphasizing interpersonal and service-oriented tasks, where women tend to have a comparative advantage.

7 Conclusion

This paper investigates the impact of information and communication technologies (ICT) on female labor market outcomes, focusing on Vietnam—a rapidly growing developing economy that underwent a large-scale expansion of its ICT infrastructure in 2007. Using detailed firm- and household-level data, we exploit exogenous variation in access to ICT infrastructure, arising from both the timing of the 2007 broadband expansion and the geographic constraints of connecting to backbone nodes. Our instrumental variable estimates show that improved ICT access significantly increases the female employment share within firms. At the household level, ICT adoption raises women's likelihood of working in the formal sector, monthly workdays, and reduces time spent on unpaid domestic work.

These results are strongest in sectors characterized by low routine intensity and high reliance on social and service-oriented tasks. These patterns support a task-based mechanism: ICT complements non-routine, interpersonal, and cognitively demanding activities—domains in which women tend to have comparative advantages. Individual-level analysis further confirms this mechanism.

Our findings suggest that expanding access to broadband ICT can be an effective tool for reducing gender disparities in labor markets, particularly by promoting women's participation

in formal, service-oriented employment. Policymakers may consider digital infrastructure not only as a productivity investment but also as a potential tool to reduce gender disparities. To maximize impact, infrastructure efforts could be paired with complementary policies that help women capitalize on new opportunities created by technological change. While our analysis is based on Vietnam, the underlying mechanisms—task reallocation and reduced frictions in labor market access—are likely to apply to other countries undergoing similar digital transitions.

Declaration of generative AI and AI-assisted technologies in the writing process *During the preparation of this work the authors used ChatGPT in order to proofread and improve the language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.*

References

ACEMOGLU, D. AND P. RESTREPO (2020): "Robots and jobs: Evidence from US labor markets," *Journal of political economy*, 128, 2188–2244.

AKERMAN, A., I. GAARDER, AND M. MOGSTAD (2015): "The skill complementarity of broadband internet," *The Quarterly Journal of Economics*, 130, 1781–1824.

AKSOY, C. G., B. ÖZCAN, AND J. PHILIPP (2021): "Robots and the gender pay gap in Europe," *European Economic Review*, 134, 103693.

AUTOR, D., D. DORN, G. HANSON, AND K. MAJLESI (2020): "Importing political polarization? The electoral consequences of rising trade exposure," *American Economic Review*, 110, 3139–3183.

AUTOR, D. H. AND D. DORN (2013): "The growth of low-skill service jobs and the polarization of the US labor market," *American economic review*, 103, 1553–1597.

AUTOR, D. H., L. F. KATZ, AND A. B. KRUEGER (1998): "Computing inequality: have computers changed the labor market?" *The Quarterly journal of economics*, 113, 1169–1213.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The skill content of recent technological change: An empirical exploration," *The Quarterly journal of economics*, 118, 1279–1333.

AZAM, M., M. S. EMRAN, AND F. SHILPI (2024): "ICT Skills and Labor Market Outcomes in India: Evidence from Cell Tower Expansion," Tech. rep., IZA Discussion Papers.

BALDWIN, R. (2013): "Trade and industrialization after globalization's second unbundling: How building and joining a supply chain are different and why it matters," in *Globalization in an age of crisis: Multilateral economic cooperation in the twenty-first century*, University of Chicago Press, 165–212.

BANERJI, S. AND R. S. CHOWDHURY (2013): "Wi-Fi & WiMAX: A Comparative Study," *arXiv preprint arXiv:1302.2247*.

BARTEL, A., C. ICHNIOWSKI, AND K. SHAW (2007): "How does information technology affect productivity? Plant-level comparisons of product innovation, process improvement, and worker skills," *The quarterly journal of Economics*, 122, 1721–1758.

BEAUDRY, P. AND E. LEWIS (2014): "Do male-female wage differentials reflect differences in the return to skill? Cross-city evidence from 1980–2000," *American Economic Journal: Applied Economics*, 6, 178–194.

BELLONI, A., D. CHEN, V. CHERNOZHUKOV, AND C. HANSEN (2012): "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, 80, 2369–2429.

BESLEY, T. AND R. BURGESS (2004): "Can Labor Regulation Hinder Economic Performance? Evidence from India," *The Quarterly Journal of Economics*, 119, 91–134.

BIGGS, A. AND K. MAHONY (2004): "Is soil science relevant to road infrastructure?" in *13th International Soil Conservation Organisation Conference, Conserving soil and water for society: Sharing solutions, Brisbane, ISCO*.

BLACK, S. E. AND A. SPITZ-OENER (2010): "Explaining women's success: technological change and the skill content of women's work," *The Review of Economics and Statistics*, 92, 187–194.

BLAU, F. D. AND L. M. KAHN (2017): "The gender wage gap: Extent, trends, and explanations," *Journal of economic literature*, 55, 789–865.

BORGHANS, L., A. L. DUCKWORTH, J. J. HECKMAN, AND B. TER WEEL (2008): "The economics and psychology of personality traits," *Journal of human Resources*, 43, 972–1059.

BRESNAHAN, T. F., E. BRYNJOLFSSON, AND L. M. HITT (2002): "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence," *The quarterly journal of economics*, 117, 339–376.

CAMPANTE, F. R. AND D. CHOR (2012): "Why was the Arab world poised for revolution? Schooling, economic opportunities, and the Arab Spring," *Journal of Economic Perspectives*, 26, 167–188.

CHIPLUNKAR, G. AND P. K. GOLDBERG (2022): "The employment effects of mobile internet in developing countries," Tech. rep., National Bureau of Economic Research.

CHUNG, S., R. M. GONZALEZ, AND A. S. MOBARIZ (2025): "Mobile Phones and Gender Norms," Manuscript.

CORTÉS, P., Y. FENG, N. GUIDA-JOHNSON, AND J. PAN (2024): "Automation and Gender: Implications for Occupational Segregation and the Gender Skill Gap," Tech. rep., National Bureau of Economic Research.

CORTES, P. AND J. PAN (2018): "Occupation and gender," *The Oxford handbook of women and the economy*, 425–452.

CRANDALL, R. W., W. LEHR, AND R. E. LITAN (2007): *The effects of broadband deployment on output and employment: A cross-sectional analysis of US data*, vol. 6, Brookings Institution.

DEMING, D. J. (2017): "The growing importance of social skills in the labor market," *The Quarterly Journal of Economics*, 132, 1593–1640.

DORN, D. (2009): "Essays on inequality, spatial interaction, and the demand for skills," Ph.D. thesis, Verlag nicht ermittelbar.

ELLIS, R., R. M. GONZALEZ, AND L. YANG (2025): "Mobile Phone Access and Insurgent Violence: Evidence from an Interval Regression Discontinuity Design in Afghanistan," Manuscript.

FARR, T., P. ROSEN, E. CARO, R. CRIPPEN, R. DUREN, S. HENSLEY, M. KOBRICK, M. PALLER, E. RODRIGUEZ, L. ROTH, D. SEAL, S. SHAFFER, J. SHIMADA, J. UMLAND, M. WERNER, M. OSKIN, D. BURBANK, AND D. ALSDORF (2007): "The shuttle radar topography mission," *Reviews of Geophysics*, 45, RG2004.

FIBER BROADBAND ASSOCIATION AND CARTESIAN (2023): "2023 Fiber Deployment Cost Study," https://fiberbroadband.org/wp-content/uploads/2024/01/Fiber-Deployment-Annual-Report-2023_FBA-and-Cartesian.pdf, accessed: 2025-04-21.

FORMAN, C., A. GOLDFARB, AND S. GREENSTEIN (2012): "The Internet and local wages: A puzzle," *American Economic Review*, 102, 556–75.

GERACI, A., M. NARDOTTO, T. REGGIANI, AND F. SABATINI (2022): "Broadband internet and social capital," *Journal of Public Economics*, 206, 104578.

GOLDIN, C. (2006): "The quiet revolution that transformed women's employment, education, and family," *American economic review*, 96, 1–21.

——— (2014): "A grand gender convergence: Its last chapter," *American economic review*, 104, 1091–1119.

GREENWOOD, J., N. GUNER, G. KOCHARKOV, AND C. SANTOS (2016): "Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation," *American Economic Journal: Macroeconomics*, 8, 1–41.

HECKMAN, J. J. AND C. PAGÉS (2003): "Law and employment: Lessons from Latin America and the Caribbean," .

HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor economics*, 24, 411–482.

HENGL, T., G. YENGOH, M. DE SOUZA, A. DE PINTO, J. TIEMANN, E. BUI, J. LEENAARS, M. KILIBARDA, AND G. HEUVELINK (2017): "SoilGrids: Global soil information based on automated mapping," *PLOS ONE*, 12, e0169746.

HJORT, J. AND J. POULSEN (2019): "The arrival of fast internet and employment in Africa," *American Economic Review*, 109, 1032–79.

HJORT, J. AND L. TIAN (2021): "The economic impact of internet connectivity in developing countries," *Annual Review of Economics*, 17.

HÖTTE, K., M. SOMERS, AND A. THEODORAKOPOULOS (2023): "Technology and jobs: A systematic literature review," *Technological Forecasting and Social Change*, 194, 122750.

HSIEH, C., H. CHANGTAI, ERIK, C. JONES, J. JONES, PETER, AND KLENOW (2019): "The Allocation of Talent and U.S. Economic Growth," *Economica*, 87, 1435–1474.

INTERNATIONAL MONETARY FUND (2018): "Pursuing Women's Economic Empowerment," Tech. rep., International Monetary Fund, accessed: 2023.

INTERNATIONAL TELECOMMUNICATION UNION (ITU) (2025): "Individuals using the Internet (% of population): Vietnam [Indicator IT.NET.USER.ZS]," World Development Indicators, The World Bank, retrieved June 10, 2025, from https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=VN.

JUHN, C., G. UJHELYI, AND C. VILLEGAS-SANCHEZ (2014): "Men, women, and machines: How trade impacts gender inequality," *Journal of Development Economics*, 106, 179–193.

KLEIBERGEN, F. AND R. PAAP (2006): "Generalized reduced rank tests using the singular value decomposition," *Journal of econometrics*, 133, 97–126.

KWON, E. (2022): "Why do improvements in transportation infrastructure reduce the gender gap in south korea," *Available at SSRN 4081288*.

LINDQVIST, E. AND R. VESTMAN (2011): "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment," *American Economic Journal: Applied Economics*, 3, 101–128.

MCCAIG, B. (2021): "Notes on the VHLSS," Accessed: 2025-06-07.

MCCAIG, B. AND N. PAVCNIK (2018): "Export markets and labor allocation in a low-income country," *American Economic Review*, 108, 1899–1941.

——— (2021): "Entry and exit of informal firms and development," *IMF Economic Review*, 69, 540–575.

MICHAELS, G., A. NATRAJ, AND J. VAN REENEN (2014): "Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years," *Review of Economics and Statistics*, 96, 60–77.

MINER, L. (2015): "The unintended consequences of Internet diffusion: Evidence from Malaysia," *Journal of Public Economics*, 132, 66–78.

NGAI, L. R. AND B. PETRONGOLO (2017): "Gender gaps and the rise of the service economy," *American Economic Journal: Macroeconomics*, 9, 1–44.

NUNN, N. AND D. PUGA (2012): "Ruggedness: The blessing of bad geography in Africa," *Review of Economics and Statistics*, 94, 20–36.

NUNN, N. AND N. QIAN (2011): "The potato's contribution to population and urbanization: evidence from a historical experiment," *The quarterly journal of economics*, 126, 593–650.

OLIVETTI, C. AND B. PETRONGOLO (2016): "The evolution of gender gaps in industrialized countries," *Annual review of Economics*, 8, 405–434.

PEKEL, J.-F., A. COTTAM, N. GORELICK, AND A. S. BELWARD (2016): "High-resolution mapping of global surface water and its long-term changes," *Nature*, 540, 418–422.

PIKETTY'S, T. (2014): "Capital in the twenty-first century," *Perspectives on Politics*.

RENDALL, M. (2017): "Brain versus brawn: the realization of women's comparative advantage," *University of Zurich, Institute for Empirical Research in Economics, Working Paper*.

——— (2018): "Female market work, tax regimes, and the rise of the service sector," *Review of Economic Dynamics*, 28, 269–289.

RUECKERT, L. AND N. NAYBAR (2008): "Gender differences in empathy: The role of the right hemisphere," *Brain and cognition*, 67, 162–167.

SCHIAVINA, M., S. FREIRE, A. CARIOLI, AND K. MACMANUS (2023): "GHS-POP R2023A - GHS population grid multitemporal (1975-2030)," https://ghs-pop.jrc.ec.europa.eu/, accessed: 2023-04-14.

SCHULTE-RÜTHER, M., H. J. MARKOWITSCH, N. J. SHAH, G. R. FINK, AND M. PIEFKE (2008): "Gender differences in brain networks supporting empathy," *Neuroimage*, 42, 393–403.

TUAN, T. M. (2011): "Vietnam-Broadband in Vietnam: forging its own path," *World Bank Group, Broadband Strategies Toolkit*.

U.S. DEPARTMENT OF LABOR, EMPLOYMENT AND TRAINING ADMINISTRATION (2025): "O*NET Database," Available at https://www.onetonline.org/.

VIDART, D. (2024): "Human capital, female employment, and electricity: Evidence from the early 20th-century United States," *Review of Economic Studies*, 91, 560–594.

VNNIC (2013): "Report on Vietnam Internet Resources," *Vietnam Internet Network Information Center*.

WELCH, F. (2000): "Growth in women's relative wages and in inequality among men: One phenomenon or two?" *American Economic Review*, 90, 444–449.

WORLD ECONOMIC FORUM (2023): "Global Gender Gap Report 2023," Tech. rep., World Economic Forum, accessed: 2023.

Table 1: SUMMARY STATISTICS ON ICT VARIABLES, FIRM LEVEL

Panel A: Firm Level					
	2004	2005	2007	2008	2009
Number of PC with Internet per Worker	0.05	0.07	0.12	0.14	0.16
	(0.15)	(0.19)	(0.22)	(0.23)	(0.25)
Number of PC with LAN per Worker	0.06	0.07	0.11	0.12	0.13
	(0.06)	(0.07)	(0.11)	(0.12)	(0.13)
Observations	32,840	41,093	44,858	40,355	43,742
Panel B: Household Level					
	2002	2004	2006	2008	
Linked to Internet (1 0)	0.003	0.009	0.013	0.040	
	(0.058)	(0.094)	(0.111)	(0.197)	
Individual N	131,864	36,659	38,262	185,220	
Household N	29,412	9,016	9,003	44,505	

Notes: This table presents summary statistics of the ICT adoption measures used in our analysis, disaggregated by year. Column headers indicate the survey year; each cell reports the mean, with the corresponding standard deviation shown in parentheses below. Panel A summarizes firm-level ICT adoption based on the Annual Enterprise Census of Vietnam. Panel B reports household-level ICT adoption using data from the Vietnam Household Living Standards Survey (VHLSS).

Table 2: FIRM AND INDIVIDUAL DISTANCE TO NEAREST DOMESTIC BACKBONE NODE SITE

Panel A	**Total Admin Units**	
	Firm Level	Individual Level
Number of Provinces	64	63
Number of Districts	665	640
Number of Communes	6,989	3,246

Panel B

	Firm Level		**Individual Level**	
	Total	Percentage	Total	Percentage
# located within 5kms from the backbone	27,922	28%	28,484	8%
# located within 10kms from the backbone	56,967	59%	91,459	24%
# located within 15kms from the backbone	75,079	78%	162,453	43%
# located within 20km from the backbone	84,429	88%	216,407	57%
# located within 25km from the backbone	89,182	92%	266,269	70%
Total	96,492	100%	377,663	100%

Notes: Panel A reports the total number of administrative units covered in our sample at the province, district, and commune levels, separately for the firm-level and household-level data. Panel B presents the total counts and cumulative percentages of firms and individuals across distance bins relative to the nearest domestic backbone node site.

Table 3: FIRST STAGE, LASSO

	# PCs w/ internet per worker (log)	# PCs w/ LAN per worker (log)
	(1)	(2)
Soil Sand (min)	0.000**	
	(0.000)	
Soil Coarse fragments (min)	0.001***	0.001***
	(0.000)	(0.000)
Soil Bulk density (min)	-0.000***	-0.000**
	(0.000)	(0.000)
The frequency water was present (mean)	-0.001***	-0.000
	(0.000)	(0.000)
The frequency water was present (std)	0.002***	0.001***
	(0.000)	(0.000)
N	184,291	179,906
N Clusters	53,152	52,343
Adj. R^2	0.702	0.673
Regression F-Stat	235.113	106.358
Firm FE	Yes	Yes
Year FE	Yes	Yes
Firm Level Control	Yes	Yes
Node ID $\times$ Post 2007	Yes	Yes
Distance to Node $\times$ Post 2007	Yes	Yes
Control Population	Yes	Yes
Distance Urban Center $\times$ Year	Yes	Yes

Notes: This table reports estimates of equation (3), which models the first-stage relationship between ICT adoption and the geographic characteristics of the *connectivity buffer*, as selected by the post-double-selection LASSO procedure. Column 1 uses the log number of PCs connected to the Internet, and Column 2 uses the log number of PCs connected to a LAN as the dependent variable. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the firm level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 4: Firm Female Labor Share and ICT Adoption

Panel A	**(1) OLS**	**(2) IV**	**(3) IV**	**(4) IV**
Number of PCs with internet per worker (log)	0.062***	0.328***	0.327***	0.356***
	(0.005)	(0.120)	(0.118)	(0.115)
N	180,791	180,791	180,791	180,791
N Clusters	52,248	52,248	52,248	52,248
Kleibergen-Paap Wald Rank F Stat		30.505	31.134	33.879
Under Id. Test (KP Rank LM Stat)		153.879	157.080	170.486
p-Value		0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)		1.624	1.656	2.097
p-Value		0.150	0.141	0.063
Panel B	**(1) OLS**	**(2) IV**	**(3) IV**	**(4) IV**
Number of PCs with LAN per worker (log)	0.039***	0.434***	0.433***	0.473***
	(0.004)	(0.164)	(0.162)	(0.159)
N	176,515	176,515	176,515	176,515
N Clusters	51,438	51,438	51,438	51,438
Kleibergen-Paap Wald Rank F Stat		18.171	18.545	20.146
Under Id. Test (KP Rank LM Stat)		72.814	74.322	80.640
p-Value		0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)		1.959	1.995	2.477
p-Value		0.098	0.092	0.042
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm Level Control	Yes	Yes	Yes	Yes
Node ID × Post 2007	Yes	Yes	Yes	Yes
Distance to Node × Post 2007	Yes	Yes	Yes	Yes
Control Population			Yes	Yes
Distance to Urban Center × Year				Yes

Notes: Each cell reports the estimated coefficient on ICT adoption from equation (1) or (2), with the specific controls listed in the column headings. Column 1 presents OLS estimates, whereas Columns 2–4 provide post-double-selection IV-Lasso estimates. Panel A uses the log number of PCs connected to the Internet, and Panel B uses the log number of PCs connected to a LAN as the dependent variable. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the firm level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 5: PRETREND OF ICT ADOPTION

Panel A:	**# PCs w/ internet per worker (log)**	**# PCs w/ LAN per worker (log)**
	(1)	**(2)**
year=2004 × Soil Sand (min)	0.000*	
	(0.000)	
year=2005 × Soil Sand (min)	0.000	
	(0.000)	
year=2008 × Soil Sand (min)	0.000***	
	(0.000)	
year=2009 × Soil Sand (min)	0.000**	
	(0.000)	
year=2004 × Soil Coarse fragments (min)	-0.000	0.000
	(0.000)	(0.000)
year=2005 × Soil Coarse fragments (min)	0.000	0.000
	(0.000)	(0.000)
year=2008 × Soil Coarse fragments (min)	-0.000	0.000
	(0.000)	(0.000)
year=2009 × Soil Coarse fragments (min)	0.000	0.000**
	(0.000)	(0.000)
year=2004 × Soil Bulk density (min)	0.000	0.000
	(0.000)	(0.000)
year=2005 × Soil Bulk density (min)	0.000	0.000
	(0.000)	(0.000)
year=2008 × Soil Bulk density (min)	-0.000***	-0.000***
	(0.000)	(0.000)
year=2009 × Soil Bulk density (min)	-0.000***	-0.000***
	(0.000)	(0.000)

Table 5: PRETREND OF ICT ADOPTION (CONTINUED)

	# PCs w/ internet per worker (log)	# PCs w/ LAN per worker (log)
	(1)	(2)
year=2004 × The frequency water was present (mean)	-0.000	-0.000
	(0.000)	(0.000)
year=2005 × The frequency water was present (mean)	-0.000	-0.000
	(0.000)	(0.000)
year=2008 × The frequency water was present (mean)	0.000**	0.000**
	(0.000)	(0.000)
year=2009 × The frequency water was present (mean)	0.000	0.000
	(0.000)	(0.000)
year=2004 × The frequency water was present (std)	0.000	0.000
	(0.000)	(0.000)
year=2005 × The frequency water was present (std)	0.000	0.000
	(0.000)	(0.000)
year=2008 × The frequency water was present (std)	-0.001***	-0.001***
	(0.000)	(0.000)
year=2009 × The frequency water was present (std)	-0.000***	-0.001***
	(0.000)	(0.000)
N	273,262	273,262
N Clusters	66,305	66,305
Panel B: Joint Significance of Geo. Char. Interacted w/ Year (P val.)		
...× 2004	0.100	0.417
...× 2005	0.169	0.102
...× 2008	0.000	0.000
...× 2009	0.000	0.000
Firm FE	Yes	Yes
Year FE	Yes	Yes
Firm Level Control	Yes	Yes
Node ID × Post 2007	Yes	Yes
Distance to Node × Post 2007	Yes	Yes
Control Population	Yes	Yes
Distance Urban Center × Year	Yes	Yes

Notes: This table reports estimates from equation (4), which tests for pre-existing differential trends in female labor share based on the geographic characteristics of the *connectivity buffer*, selected using the post-double-selection LASSO procedure. Column 1 uses the set of instruments selected based on the log number of Internet-connected PCs per worker as the ICT adoption measure, while Column 2 uses those selected based on the log number of LAN-connected PCs per worker. Panel B test joint significance (P value) tests of the full set of β_{1t} coefficients of all instruments separately for each year. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the firm level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 6: INDIVIDUAL LABOR MARKET OUTCOMES AND ICT BY GENDER, VHLSS

	Formal (1 0) **(1)**	**Monthly Work Days** **(2)**	**Second Job (1 0)** **(3)**	**Housework (1 0)** **(4)**	**Housework Hrs** **(5)**
Linked to Internet (1 0)	-0.132	0.912	-0.424	0.873*	-0.035
	(0.422)	(5.241)	(0.472)	(0.476)	(1.184)
...× Female	0.144*	1.879*	-0.094	-0.234**	-2.285***
	(0.083)	(1.028)	(0.087)	(0.097)	(0.257)
N	37,980	54,429	54,429	54,429	43,339
N Clusters	21,278	31,150	31,150	31,150	27,061
Kleibergen-Paap Wald Rank F Stat	2.255	3.393	3.393	3.393	2.720
Under Id. Test (KP Rank LM Stat)	14.083	20.957	20.957	20.957	16.898
p-Value	0.015	0.001	0.001	0.001	0.005
AR Wald Test (Weak IV Robust Inf.)	3.041	1.624	0.878	10.792	50.348
p-Value	0.006	0.136	0.510	0.000	0.000
Commune FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Individual Level Control	Yes	Yes	Yes	Yes	Yes
Node ID × Post 2007	Yes	Yes	Yes	Yes	Yes
Distance to Node × Post 2007	Yes	Yes	Yes	Yes	Yes
Control Population	Yes	Yes	Yes	Yes	Yes
Distance Urban Center × Year	Yes	Yes	Yes	Yes	Yes

Notes: Each cell reports the estimated coefficient on ICT adoption and its interaction with the female indicator from equation (5), with the corresponding dependent variables listed in the column titles. The control variables are detailed in the bottom panel. All estimates are obtained using a post-double-selection IV-Lasso procedure to select the optimal instruments. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the household level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 7: DIFFERENTIAL IMPACT OF ICT BY INDUSTRY-LEVEL TASK INTENSITY

	(1)	(2)	(3)	(4)	(5)	(6)
Number of PCs with internet per worker (log)	0.406***	0.644***	0.443***	0.170	0.212	0.174
	(0.157)	(0.083)	(0.120)	(0.226)	(0.210)	(0.223)
...× Task Intensity: Routine		-0.076**				
		(0.032)				
...× Task Intensity: Manual			-0.095**			
			(0.047)			
...× Task Intensity: Social				0.033**		
				(0.013)		
...× Task Intensity: Customer					0.027***	
					(0.010)	
...× Task Intensity: Service						0.066***
						(0.023)
N	180,791	162,946	162,946	162,946	162,946	162,946
N Clusters	52,248	47,053	47,053	47,053	47,053	47,053
Kleibergen-Paap Wald Rank *F* Stat	19.600	7.492	6.750	8.490	8.572	8.666
Under Id. Test (KP Rank LM Stat)	101.419	77.038	69.942	87.711	88.394	89.388
p-Value	0.000	0.000	0.000	0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)	3.284	11.195	11.342	9.500	9.383	9.519
p-Value	0.006	0.000	0.000	0.000	0.000	0.000

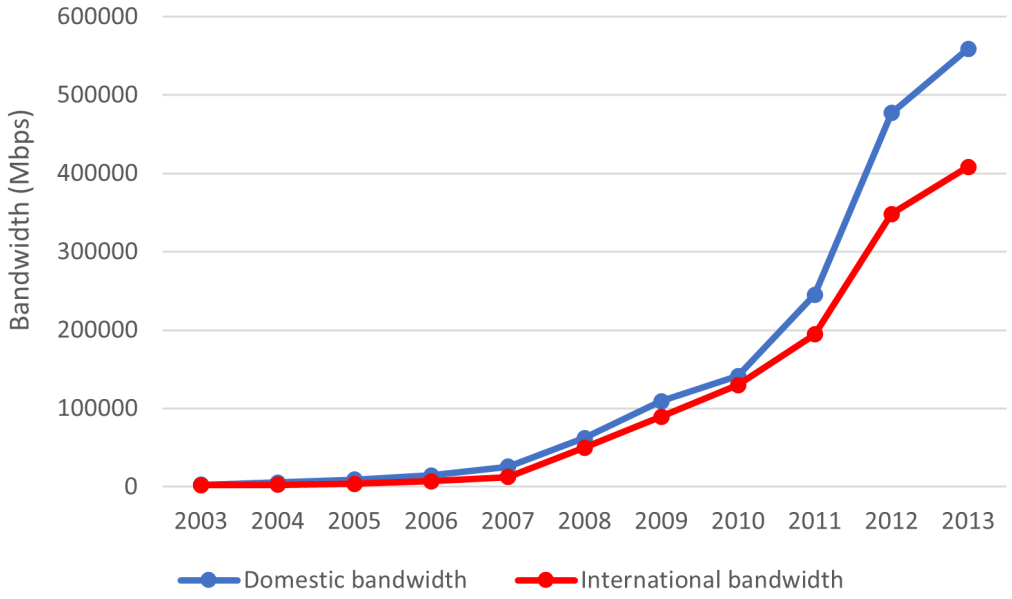
Notes: Each cell reports the estimated coefficient on ICT adoption and its interaction with task intensity measures from equation (6), with the specific controls indicated in the column headings. The occupational task intensity measures—routine, manual, social, customer, and service—are constructed by linking our firm-level data to O*NET-based task classifications, weighted by industry-level occupational shares from Indonesia's 2010 National Labor Survey. All estimates are obtained using a post-double-selection IV-Lasso procedure to select the optimal instruments. All estimates are obtained using a post-double-selection IV-Lasso procedure to select the optimal instruments. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the firm level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 8: INDIVIDUAL OCCUPATIONAL TASK INTENSITY BY GENDER, VHLSS

	Routine (1)	**Manual (2)**	**Social (3)**	**Customer (4)**	**Service (5)**
Linked to Internet (1 0)	-0.079	2.169*	-3.721	0.425	-1.744
	(2.135)	(1.285)	(2.659)	(2.365)	(1.510)
...× Female	-1.603**	-1.733***	2.022**	1.583**	3.775***
	(0.704)	(0.420)	(0.851)	(0.764)	(0.530)
N	49,567	49,567	49,567	49,567	49,567
N Clusters	29,147	29,147	29,147	29,147	29,147
Kleibergen-Paap Wald Rank F Stat	4.630	4.630	4.630	4.630	4.630
Under Id. Test (KP Rank LM Stat)	28.695	28.695	28.695	28.695	28.695
p-Value	0.000	0.000	0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)	3.782	3.283	2.688	2.688	21.953
p-Value	0.001	0.003	0.013	0.013	0.000
Commune FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Individual Level Control	Yes	Yes	Yes	Yes	Yes
Node ID × Post 2007	Yes	Yes	Yes	Yes	Yes
Distance to Node × Post 2007	Yes	Yes	Yes	Yes	Yes
Control Population	Yes	Yes	Yes	Yes	Yes
Distance Urban Center × Year	Yes	Yes	Yes	Yes	Yes

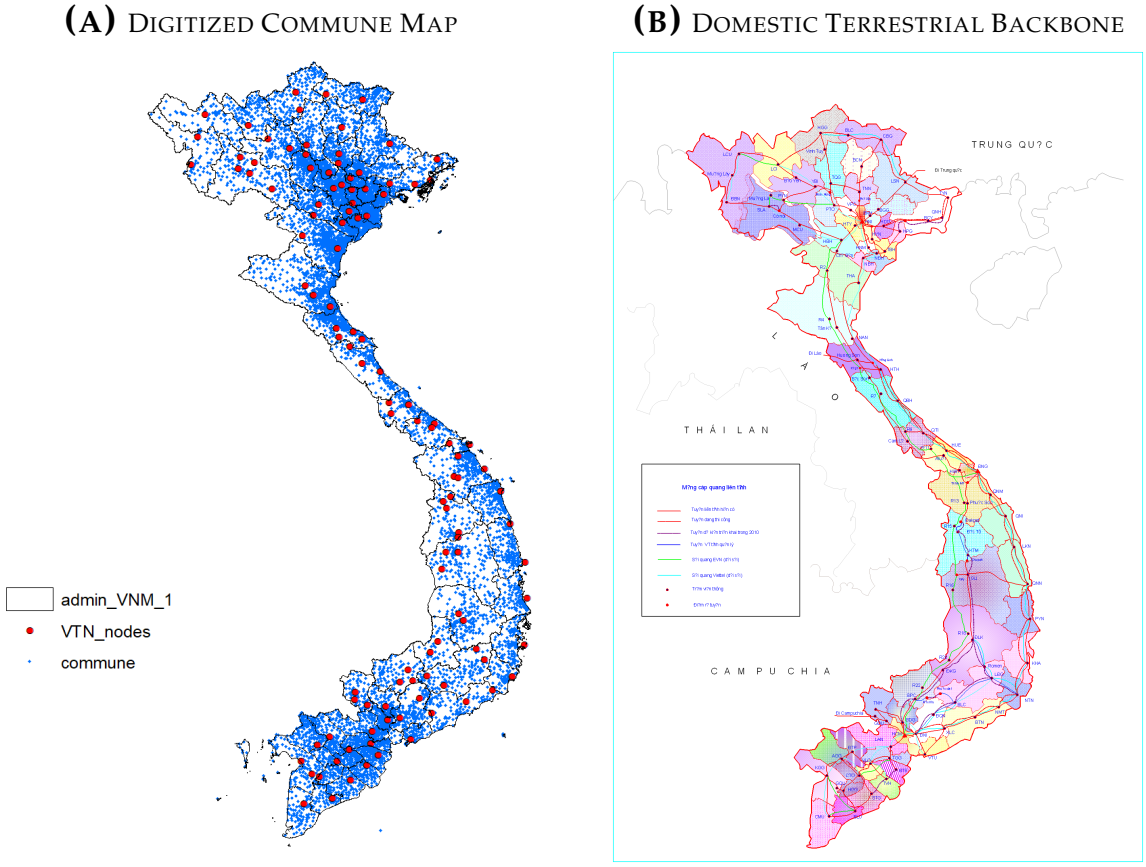
Notes: Each cell reports the estimated coefficient on ICT adoption and its interaction with a female indicator from equation (5), with the corresponding dependent variables listed in the column titles. The outcome variables are occupational task intensity measures—routine, manual, social, customer, and service—constructed by linking VHLSS occupational codes to O*NET-based task measures following Autor et al. (2003) and Deming (2017). Control variables are detailed in the bottom panel. All estimates are obtained using a post-double-selection IV-Lasso procedure to select the optimal instruments. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the household level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Figure 1: Domestic and International Bandwidth



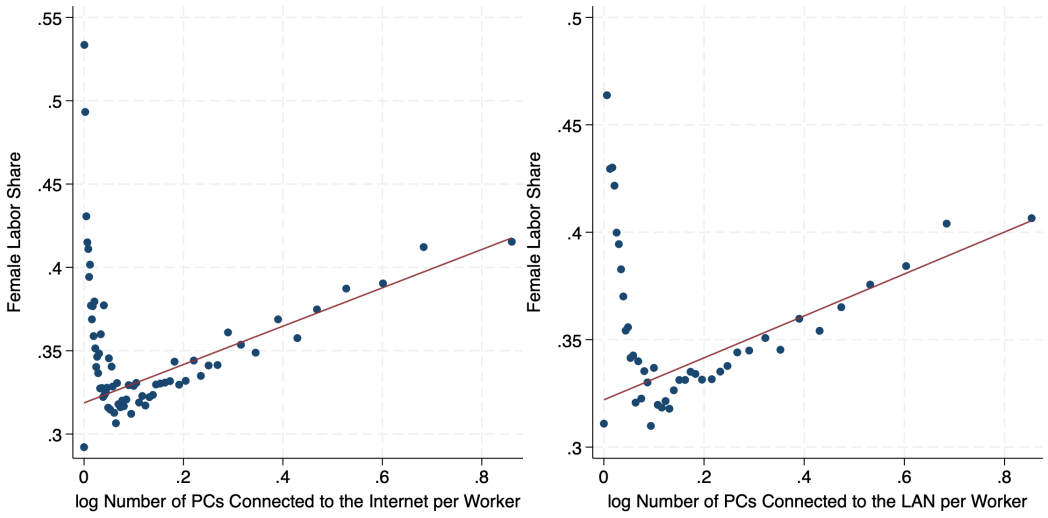
Notes: This figure plots the average domestic and international bandwidth in Vietnam from 2003 to 2013, using data from the Vietnam Internet Network Information Center (VNNIC, 2013). The red line represents international bandwidth, while the blue line represents domestic bandwidth.

Figure 2: Main Terrestrial Backbone Node Sites and Vietnamese Communes



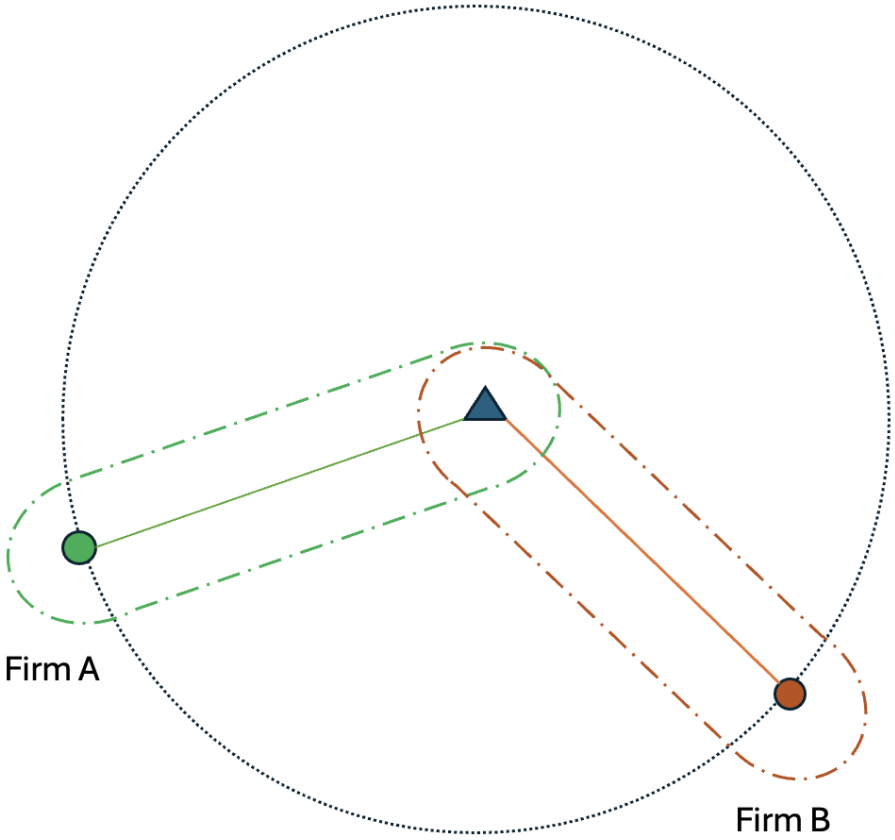
Notes:Panel A displays the main backbone node sites in Vietnam (in red) alongside the centroids of Vietnamese communes (in blue). The backbone network map is digitized from the original ITU transmission network data. Administrative boundaries correspond to Vietnam's first-level divisions (provinces). Panel B presents the original map of Vietnam's main terrestrial backbone and node sites, provided by the ITU Transmission Network Data and Vietnam Telecom National (VTN). The backbone network consists of fiber optic transmission systems connecting Hanoi and Ho Chi Minh City, as well as regional networks in the North, Northeast, Central, and Southern regions. Existing inter-provincial backbone lines are shown in dark red, with major backbone nodes—telecommunication stations—indicated by black dots outlined in red. Lines under construction appear in light red; planned lines in purple; and provincial-level backbone lines in blue. Planned EVN and Viettel fiber optic routes are shown in green and teal, respectively. Red dots mark branching points within the network.

Figure 3: Female Labor Share and Firm ICT Adoption



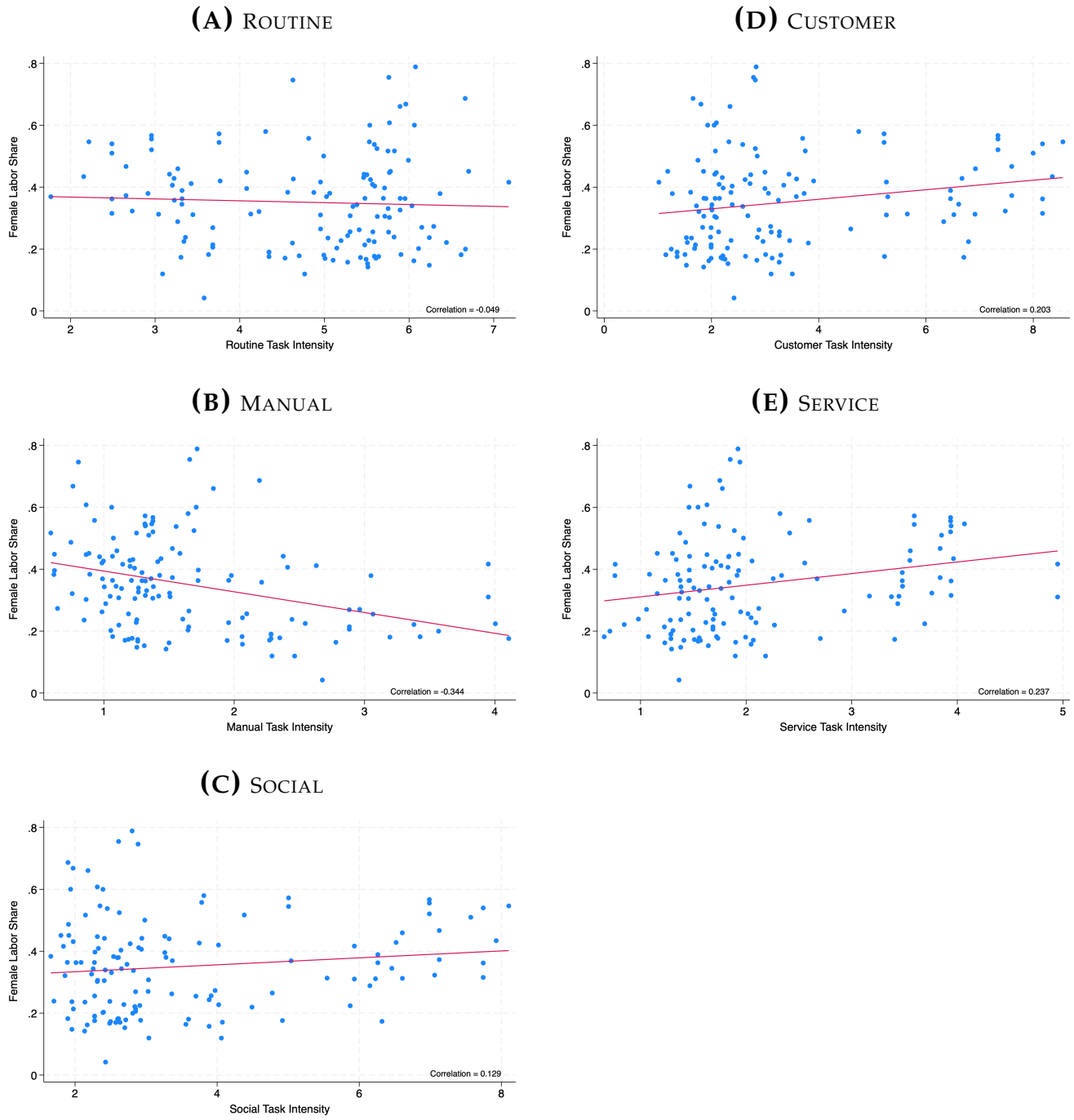
Notes: This figure presents binscatter plots of firms' female labor share against ICT adoption measures. Panel A uses the log number of Internet-connected PCs per worker, while Panel B uses the log number of LAN-connected PCs per worker.

Figure 4: The Construction of the IV



Notes: This figure illustrates the construction of the geographic instruments used in our analysis. The blue triangle represents a backbone node, while the green and red circles denote two different firms. The grey circle is centered on the backbone node, indicating equal distance to the node for both firms. The green and red shaded areas around firms A and B represent the *connectivity buffers*, which are used to construct the geographic instrumental variables.

Figure 5: Female Labor Share and Task Intesntiy



Notes: This figure shows the correlation between industry-level female labor share and task intensity. Each panel corresponds to a different type of task intensity, as indicated in the panel titles. The sectoral task intensity measures—routine, manual, social, customer, and service—are constructed by linking our firm-level data to O*NET-based task classifications, weighted by industry-level occupational shares from Indonesia's 2010 National Labor Survey.

A Appendix Tables and Figures

Table A.1: QUALITY OF BROADBAND INTERNET ACROSS SELECT COUNTRIES

Country	Speed (bits per second per Internet user)	Affordability ($ per month)
PRC	6,530	19
Indonesia	6,584	29
Cambodia	17,792	12
Viet Nam	24,374	3
Malaysia	34,119	10
Philippines	37,409	22
Thailand	64,907	19
United States	99,147	16

Notes: This table shows cross-country comparisons of broadband Internet quality, measured by average connection speed and monthly affordability. Internet speed is reported in bits per second per Internet user, and affordability reflects the average monthly cost in U.S. dollars. Data are from the World Bank's World Development Indicators and represent the latest available year.

Table A.2: VIETNAM MONTHLY WAGES FOR WAGE WORKERS BY GENDER AND EDUCATION ($)

	2007			2013		
	Male	Female	Ratio	Male	Female	Ratio
Primary	1,057	769	1.37	2,974	2,314	1.29
Some secondary	1,044	754	1.38	3,225	2,791	1.16
Secondary	1,249	1,055	1.18	3,774	3,189	1.18
Secondary tech	1,648	1,406	1.17	4,119	3,567	1.15
College	2,537	2,085	1.22	5,815	4,751	1.22

Notes: This table shows average monthly wages (in VND) for wage workers in Vietnam by gender and educational attainment for the years 2007 and 2013. The "Ratio" column reports the male-to-female wage ratio within each education category. Statistics are based on individuals aged 15–65. Source: Vietnam Labor Force Surveys.

Table A.3: VIETNAM LABOR FORCE STATISTICS BY GENDER

	2007		2013	
	Male	**Female**	**Male**	**Female**
Labor force participation rate	81.4	75.1	84.6	76.9
Employment status of those in labor force				
Unemployed	2.2	2	1.5	1.2
Employer	3.8	2.8	3.4	1.5
Own-account	46.2	56.5	42.8	45.1
Unpaid/Other	12.2	14.1	12	22.5
Wage	35.7	24.6	40.4	29.8
Education of those in labor force				
Primary	41.1	47.1	35.5	41.3
Some secondary	29.8	28.9	30.8	31.5
Secondary	13.4	11.2	17.9	12.8
Secondary tech	8.6	6.5	6.7	5
College	7.1	6.4	9.1	9.5

Notes: This table shows labor force statistics in Vietnam by gender for the years 2007 and 2013. It reports labor force participation rates, employment status, and educational attainment for individuals aged 15–65. All values are expressed as percentages. Source: Vietnam Labor Force Surveys.

Table A.4: GEOGRAPHIC CHARACTERISTICS LIST

Variable Name	**Description**	**Source**	**Resolution**
Hydrology Data			
occurrence	The frequency with which water was present.	JRC Global Surface Water	250m
max_extent	Equals one where water has ever been detected.	JRC Global Surface Water	250m
Soil Data			
bdod	Bulk density of soil.	SoilGrids	250m
cfvo	Coarse fragments in the soil.	SoilGrids	250m
sand	Percentage of sand in the soil.	SoilGrids	250m
Geographic Data			
elevation	Elevation in meters.	NASA SRTM	30m
TRI	Ruggedness Index by Nunn and Puga (2012)	Nunn and Puga (2012)	30 arc-second

Notes: This table provides the description, data source, and spatial resolution of the potential geographic, soil, and hydrological variables used in the LASSO selection procedure.

Table A.5: ICT AND FEMALE LABOR DEMAND, CONNECTIVITY BUFFER

	Baseline		Leave out Immediate Neighbors		Smaller Buffer Size (0.5 km)	
	(1)	(2)	(3)	(4)	(5)	(6)
Number of PCs with internet per worker (log)	0.356***		0.320**		0.357**	
	(0.115)		(0.145)		(0.165)	
Number of PCs with LAN per worker (log)		0.473***		0.675***		0.432**
		(0.159)		(0.226)		(0.204)
N	180,791	176,515	180,791	176,515	153,292	149,367
N Clusters	52,248	51,438	52,248	51,438	44,363	43,616
Kleibergen-Paap Wald Rank F Stat	33.879	20.146	27.674	19.580	15.493	12.708
Under Id. Test (KP Rank LM Stat)	170.486	80.640	116.279	39.539	77.647	50.874
p-Value	0.000	0.000	0.000	0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)	2.097	2.477	3.213	5.488	2.425	2.000
p-Value	0.063	0.042	0.012	0.004	0.033	0.092
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm Level Control	Yes	Yes	Yes	Yes	Yes	Yes
Node ID × Post 2007	Yes	Yes	Yes	Yes	Yes	Yes
Distance to Node × Post 2007	Yes	Yes	Yes	Yes	Yes	Yes
Control Population	Yes	Yes	Yes	Yes	Yes	Yes
Distance Urban Center × Year	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents estimates of the effect of firms' ICT adoption on the female labor share, using alternative specifications of the *connectivity buffer* used to construct the instrumental variables. Columns 1 and 2 show the baseline results, using the log number of Internet-connected PCs and LAN-connected PCs per worker, respectively, as the dependent variable. These correspond to Column 4 of Panels A and B in Table 4. Columns 3 and 4 modify the instrument by excluding a 250-meter radius around each firm from the *connectivity buffer*. Columns 5 and 6 further adjust the specification by using a narrower 0.5 km *connectivity buffer* in instrument construction. All regressions are estimated using a post-double-selection IV-Lasso procedure. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the firm level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.6: ICT AND FEMALE LABOR DEMAND, VARYING LOCATION RESTRICTIONS

	Distance: 20 km		Distance: 25 km		Distance: 15 km	
	(1)	(2)	(3)	(4)	(5)	(6)
Number of PCs with internet per worker (log)	0.356***		0.360***		0.393***	
	(0.115)		(0.121)		(0.111)	
Number of PCs with LAN per worker (log)		0.473***		0.410**		0.487***
		(0.159)		(0.170)		(0.146)
N	180,791	176,515	191,301	186,651	158,617	154,921
N Clusters	52,248	51,438	55,270	54,374	46,017	45,320
Kleibergen-Paap Wald Rank F Stat	33.879	20.146	27.509	17.865	32.344	21.687
Under Id. Test (KP Rank LM Stat)	170.486	80.640	163.574	71.549	163.338	87.101
p-Value	0.000	0.000	0.000	0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)	2.097	2.477	1.957	1.592	2.922	3.365
p-Value	0.063	0.042	0.068	0.173	0.012	0.009
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm Level Control	Yes	Yes	Yes	Yes	Yes	Yes
Node ID × Post 2007	Yes	Yes	Yes	Yes	Yes	Yes
Distance to Node × Post 2007	Yes	Yes	Yes	Yes	Yes	Yes
Control Population	Yes	Yes	Yes	Yes	Yes	Yes
Distance Urban Center × Year	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents estimates of the effect of firms' ICT adoption on the female labor share, using alternative location-based sample restrictions. Columns 1 and 2 report the baseline results, using the log number of Internet-connected PCs and LAN-connected PCs per worker, respectively, as the dependent variable. These correspond to Column 4 of Panels A and B in Table 4, where the sample is restricted to firms located within 20 km of the nearest backbone node. Columns 3 and 4 extend the sample to include firms within 25 km of the nearest backbone node, while Columns 5 and 6 restrict the sample to firms within 15 km. All regressions are estimated using a post-double-selection IV-Lasso procedure. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the firm level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.7: Firm Female Labor Share and ICT Adoption, Different Ways of clustering

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of PCs with internet per worker (log)	0.356***		0.356**		0.356***		0.356***	
	(0.115)		(0.144)		(0.106)		(0.119)	
Number of PCs with LAN per worker (log)		0.473***		0.473**		0.473***		0.473***
		(0.159)		(0.189)		(0.147)		(0.164)
N	180,791	176,515	180,791	176,515	180,791	176,515	180,791	176,515
N Clusters	52,248	51,438	3,704	3,662	15,653	15,399	15,653	15,399
Kleibergen-Paap Wald Rank F Stat	33.879	20.146	16.345	9.617	32.592	19.086	27.474	16.913
Under Id. Test (KP Rank LM Stat)	170.486	80.640	80.815	44.450	193.726	97.394	124.417	63.300
p-Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)	2.097	2.477	1.430	1.689	2.424	2.854	1.945	2.298
p-Value	0.063	0.042	0.210	0.150	0.033	0.022	0.084	0.057
Cluster Level:								
Firm	Yes	Yes						
Commune			Yes	Yes				
Commune×Year					Yes	Yes		
Commune×Year; Firm							Yes	Yes

Notes: This table presents estimates of the effect of firms' ICT adoption on the female labor share, using alternative clustering methods for standard errors. Columns 1 and 2 report the baseline results, where the dependent variables are the log number of Internet-connected PCs and LAN-connected PCs per worker, respectively. These correspond to Column 4 of Panels A and B in Table 4, with standard errors clustered at the firm level. Columns 3–8 apply alternative clustering approaches, as specified in the bottom panel of the table. All regressions are estimated using a post-double-selection IV-Lasso procedure. */**/*** denotes significant at the 10% / 5% / 1% levels.

B Data Appendix: Task Intensity

To examine the potential mechanisms, we merge occupational-level task intensity measures with Vietnam's household and firm survey data. These task measures are originally defined using U.S. Census occupational codes and data from the U.S. Department of Labor's Occupational Information Network (O*NET), which provides detailed information on the nature of work across occupations. Following established frameworks in the literature (Autor et al., 2003; Deming, 2017), these measures capture the extent to which each occupation relies on routine, social, customer-based, and service-oriented tasks, etc. For the household-level analysis, we directly merge O*NET task intensity measures using the occupation codes reported for each worker. For example, sewing machine operators in garment factories and construction workers are high in manual task intensity; data entry clerks and call center agents are high in routine and communication intensity; while sales agents require high levels of social task intensity.

For the firm-level data, we lack detailed occupational breakdowns or firm-level employment by occupation. To address this, we construct industry-level average task intensity measures by combining occupational task intensity scores from the U.S. O*NET database with the occupational composition of each industry. To avoid issues related to factor intensity reversal—i.e., the possibility that the relative task intensities across occupations differ between countries—we use the 2010 Indonesian National Labor Force Survey to derive occupation shares within industries. Indonesia provides a reasonable benchmark given its similar stage of economic development and comparable industrial structure to Vietnam.

Appendix Tables B.8 and B.9 list the five sectors with the highest and lowest task intensities across four categories—manual, social, customer-facing, and service-oriented—based on the Indonesian labor force survey. In our data, manual task–intensive industries include sectors such as transport via pipeline, land and air transport, and services to buildings and landscape activities, which are characterized by high physical demands and low automation. Industries with high social task intensity include retail sectors such as non-specialized stores and specialized food and beverage stores, which require frequent interpersonal interactions. In contrast, manufacturing industries, such as the manufacture of paper and paper products or pharmaceuticals, generally rank among the lowest in social task intensity, consistent with their reliance on standardized production processes and limited interaction with customers or coworkers. High customer-facing task intensity is concentrated in retail and service delivery sectors, including non-specialized retail and steam and air conditioning supply, where frequent client engagement and responsiveness are critical. Conversely, service-oriented task intensity is highest in logistics and infrastructure services such as freight air transport and steam supply, while technical and industrial sectors such as engineering services and tobacco manufacturing rank among the lowest.

Table B.8: SECTORS BY DIFFERENT TASK INTENSITY, PART I

Rank	Sectors (Code)	Task Intensity
Routine		
1	Transport via pipeline (493)	4.11
2	Other Land Transport (492)	4.00
3	Passenger Air Transport (511)	3.95
4	Freight Air Transport (512)	3.95
5	Services to Buildings and Landscape Activities (81)	3.56
...		
145	Music Publishing (592)	0.62
146	Reproduction of Recorded Media (182)	0.62
147	Manufacture of Pharmaceuticals, Medicinal Chemical and Botanical Products (210)	0.59
148	Manufacture of Paper and Paper Products (17)	0.26
149	Activities of Other Membership Organizations (949)	0.26
Manual		
1	Transport via pipeline (493)	4.11
2	Other Land Transport (492)	4.00
3	Passenger Air Transport (511)	3.95
4	Freight Air Transport (512)	3.95
5	Services to Buildings and Landscape Activities (81)	3.56
...		
145	Music Publishing (592)	0.62
146	Reproduction of Recorded Media (182)	0.62
147	Manufacture of Pharmaceuticals, Medicinal Chemical and Botanical Products (210)	0.59
148	Manufacture of Paper and Paper Products (17)	0.26
149	Activities of Other Membership Organizations (949)	0.26
Social		
1	Retail Sale of Food, Beverages and Tobacco in Specialized Stores (472)	1.44
2	Retail Sale in Non-Specialized Stores (471)	1.46
3	Retail Sale of Other Household Equipment in Specialized Stores (475)	1.32
4	Retail Sale of Information and Communications Equipment in Specialized Stores (474)	1.32
5	Retail Sale of Cultural and Recreation Goods in Specialized Stores (476)	1.32
...		
145	Services to Buildings and Landscape Activities (81)	3.56
146	Scientific Research and Development (72)	3.34
147	Manufacture of Man-Made Fibres (203)	0.89
148	Activities of Other Membership Organizations (949)	0.26
149	Manufacture of Paper and Paper Product (17)	0.26

Notes: This table lists the top five and bottom five industries (at the 3-digit level) by task intensity, based on O*NET occupational task measures. Task intensities are aggregated to the industry level by averaging occupational measures, weighted by industry-level occupational shares from Indonesia's 2010 National Labor Survey. Each panel presents rankings for a specific task type—routine, manual, social, customer, and service.

Table B.9: SECTORS BY DIFFERENT TASK INTENSITY, PART II

Rank	Sectors (Code)	Task Intensity
Customer		
1	Retail Sale in Non-Specialized Stores (471)	8.55
2	Retail Sale of Food, Beverages and Tobacco in Specialized Stores (472)	8.35
3	Steam and Air Conditioning Supply (353)	8.18
4	Retail Sale of Cultural and Recreation Goods in Specialized Stores (476)	8.17
5	Retail Sale of Information and Communications Equipment in Specialized Stores (474)	8.17
...		
145	Services to Buildings and Landscape Activities (81)	1.24
146	Manufacture of Man-Made Fibres (203)	1.18
147	Scientific Research and Development (72)	1.15
148	Materials Recovery (383)	1.13
149	Manufacture of Tobacco Products (120)	1.02
Service		
1	Freight Air Transport (512)	4.95
2	Air Transport (511)	4.95
3	Steam and Air Conditioning Supply (353)	4.54
4	Retail Sale in Non-Specialized Stores(471)	4.07
5	Retail Sale of Food, Beverages and Tobacco in Specialized Stores (472)	3.76
...		
145	Manufacture of Structural Metal Products (251)	1.67
146	Activities of Extraterritorial Organizations and Bodies (99)	1.06
147	Sawmilling and Planing of Wood (161)	0.97
148	Architectural and Engineering Activities; Technical Testing and Analysis (71)	0.84
149	Manufacture of Tobacco Products (120)	0.76

Notes: This table lists the top five and bottom five industries (at the 3-digit level) by task intensity, based on O*NET occupational task measures. Task intensities are aggregated to the industry level by averaging occupational measures, weighted by industry-level occupational shares from Indonesia's 2010 National Labor Survey. Each panel presents rankings for a specific task type—routine, manual, social, customer, and service.