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 2026

Abstract

We study the impact of Information and Communication Technologies (ICT) adoption on female employment outcomes in Vietnam using rich firm- and individual-level data. Our identification exploits exogenous variation in ICT access driven by geographic connectivity conditions and a nationwide network upgrade in 2007. We find that ICT adoption increases the female share of employment within firms, primarily through within-firm reallocation in the formal sector. We further provide evidence consistent with a task-reallocation channel: ICT is associated with a shift in labor demand away from routine and manual tasks toward social, customer-facing, and service-oriented activities, where women are more likely to have 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 Agricultural, Environmental, and Development Economics (AEDE) Faculty Brownbag, the Southern Economic Association (SEA), the Chinese Economists Society (CES), and Hong Kong University 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 (Piketty, 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). A growing literature documents how ICT has profoundly transformed labor market outcomes, including its effects on women's employment, earnings, and occupational outcomes (Kuhn and Mansour, 2014; Suri and Jack, 2016; Viollaz and Winkler, 2022; Chiplunkar and Goldberg, 2022; Bahia et al., 2024; Kusumawardhani et al., 2023; Ho et al., 2024; Chung et al., 2025). However, relatively little is known about the labor demand side of this transformation, specifically, how firm-level ICT adoption reshapes the gender composition of employment within firms and the mechanisms through which these changes occur, particularly in developing-country contexts.

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 study this question in the context of a large developing country, Vietnam. Vietnam provides a particularly compelling setting, as it has experienced both a relatively recent expansion of ICT infrastructure and rapid economic development. Vietnam also offers rich data that allow us to study the effects of ICT expansion, most importantly the country's annual enterprise census, a panel dataset that provides detailed information on firm-level ICT adoption and a rich set of production variables, including female employment.

Identifying the causal impact of ICT expansion on labor market outcomes is challenging, as ICT infrastructure is not randomly placed, and firms may endogenously sort into areas with

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

better access. To address this, our paper leverages plausibly exogenous variation in firms' ICT access. First, we exploit temporal variation from the nationwide expansion of ICT infrastructure in 2007, when Vietnam implemented a major upgrade to its domestic backbone and submarine cable systems, substantially improving internet capacity and quality. Second, we exploit geographic variation in the effective cost of connecting to the ICT backbone. Not all firms benefited equally from the nationwide ICT expansion, as treatment intensity depends on the geographic conditions of connecting to the nearest backbone node, which is the core infrastructure that transmits high-capacity internet across regions. To capture this heterogeneity, we construct a measure of geographic connectivity conditions based on terrain, soil characteristics, and hydrological features along the path from each firm to its closest backbone node. Features such as water bodies, unstable soil, and rugged terrain increase the difficulty of laying and maintaining fiber connections, lengthen effective connection routes, and reduce the quality of internet access. As a result, firms facing more difficult geographic conditions experience smaller improvements in ICT access following the nationwide expansion. We interact these geographic connectivity measures with the post-2007 expansion to construct instruments for firm-level ICT adoption. The intuition behind our instrument is similar to a generalized difference-in-differences strategy (Nunn and Qian, 2011): we compare outcomes across firms located in areas with different levels of geographic connectivity conditions to the ICT backbone before and after the nationwide expansion.

Our identification strategy relies on two key assumptions. First, relevance: firms and households facing more favorable geographic connectivity conditions experience larger improvements in ICT access following the 2007 expansion. Second, the exclusion restriction: the instrument affects female labor share only through its impact on ICT adoption. In other words, there are no concurrent shocks around 2007 that are systematically correlated with geographic connectivity conditions to backbone nodes and that independently affect labor market outcomes. Importantly, identification relies on the exogeneity of the interaction between the nationwide ICT expansion and geographic connectivity conditions, rather than requiring each component to be exogenous on its own. To address potential concerns regarding the exclusion restriction,

including concurrent urbanization trends and improvements in other infrastructure, we allow for regional trends and flexible differential trends by proximity to other infrastructure, and additionally control for local population growth and urbanization.

We find that greater firm-level ICT adoption significantly increases the share of female workers in firm employment. Our first-stage results show that firms' post-2007 ICT adoption depends critically on local terrain, hydrological, and soil conditions along the path to the nearest ICT backbone node. Notice that because many geographic conditions may influence connectivity, we follow Belloni et al. (2012) and implement a LASSO procedure to select the relevant instruments, thereby mitigating concerns about overfitting and weak instruments. In the second stage, we estimate a robust and statistically significant positive effect of ICT adoption on female employment share. Specifically, our baseline estimates imply that a one-standard-deviation increase in internet-connected PCs per worker raises the female labor share by about 6.6 percentage points, and a comparable increase in LAN-connected PCs per worker by about 11.5 percentage points. We also find evidence that these effects persist into the medium term. Importantly, the results appear to be driven primarily by compositional shifts across gender, as we find no statistically significant changes in firms' total employment or firm entry. Pre-trend estimates are statistically insignificant over an extended 2001–2006 pre-period, and our results are robust to alternative instruments and specification choices.

We complement the firm-level analysis with data from the Vietnam Household Living Standards Survey (VHLSS) to provide additional insights into labor market outcomes, including those in the informal economy. To mirror the treatment variation in the firm-level analysis, we assign households the same geographically determined ICT treatment intensity as firms located in the same area. We then implement a triple-DiD framework using the household data. The results suggest that the gender effects observed at the firm level are primarily driven by labor reallocation within the formal sector. In particular, ICT expansion increases women's likelihood of working in non-agricultural activities relative to men, while raising wages overall with no significant gender differences. We find no significant effects on transitions from informal to formal employment or on overall employment rates.

To better understand the mechanisms through which ICT adoption increases firms' female labor shares, we examine labor reallocation across tasks requiring different skills. This analysis is motivated by prior research showing that ICT tends to substitute for routine and manual labor while complementing non-routine, social, and service-oriented tasks (Autor et al., 2003; Autor and Dorn, 2013; Deming, 2017; Acemoglu and Restrepo, 2020; Hötte et al., 2023), which are activities in which women are more likely to have a comparative advantage (Rueckert and Naybar, 2008; Schulte-Rüther et al., 2008; Cortes and Pan, 2018). Consistent with this mechanism, we find that the positive effect of ICT adoption on female labor share is stronger in industries with lower routine and manual task intensity and higher service, social, and customer-facing task intensity. At the individual level, women are more likely than men to transition into occupations emphasizing social, customer-facing, and service-oriented tasks, and less likely to remain in routine roles, following ICT expansion.

This paper contributes to the growing literature on the labor-market effects of ICT (Autor et al., 1998; Crandall et al., 2007; Forman et al., 2012; Michaels et al., 2014; Akerman et al., 2015), including recent work on developing countries (e.g., Kuhn and Mansour, 2014; Hjort and Poulsen, 2019; Viollaz and Winkler, 2022; Azam et al., 2024; Bahia et al., 2024). As reviewed by Hjort and Tian (2025), ICT can generate substantial employment and productivity gains, although its effects and underlying mechanisms vary across settings. We contribute by studying gender-specific employment outcomes in Vietnam, one of the fastest-growing developing economies, using firm-level data. Methodologically, we develop an identification strategy that uses geographic characteristics along the path to ICT backbone nodes as proxies for the engineering costs of connection.[2] This strategy complements conventional distance-to-infrastructure instruments (e.g., Miner, 2015; Hjort and Poulsen, 2019; Geraci et al., 2022) by incorporating geographic conditions that more directly capture connection costs and are less closely tied to the endogenous placement of infrastructure.

Second, our study contributes to the literature on technological change, task composition, and gender inequality (Autor et al., 2003; Black and Spitz-Oener, 2010; Beaudry and Lewis,

[2]This approach is conceptually related to Ellis et al. (2025), who use terrain variation and the spatial distribution of mobile phone towers to predict mobile signal strength.

2014; Greenwood et al., 2016; Olivetti and Petrongolo, 2016; Ngai and Petrongolo, 2017; Aksoy et al., 2021; Kwon, 2022; Cortés et al., 2024; Vidart, 2024).[3] ICT tends to substitute for routine and manual tasks while complementing work involving information processing, interpersonal interaction, and service provision, partly by enabling more flexible workplace organization and task design (Bresnahan et al., 2002; Autor et al., 2003; Bartel et al., 2007; Autor and Dorn, 2013; Deming, 2017; Acemoglu and Restrepo, 2020; Hötte et al., 2023). At the same time, women tend to have a comparative advantage in cognitive, social, customer-facing, and service-oriented work relative to physically intensive manual tasks (Welch, 2000; Rendall, 2018; Blau and Kahn, 2017; Cortes and Pan, 2018). ICT may therefore increase the relative demand for female labor by shifting production toward tasks in which women are comparatively more productive, even without expanding firms' total employment. We test this mechanism directly by examining whether the effects of ICT vary with industries' task intensity and whether ICT exposure shifts women's occupational allocation, relative to men's, toward social, customer-facing, and service-oriented tasks.

Our paper is closely related to Juhn et al. (2014), who show that trade-induced technological change improved female labor-market outcomes in Mexico by reducing the importance of physical tasks, and to Chiplunkar and Goldberg (2022), who find that the rollout of 3G mobile internet increased female labor-force participation in developing countries. We differ by studying broadband infrastructure, firm-level employment composition, and the task-based mechanisms underlying the gender effects of ICT. We also connect to research showing that interpersonal and non-cognitive skills are increasingly rewarded in modern labor markets (Heckman et al., 2006; Borghans et al., 2008; Lindqvist and Vestman, 2011; Deming, 2017). Our findings suggest that ICT-driven shifts toward social, customer-facing, and service-oriented tasks are one channel through which broadband expansion increases female employment.

The remainder of the paper is organized as follows. Section 2 provides the institutional background. Section 3 describes the data. Section 4 outlines the identification strategy and the construction of the instrumental variable. Section 5 presents the firm-level results. Section 6 extends the analysis using household data. Section 7 examines the underlying mechanisms.

[3]See Goldin (2006) for a broader discussion of the forces shaping female labor-market outcomes.

Section 8 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* (renovation) 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.[4] 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 1 displays the early terrestrial backbone and node sites, based on data sourced from the VTN. 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 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 3.6 gigabits per second (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 10 Gbps capacity upgrade. This was the

[4]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.

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) 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 in 2007 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 A.1, international bandwidth surged from 3,615 megabits per second (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 kilobits per second (Kbps) per person, and entry-level broadband plans were available for as little as $3 per month, making Internet access broadly affordable.

Importantly, the 2007 ICT expansion in Vietnam is best understood primarily as an upgrade of existing backbone infrastructure rather than the construction of entirely new network coverage. The core terrestrial backbone network had largely been established prior to 2007, and the reform mainly involved substantial increases in transmission capacity, both through upgrades to the SEA-ME-WE 3 submarine cable and improvements to the domestic backbone system. While some complementary investments, such as extending local connections and improving last-mile access, accompanied this process, these changes were incremental relative to the existing backbone footprint. As a result, the expansion primarily improved the quality, speed, and reliability of connectivity rather than expanding access to entirely new locations.

At the same time, local and last-mile access continued to expand, although commune-level public access still relied largely on narrowband connections. VNPT expanded access through its network of Communal Post and Cultural Points, beginning in 2003 to supply mainly dial-up Internet to around 3,000 of them. Meanwhile, the Universal Service Fund, established in 2004, supported a program that aimed to provide public telephone access points in all communes and public Internet access points in 70 percent of communes by 2010. (Lee, 2011; Tuan, 2011). The 2007 backbone upgrade therefore complemented, rather than replaced, this ongoing last-mile build-out: greater backbone capacity encouraged firms to establish or upgrade local connections in subsequent years. Our instrument exploits the geographically determined cost of making these last-mile connections.

Gender Differences in the Labor Market Vietnam's labor market exhibits relatively small gender gaps in labor force participation compared to other lower-middle-income countries, yet substantial disparities persist in employment type and earnings. In 2012, 82% of men and 73% of women aged 15 and older participated in the labor force (World Bank, 2026). Although gender wage gaps narrowed between 2004 and 2012, particularly among less-educated workers, men continued to earn approximately 12–44% more than women on average (Appendix Table A.2).

As in many developing countries, informal employment and unpaid work remain prevalent. Women are also less likely to hold wage employment in Vietnam: as shown in Appendix Table A.3, in 2012 only 30% of women were wage employees, compared to 47% of men. Men are also more likely to hold primary jobs in the formal sector, with rates of 25% for men and 22% for women.

3 Data

Firm-level Data The firm-level data used in our main analysis come from the Annual Enterprise Census of Vietnam, conducted by the General Statistics Office (GSO), covering the period from 2004 to 2009; in some analyses, we extend the sample to 2011. These annual censuses in-

clude 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 distinctive feature of the census is its detailed information on firms' use of information and communication technologies, including personal computers, Internet access, and local area networks (LANs).[5] Our two primary measures are the number of Internet-connected PCs per worker and the number of LAN-connected PCs per worker. Appendix Table A.4 reports their evolution from 2004 to 2009. ICT information was not collected in 2006. Both measures increased substantially over the sample period: Internet-connected PCs per worker rose from 0.05 in 2004 to 0.15 in 2009, while LAN-connected PCs per worker increased from 0.05 to 0.12.

We also examine the raw correlation between ICT adoption and firms' female labor share. Figure 2 presents binned scatter plots illustrating these relationships. To show *when* this association emerges, we split the sample by period: the left panel plots firms' pre-expansion (2004–2006) female labor share against their eventual (post-2007) Internet adoption, and the right panel plots their post-expansion (2007–2009) female labor share against the same measure. A positive association is already present before the expansion—indicating some sorting of high-female-share firms into eventual ICT adoption—but the slope is roughly twice as steep after 2007 (fitted slope of 0.053 pre-2007 versus 0.100 post-2007). This steepening, with the association strengthening precisely when ICT becomes available, is consistent with a causal effect of ICT adoption rather than with pre-existing differences alone. For this descriptive figure we plot ICT adoption on the $\log(1+x)$ scale for readability; our main analysis, however, uses ICT adoption in levels. Appendix Figure A.2 repeats this exercise using the LAN measure and shows the same steepening (fitted slope 0.057 pre-2007 versus 0.101 post-2007).

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

[5] A local area network connects computers and other devices within a limited area, such as an office or building, allowing users to share data, applications, and resources.

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.[6] Panel A of Figure 1 displays the commune centroids of the firms in blue, as reported in the Vietnam Enterprise Surveys.

To ensure comparability in connectivity environments, we restrict the estimation sample to firms located within 20 km of an ICT backbone node, which we introduce later in this section. This restriction retains approximately 88% of surveyed firms but disproportionately excludes remote rural communes. Our estimates should therefore be interpreted as applying primarily to firms in Vietnam's more connected and economically active areas. Table 1 reports descriptive statistics for the resulting sample. Women account for about one-third of firm employment, with a mean female labor share of 0.331. Firms have, on average, 0.112 Internet-connected PCs and 0.101 LAN-connected PCs per worker, with substantial dispersion in both measures. Mean log employment is 3.805, corresponding to approximately 45 workers. Foreign-owned and state-owned firms account for 9.3% and 10.0% of the sample, respectively, with domestically owned private firms serving as the omitted category.

Individual-level Data To complement our firm-level analysis and shed light on broader labor market outcomes, particularly in the informal economy, we also use individual-level data from the Vietnam Household Living Standards Survey (VHLSS), which spans 2002–2012 and is conducted biennially. The VHLSS is a nationally representative, multi-topic household survey administered by Vietnam's GSO to monitor living conditions and socioeconomic outcomes. It covers workers across all industries and reports detailed demographic characteristics, including gender, age, education, and marital status. The labor module provides rich information on employment outcomes, such as employment status, monthly workdays, wages, sector of employment, self-employment, job formality, secondary employment, and engagement in house-

[6] 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.

work. Throughout the individual-level analysis, we classify a worker as formally employed if the job is in a private registered firm, a state-owned enterprise, a public institution, or a foreign-invested firm; all other employment is classified as informal.[7]

Importantly, the survey includes commune (*xa*)-level geographic identifiers, which allow us to geocode households and construct ICT exposure measures consistent with those derived from the firm census. We describe this procedure in greater detail in Section 6. Although the survey reports household access to broadband internet, adoption rates are very low—0.3% in 2002 and 4% in 2008. Since our focus is on the effects of broadband expansion on labor demand (firm-level ICT adoption), rather than household-level adoption per se, we do not treat this measure as capturing the relevant policy variation in our setting.

Terrestrial Backbone Nodes Location We obtain historical maps of Vietnam's terrestrial backbone network from the International Telecommunication Union (ITU) and Vietnam Telecom National (VTN). Panel B of Figure 1 displays the original map, which depicts the fiber-optic transmission system linking 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 as solid red lines, while 115 ICT backbone nodes (telecommunication stations) are marked by black dots. Lines under construction are indicated by dashed red lines.

Since backbone nodes 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 ICT adoption. Panel A of Figure 1 illustrates the spatial distribution of Vietnam's terrestrial backbone nodes (red dots). 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. Also, the backbone network is broadly distributed along the country's north–south axis. The majority of communes are located in close proximity to at least one backbone node.

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

Table 2 presents summary statistics on the spatial distribution of firms and individuals relative to backbone nodes. The Annual Enterprise Census (2004–2009) includes 117,333 unique firms, spanning 64 provinces, 665 districts, and 6,997 communes. In contrast, the individual-level data have more limited geographic coverage, encompassing 3,496 communes. Approximately 88% of firms and 58% of individuals are located within 20 km of the nearest terrestrial backbone node.

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 satellite-based data to capture localized geographic features, including: (1) hydrological conditions from the JRC Global Surface Water Mapping Layers (Pekel et al., 2016); (2) soil characteristics from SoilGrids (Hengl et al., 2017); (3) elevation from the NASA SRTM Digital Elevation Model (Farr et al., 2007); and (4) terrain ruggedness from Nunn and Puga (2012). All geographic variables are time-invariant. Soil characteristics and ruggedness are inherently fixed, while the hydrological measures are constructed as long-run aggregates from historical time-series satellite imagery.

Appendix Table A.5 summarizes the variables that may affect the engineering feasibility of ICT infrastructure deployment. For instance, higher elevation increases 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.

Other Data We use the Global Human Settlement Layer population dataset (GHS-POP) (SCHIAVINA et al., 2025) to measure population. GHS-POP is a globally harmonized dataset produced by the European Commission's Joint Research Centre that provides high-resolution gridded population estimates by combining census data with satellite-derived built-up area information. Because population data are available only for 2005 and 2010, we compute the annual population growth rate for each spatial unit and interpolate annual population levels for the intervening years. In addition, we use the 2019 GHS Urban Centre Database to obtain the ge-

ographic coordinates of urban centers in Vietnam. Data on roads, railways, and rivers (circa 2007/2008) are obtained from the NYU Spatial Data Repository (SDR) (International Steering Committee for Global Mapping, 2007b,a, 2008). For our mechanism analysis, we use task measures from the U.S. Department of Labor's Occupational Information Network (O*NET), which provides detailed information on the nature of work across occupations (U.S. Department of Labor, Employment and Training Administration, 2025). Further details on the O*NET data are provided in Section 7 and Appendix Section B.

4 Empirical Strategy

4.1 OLS Specification

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

$$FemaleS_{it} = \alpha_0 + \beta_1 ICT_{it} + \gamma_i + \theta_{nt} + \mathbf{X}'_{it}\beta_2 + \mathbf{Z}'_{it}\beta_3 + \epsilon_{it}, \tag{1}$$

where $FemaleS_{it}$ denotes the share of female workers in total employment for firm i in year t, and ICT_{it} measures firm-level ICT adoption. We consider two alternative measures: (i) the number of internet-connected PCs per worker and (ii) the number of LAN-connected PCs per worker. The specification includes firm fixed effects, γ_i, which absorb time-invariant firm characteristics, and (nearest) backbone node-by-year fixed effects, θ_{nt}, which absorb node-specific time-varying shocks. Because backbone nodes are closely associated with major regional centers (see Figure 1), these fixed effects also capture broader region-specific time trends. The vector $\mathbf{X}_{it}$ includes time-varying firm-level controls, such as indicators for foreign and state ownership, as well as total sales, fixed assets, and total employment. In addition, $\mathbf{Z}_{it}$ includes time-varying locational controls and baseline location characteristics interacted with year fixed effects, allowing for flexible differential trends across locations with different baseline characteristics that may be correlated with both ICT adoption and female employment. We provide further details on the construction of $\mathbf{Z}_{it}$ in the following section.

4.2 Instrumental Variable Approach

Sources of Exogenous Variation for the Instrument Even after controlling for firm fixed effects, flexible regional time trends, and time-varying firm-level characteristics, our estimates may still be subject to endogeneity concerns. First, the placement of ICT infrastructure is unlikely to be random. Backbone nodes are typically located based on strategic policy considerations and may correlate with local socio-economic conditions that also affect female employment outcomes, thereby generating potential omitted-variable bias. Second, reverse causality may arise if firms with a higher female employment share undergo organizational or technological changes that increase their demand for ICT adoption.

To address these concerns, we exploit both temporal and spatial sources of variation to construct an instrumental variable for firm-level ICT adoption. On the temporal dimension, we leverage the nationwide ICT capacity expansion in 2007, when Vietnam's domestic terrestrial backbone and submarine cable systems underwent substantial upgrades, sharply increasing Internet transmission speeds, as described in Section 2. Because this reform coincided with other transformative economic developments, including Vietnam's accession to the WTO and broader processes of globalization and urbanization, we further isolate ICT-specific effects by exploiting cross-sectional heterogeneity in firms' connectivity costs to the backbone network.

On the spatial dimension, firms differ in the ease of accessing the upgraded backbone network. Although capacity expanded nationwide, the extent to which firms benefited depended on geographic conditions affecting their connectivity to the nearest backbone node. These conditions influence the engineering difficulty of expanding and maintaining last-mile infrastructure and, more broadly, the quality and reliability of access. Engineering evidence indicates that terrain, soil stability, and hydrological barriers play a central role in determining the cost of last-mile infrastructure deployment (Biggs and Mahony, 2004). For example, fiber installation is substantially more expensive in rocky terrain than in soft soil, and additional obstacles such as waterways or unstable ground further increase construction difficulty (Fiber Broadband Association and Cartesian, 2024). As a result, firms located in areas with flatter terrain, more stable soils, and fewer hydrological barriers along the path to the ICT backbone experienced

larger effective improvements in access following the 2007 expansion.

Motivated by these engineering considerations, we proxy the cost of connecting each firm to the backbone using terrain, soil, and hydrological characteristics *along the path between the firm and its nearest backbone node*. We interact these predetermined characteristics with an indicator for the post-2007 ICT expansion. The intuition is that, following the nationwide capacity upgrade, firms facing more favorable connection conditions experienced larger improvements in effective ICT access and therefore stronger increases in ICT adoption. Because the backbone network was largely in place before 2007, the reform primarily increased the capacity of existing infrastructure rather than extending backbone coverage. Identification therefore comes from differential post-upgrade changes associated with predetermined geographic conditions, rather than from either geography or the timing of the expansion alone. The instrument is constructed exclusively from backbone nodes that were operational before the 2007 upgrade; links that were planned or under construction at the time, as indicated in Figure 1, play no role in its construction.

The nearest-node assignment should be interpreted as a proxy for the engineering cost of extending high-quality connectivity rather than as a literal representation of network routing. VNPT's national backbone connected its provincial operating companies across the country, so actual routing may also have reflected the organization of provincial networks and other engineering considerations (Tuan, 2011). However, the terrain, soil, and hydrological characteristics along the firm–node path capture variation in the difficulty of extending local connectivity. Only about 8% of firms are assigned to a nearest node outside their province, and Section 5.6 shows that excluding these firms leaves the results unchanged.

Instrumental Variable Construction To operationalize this, 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-kilometer (3 km) buffer around this path. We term this buffer the *connectivity buffer* for future reference. Within this buffer, we compute summary statistics (mean, standard deviation, minimum, and maximum) of geographic characteristics, including hydrological features, soil properties, ruggedness, and elevation (see Table A.5 for

the full list).

Because we include backbone node-by-year fixed effects, θ_{nt}, identification is based on comparing changes in ICT adoption across firms within the same backbone-node catchment area. In addition, we flexibly control for differential time trends by proximity to backbone infrastructure by interacting the log distance to the nearest node, $\log(DistNode_i)$, with year dummies. While prior studies use distance to backbone networks as a proxy for connection costs (Miner, 2015; Geraci et al., 2022), motivated by the notion that effective connectivity declines with distance (Banerji and Chowdhury, 2013), backbone nodes are typically strategically located near cities and population centers, as shown in Panel A of Figure 1. As a result, distance may capture not only connectivity-related frictions but also broader factors such as agglomeration forces and historical development, which can independently influence local economic conditions and firm behavior. Although firm fixed effects absorb time-invariant spatial differences, distance may still correlate with differential post-2007 trends; interacting distance with year fixed effects ensures that our identification does not mechanically rely on proximity to backbone nodes.

Figure A.3 provides a schematic illustration of the geographic variation used in our identification strategy. The blue triangle denotes a backbone node. Firm A (green circle) and Firm B (orange circle) both identify this node as their nearest backbone connection. The large black dotted circle, centered at the node, illustrates that the two firms are located at a similar distance from the backbone node for expositional purposes.

The solid green and orange lines represent the straight-line paths connecting each firm to the node. Around each path, we construct a *connectivity buffer* (shown as dash-dotted corridors in green and orange), within which geographic characteristics such as terrain, soil, and hydrological features are aggregated. Although the two firms are equidistant from the same backbone node, the geographic conditions within their respective connectivity buffers differ. These differences imply heterogeneous connection costs and generate differential increases in ICT adoption following the post-2007 broadband expansion. We use this differential exposure to identify the effect of ICT adoption on female employment shares.

The exclusion restriction requires that, conditional on firm fixed effects, node-by-year fixed

effects, and flexible controls for proximity to backbone, the interaction between geographic connectivity conditions and the post-2007 indicator affects changes in a firm's female employment share only through its impact on ICT adoption. Since firm fixed effects absorb all time-invariant channels through which local geography could influence employment composition, and node-by-year fixed effects capture common local shocks, any remaining violation would have to arise from *time-varying* local developments that coincide with the 2007 expansion.

A remaining concern is that geographic characteristics may influence urbanization, population growth, or the placement of other infrastructure, generating differential post-2007 trends correlated with our instrument. To address this, we control for time-varying population density within the *connectivity buffer*, $\log(pop)_{it}$, and allow for flexible trends by proximity to urban centers by interacting $\log(\text{DistCBD})_i$ with year fixed effects. We further mitigate concerns about correlated infrastructure expansion by allowing for flexible time trends by proximity to other major infrastructure networks, implemented by interacting $\log(\text{DistInfras}_i)$ with year fixed effects, where $\log(\text{DistInfras}_i)$ includes log distances to railways, major roads, and rivers.

A related concern is that Vietnam's 2007 WTO accession, which coincided with the ICT expansion, may have placed more trade-exposed localities on different growth trajectories. In addition, internationally engaged and foreign-invested firms may employ larger shares of women. To address these concerns, all specifications control for communes' pre-2007 shares of foreign-invested and trade-participating firms, $\overline{\text{FDI}}_c$ and $\overline{\text{Trade}}_c$, each interacted with year fixed effects ($\overline{\text{FDI}}_c \times \theta_t$ and $\overline{\text{Trade}}_c \times \theta_t$). These interactions allow communes with different initial exposure to foreign investment and trade to follow flexible year-specific trajectories.

Following Abadie et al. (2023), who emphasize that standard errors should be clustered at the level where treatment assignment or key identifying variation occurs, we cluster standard errors at the commune level, where the geographic variables proxying ICT connectivity costs are defined and where the key variation in our analysis arises. To ensure comparability in connectivity environments, we restrict the sample to firms located within 20 km of a backbone node.

5 Main Results

5.1 First Stage

Given the large number of geographic characteristics that may affect firms' access to ICT infrastructure, including all candidate instruments in the first stage could lead to overfitting and weak-instrument concerns (Bekker, 1994; Hansen et al., 2008). To address this issue, we employ a Least Absolute Shrinkage and Selection Operator (LASSO) procedure to identify the optimal set of instrumental variables, following Belloni et al. (2012). Intuitively, LASSO solves a penalized first-stage regression that retains only the subset of instruments jointly predictive of the endogenous regressor, shrinking the coefficients on uninformative variables to zero. Following the notation of Liang et al. (2025), we refer to this step as the "zero-stage" LASSO regression.[8]

Our candidate instruments consist of interaction terms between geographic characteristics of the *connectivity buffer* (geographic connectivity conditions), denoted Geo_i, and the post-2007 indicator, $Post07_t$. In the zero-stage, we apply LASSO to select the relevant subset of interactions $Geo_i \times Post07_t$, denoted $selectGeo_i \times Post07_t$.

After applying the LASSO procedure in the zero stage, we estimate the first-stage regression using the selected instrumental variables $selectGeo_i \times Post07_t$. We run the following specification:

$$ICT_{it} = \alpha_0 + \beta_1 selectGeo_i \times Post07_t + \gamma_i + \theta_{nt} + \mathbf{X}'_{it}\beta_2 + \mathbf{Z}'_{it}\beta_3 + \epsilon_{it}, \tag{2}$$

where all other variables besides the selected instrumental variables $selectGeo_i \times Post07_t$ are defined as in equation (1).

Table 3 presents the estimation results. Out of 28 potential geographic characteristics, only 2–5 are selected by the LASSO procedure. To facilitate interpretation, we rescale all geographic variables by dividing them by 10, so that coefficients correspond to economically meaningful changes.

[8]This procedure has been widely applied in empirical work, including Beraja et al. (2023), Civelli et al. (2023), and Liang et al. (2025).

In column 1, where the dependent variable is the number of Internet-connected PCs per worker, soil conditions are important predictors of ICT adoption. Mean soil bulk density enters negatively: a one-standard-deviation increase is associated with 0.018 fewer Internet-connected PCs per worker, equivalent to about 16% of the sample mean of 0.112.[9] Minimum soil coarse-fragment content, by contrast, enters positively: a one-standard-deviation increase is associated with 0.011 additional PCs per worker, or about 10% of mean adoption. These relationships are consistent with local broadband infrastructure being less costly to install and maintain where soil conditions are more favorable for trenching.

Hydrological conditions also play an important role. Greater mean water coverage is a strong geographic predictor of lower adoption: a one-standard-deviation increase in surface water coverage (about 1.3 percentage points) lowers adoption by about 0.017 PCs per worker, roughly 15 percent of mean adoption, reflecting the high cost and technical difficulty of laying cables through water-covered areas. In contrast, greater variability in water presence is associated with higher adoption: a one-standard-deviation increase in the variability of water occurrence raises adoption by about 0.013 PCs per worker (roughly 12 percent of mean adoption). This pattern is consistent with infrastructure being more easily routed in areas with greater flexibility, such as along seasonally dry or elevated corridors that avoid persistently wet terrain.

Finally, terrain ruggedness along the connection path reduces ICT adoption. A one-standard-deviation increase in mean ruggedness lowers adoption by about 0.015 PCs per worker (roughly 13 percent of mean adoption), consistent with higher installation costs in uneven terrain.

Column 2 reports analogous results using LAN-connected PCs per worker. The estimates are broadly similar, with soil conditions and water variability remaining the key predictors. However, fewer variables are selected, and the first stage is weaker (see Table 4). This pattern is consistent with our instruments capturing constraints to accessing external Internet infrastruc-

[9]For each geographic variable, we calculate the reported magnitude by multiplying its first-stage coefficient by its standard deviation in the estimation sample. We express the result both in PCs per worker and as a percentage of mean Internet adoption in that sample. Using one-standard-deviation changes places variables measured in different physical units on a common scale. The estimation-sample standard deviations are 0.468 for mean soil bulk density, 2.085 for minimum soil coarse fragments, 1.006 for the standard deviation of water occurrence, 0.013 for mean water coverage, and 4.865 for mean ruggedness.

ture, which are less directly relevant for internal networking decisions. Accordingly, we use internet-connected PCs per worker as our primary measure of ICT adoption in the subsequent robustness and mechanism analyses.

5.2 Results on Female Employment Shares

Panel A of Table 4 presents the estimated effect of ICT adoption using the number of Internet-connected PCs per worker as the ICT measure. Each entry reports the coefficient on β_1 from a specific specification of regression (1). 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, foreign and state ownership, as well as year and firm fixed effects. It also includes node-specific time trends and distance-to-backbone-by-year trends to account for regional trends and the non-random placement of backbone nodes. We find a positive and statistically significant correlation between ICT adoption and the female labor share. The OLS estimate of 0.037 implies that a one-standard-deviation increase in internet-connected PCs per worker (0.215) is associated with a 0.8 percentage-point higher female labor share (0.215×0.037), a modest partial correlation.

Column 2 addresses endogeneity by instrumenting ICT adoption with the interactions between post-2007 and geographic conditions along the path from each firm to its nearest backbone node, as described in Section 4, while retaining the controls from column 1. Column 3 additionally controls for population growth and allows trends to vary with distance to the nearest urban center. Our baseline specification in column 4 further includes year-specific trends in distance to major infrastructure networks, $\log(\text{DistInfras}_i) \times \theta_t$, to account for the possibility that topography jointly shapes infrastructure placement and post-2007 outcomes. It also adds the commune-level trade-exposure controls, $\overline{\text{FDI}}_c \times \theta_t$ and $\overline{\text{Trade}}_c \times \theta_t$, to absorb differential trends by initial trade exposure.

Across all IV regressions (columns 2–4, panel A), 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 instru-

ments are relevant for the endogenous ICT adoption variable. In addition, the Anderson–Rubin Wald test yields p-values below 0.05, suggesting that the IV estimates are robust to potential weak-instrument concerns.

The results are consistent across columns 2–4 and indicate a substantially larger positive effect of ICT adoption on the female labor share than the OLS estimate reported in column 1. Our baseline estimate (column 4) of 0.305 implies that a one-standard-deviation increase in internet-connected PCs per worker (0.215) raises the female labor share by about 6.6 percentage points (0.215×0.305), roughly one-fifth of the sample mean female labor share of 0.33. Equivalently, moving a firm from the median to the 90th percentile of ICT adoption (from 0.03 to 0.31) raises its female labor share by about 8.5 percentage points. These are economically meaningful effects.

Panel B of Table 4 reports OLS and IV results using an alternative measure of ICT adoption: the 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. The baseline IV estimate of 0.522 implies that a one-standard-deviation increase in LAN-connected PCs per worker (0.22) raises the female labor share by about 11.5 percentage points.

Our IV estimates are larger than the OLS estimates for several reasons. First, OLS is likely biased downward due to non-random selection into ICT adoption. Informality is pervasive in developing countries (La Porta and Shleifer, 2014), and women face higher barriers to formal employment (Duflo, 2012). As shown in Table A.3, in Vietnam, women are underrepresented in the formal sector, where ICT adoption is more prevalent. This suggests that ICT expansion may be concentrated in initially male-dominated areas or areas with lower female labor share, attenuating the OLS estimates. Consistent with this interpretation, we find that the change in ICT adoption is negatively correlated with baseline female labor share (significant at the 1% level).

Second, the IV estimates capture a local average treatment effect (LATE), which may exceed the average partial correlation estimated by OLS. Our instrument, based on geographic connectivity conditions interacted with the post-2007 expansion, primarily shifts ICT adoption

for firms with more favorable connectivity, which may have greater scope for adjustment and thus exhibit larger effects on female labor share.

Finally, measurement error in firm-level ICT variables may also contribute to the gap. Our ICT measures are imperfect proxies for underlying digital adoption, and classical measurement error would attenuate OLS estimates toward zero. The IV approach helps recover the effect of the underlying adoption margin more accurately.

5.3 Parallel Trends

A key identifying assumption underlying our IV strategy is the exclusion restriction: the interaction between geographic connectivity conditions and the post-2007 indicator affect female labor share solely through ICT adoption, and not through any other channel that differentially emerges after 2007. While this assumption is not directly testable, a necessary implication is that, in the absence of the 2007 ICT expansion, female labor share would have evolved similarly across firms facing different geographic connectivity conditions. This condition closely corresponds to the parallel-trends assumption in difference-in-differences designs. Violations of pre-trends would therefore provide direct evidence against the exclusion restriction, whereas parallel pre-trends, while not sufficient to establish it, would indicate that geographic connectivity conditions were not differentially predicting female labor share through channels other than ICT adoption prior to the shock.

Indeed, if firm-level ICT adoption were not directly observed, our setting could be interpreted as a difference-in-differences framework with continuous treatment intensity (Nunn and Qian, 2011; Geraci et al., 2022; Parker and Vogl, 2023; Callaway et al., 2024). Cross-sectional variation in exposure is determined by each firm's connectivity suitability, captured by the geographic characteristics along the path to its nearest ICT backbone node, while time-series variation arises from the nationwide ICT capacity expansion in 2007.

To assess the validity of this assumption, we construct a firm-level measure of treatment intensity that isolates variation driven solely by predetermined geographic connectivity conditions. Using the first-stage estimates in column 1 of Table 3 (based on equation (2)), we define

treatment intensity as the fitted component of ICT adoption explained by the geographic connectivity conditions selected by the LASSO procedure in the "zero-stage":

$$T_i = \hat{\beta}_1 selectGeo_i, \tag{3}$$

where $\hat{\beta}_1$ is estimated from equation (2). By construction, T_i captures only variation in connectivity suitability arising from predetermined geographic characteristics. It is therefore time-invariant and varies across firms. Because firms are assigned to communes, T_i varies at the commune level and is common to all firms within the same commune.

The baseline 2004–2009 sample provides limited leverage for assessing pre-trends, particularly because data are unavailable for 2006. We therefore extend the event study back to 2001 using outcome data from the 2001–2003 annual enterprise surveys. Because harmonized commune identifiers are consistently available only from 2004 onward, we link earlier firm records to their post-2004 commune locations using time-invariant firm identifiers and assign the corresponding predetermined treatment intensity. Since treatment intensity is time invariant, this extension requires neither ICT data nor reconstruction of the instrument for the earlier years. The extended event study is therefore limited to firms observed in the 2004–2009 sample. We do not impose this linkage in the remaining analyses, as restricting the sample to firms observed across coding regimes could introduce survivor bias; all other results use the baseline sample.[10] We then estimate the following event-study specification:

$$FemaleS_{it} = \alpha_0 + \sum_{t=2001}^{2009} \beta_{1t}(\mathbb{1}\{Year_t = 1\} \times T_i) + \gamma_i + \theta_{nt} + \mathbf{X}'_{it}\beta_2 + \mathbf{Z}'_{it}\beta_3 + \epsilon_{it}. \tag{4}$$

All variables are defined as in the baseline specification in equation (1) (column 4 of Table 4). The only difference relative to the baseline specification, besides not using IV, is that we replace the endogenous ICT adoption variable ICT_{it} with the treatment intensity T_i, interacted with year indicators, $(\mathbb{1}\{Year_t = 1\} \times T_i)$. The resulting coefficients, β_{1t}, trace out the year-by-year relationship between geographically predicted connectivity intensity and the female labor share.

[10] We omit 2006 throughout to maintain a consistent set of estimation years.

Panel A of Figure 3 plots the estimated coefficients, β_{1t}. The figure traces the differential evolution of female labor shares associated with treatment intensity relative to the pre-expansion period. The absence of systematic differences in the pre-2007 coefficients provides evidence consistent with the generalized parallel trends assumption in a continuous-treatment setting.[11] After 2007, the coefficients become positive and remain elevated. The effects are modest in the first year after the expansion and appear slightly attenuated in 2008, potentially reflecting broader macroeconomic disruptions associated with the 2008 global financial crisis. Thereafter, the effects persist in subsequent years. As a further check, we re-estimate the baseline 2SLS on this extended 2001–2009 panel: the estimate is essentially unchanged for both the internet and LAN measures and remains significant under weak-IV-robust inference (Appendix Table A.6).

To provide a more transparent treated-versus-control comparison, we discretize treatment intensity by assigning firms to quartiles of T_i within each backbone-node catchment area. Firms in the lowest quartile form the control group, while those in the second, third, and fourth quartiles define the low-, medium-, and high-exposure groups, denoted by T_i^l, T_i^m, and T_i^h. We then re-estimate specification (4), replacing the continuous treatment measure with indicators for these three exposure groups. Panel B of Figure 3 reports the results. We find no evidence of differential pre-trends before the 2007 expansion. After 2007, firms in the medium- and high-exposure groups experience larger increases in female labor share relative to the lowest-exposure quartile, although the differences across treated groups are not statistically significant. This pattern is consistent with the IV estimates identifying effects among firms whose ICT adoption is most responsive to the instrument, particularly those facing more favorable connectivity conditions and greater scope to adjust.

[11] In a difference-in-differences setting with continuous treatment, identification relies on the assumption that, absent the ICT shock, firms with different geographic connectivity conditions would have followed similar trends (Callaway et al., 2024). Although this condition cannot be tested directly, the absence of differential pre-trends in Figure 3, together with the fact that exposure intensity is predetermined and our specification flexibly absorbs location-specific time variation, supports its plausibility.

5.4 Medium-term Effects

In our instrumental variables analysis, firm-level ICT adoption is observed only through 2009, which limits the main IV analysis to the 2004–2009 period. However, the effects of large-scale infrastructure expansion may take time to fully materialize. In this section, we therefore provide additional evidence on the medium-term effects of ICT expansion using alternative empirical strategies.

Our first approach exploits the fact that, while ICT adoption is observed only through 2009, the outcome variable (female labor share) is available through 2011. This allows us to examine medium-term outcome dynamics using the generalized difference-in-differences framework developed in the parallel-trends analysis in Section 5.3. Rather than estimating a fully dynamic event-study specification, we summarize the post-2007 effects by estimating:

$$FemaleS_{it} = \alpha_0 + \beta_1 T_i \times Post07_t + \gamma_i + \theta_{nt} + \mathbf{X}'_{it}\beta_2 + \mathbf{Z}'_{it}\beta_3 + \epsilon_{it}, \tag{5}$$

where T_i denotes treatment intensity, defined in equation (3) as the fitted component of ICT adoption explained solely by the geographic connectivity conditions. This specification is a parsimonious version of the event-study regression in equation (4), replacing the full set of interactions with a single post-2007 term so that β_1 captures the average post-expansion effect.

We estimate this specification over two periods: 2004–2009, corresponding to our main sample, and 2004–2011 to capture medium-term effects.[12] Panel A of Appendix Table A.7 reports the results, with treatment intensity standardized for ease of interpretation. In column 1, a one-standard-deviation increase in treatment intensity is associated with a 1 percentage-point larger post-2007 increase in female labor share. Extending the sample through 2011 yields a similar estimate of 1.1 percentage points, indicating that the effects of ICT expansion persist over the medium term.

Figure A.4 provides complementary evidence from a longer-horizon (2001–2011) event-study based on specification (4). The pre-treatment coefficients are small and statistically in-

[12]These estimates differ from the second stage of the baseline 2SLS results because they are based on a reduced-form relationship using the treatment-intensity measure, which relies solely on variation predicted by geographic conditions interacted with the post-2007 indicator.

distinguishable from zero, consistent with the parallel trends assumption. After 2007, the coefficients become positive and remain elevated through 2011, indicating that the effects of ICT expansion on female labor share persist in the medium term.

As a complementary approach, we implement a two-sample two-stage least squares (TS2SLS) estimator (Angrist and Krueger, 1992). In the first stage, we estimate the relationship between ICT adoption and the geographic connectivity conditions interacted with the policy shock (the IVs) using the subsample with observed ICT data, and use the estimated coefficients to construct predicted ICT adoption for the full sample. In the second stage, we regress female labor share on this predicted measure of ICT adoption. This approach differs from the generalized difference-in-differences framework above by recovering the structural effect of ICT adoption rather than reduced-form exposure effects. At the same time, it is conceptually similar to our baseline 2SLS specification, but extends it by leveraging a larger sample through predicted ICT adoption.

Panel B of Appendix Table A.7 reports the TS2SLS estimates using the number of internet-connected PCs per worker as the measure of ICT adoption. As in Panel A, we present results for both the 2004–2009 sample and the extended 2004–2011 sample. The estimates indicate a positive and statistically significant effect of ICT adoption on female labor share. In column 1 (2004–2009), the two-sample IV estimate is 0.190; extending the sample to 2004–2011 in column 2 yields 0.236, similar in magnitude, indicating that the effects persist over the medium term, consistent with the patterns observed in Panel A. Panel C reports similar results using the number of LAN-connected PCs per worker as the measure of ICT adoption.

To assess persistence, we estimate equation (5) over 2004–2011, separating the post-expansion period into 2007–2009 and 2010–2011. The medium-run estimates are close to the short-run effects, and equality cannot be rejected for either Internet intensity ($F = 0.15$, $p = 0.70$) or LAN intensity ($F = 0.87$, $p = 0.35$). The results therefore indicate persistence rather than continued growth. This persistence is consistent with the structure of Vietnam's ICT expansion, which primarily increased capacity along an already established backbone network rather than building entirely new infrastructure. As a result, although some additional investments, such as extend-

ing local connections or expanding bandwidth, were required, these "last-mile" improvements could be implemented relatively quickly, allowing firms to benefit from enhanced connectivity soon after the expansion.

5.5 Other Firm-level Outcomes

Effects on Employment Levels We have shown that ICT adoption increases the firm-level female labor share. This effect could arise either from firm expansion or from within-firm substitution between male and female workers. To distinguish between these channels, we examine log total, female, and male employment under our baseline specification (1).

Panel A of Table 5 shows no significant effect on total employment (column 1), suggesting that ICT adoption does not induce meaningful expansion in firm size. Instead, ICT adoption significantly increases female employment (column 2) while reducing male employment (column 3). These results suggest that the rise in the female labor share is driven primarily by within-firm workforce recomposition—more female and fewer male workers—rather than by overall employment growth. However, because the Anderson–Rubin Wald test yields a p value of 0.079 for female employment, we interpret this evidence as suggestive.

Effects on Firm Entry and Exit We next examine whether ICT expansion affects firm dynamics along the extensive margin, namely firm exit and entry. For exit, we construct a firm-level indicator equal to one in the last year a firm appears in the census and estimate a discrete-time hazard model relating exit to ICT adoption, instrumented using our baseline strategy. For entry, because firms are not observed prior to entry, we instead analyze entry at the commune–year level. We define the entry rate as the number of new firms entering a commune in a given year divided by the total number of active firms, and estimate IV regressions of entry rates on commune-level ICT exposure using the same instruments as in our baseline strategy. Additional details are provided in Appendix Section C.

Panel B of Table 5 finds no statistically significant effects of ICT expansion on firm exit or entry (columns 1 and 2). However, the Anderson–Rubin weak-instrument-robust tests also fail

to reject a zero effect, so the null results should be interpreted cautiously given their limited precision. To further assess potential selection, we re-estimate the baseline specification using a balanced panel of firms observed throughout 2004–2009 (column 3). The estimated effect is similar to, and slightly larger than, that obtained from the full sample.

In sum, these findings suggest that our main results are not driven by selection and compositional changes in the firm population and instead reflect within-firm adjustments in labor demand.

5.6 Robustness Checks

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.8, 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 allows for deviations from the straight-line connection path due to terrain and reflects the 250 m spatial resolution of our geographic data. As a robustness check, columns 5 and 6 of Appendix Table A.8 use a narrower 0.5 km buffer. The estimates remain positive but smaller and less precise, and the Anderson–Rubin tests narrowly fail to reject the null at the 5% level ($p = 0.073$ and $p = 0.055$). We therefore view these results as directionally consistent with, but weaker than, the baseline evidence.

Third, in our baseline specification, we restrict the sample to firms located within 20 km of the nearest domestic backbone node to ensure comparable pre-2007 trends. Appendix Table A.9 shows that our findings are robust to alternative distance thresholds of 25 km and 15 km.

Fourth, Appendix Table A.10 shows that our results are robust to alternative clustering schemes, including clustering at the district, nearest-node, and province levels.

Fifth, Appendix Table A.11 examines the robustness of the results to additional fixed effects

and alternative sample restrictions. Column 2 adds province×year fixed effects, absorbing all time-varying shocks common to firms within a province. The resulting estimate of 0.261 is close to the baseline estimate of 0.305 in column 1 and remains significant under weak-instrument-robust inference, although the first-stage F-statistic declines because little geographic variation remains after including these fixed effects. Column 3 restricts the sample to firms whose nearest backbone node lies in the same province, approximately 92% of the sample, and yields a similar estimate of 0.279. Column 4 restricts the sample to firms with positive ICT adoption and produces an estimate of 0.360. Finally, column 5 drops the top 1% of the ICT distribution and leaves the estimate virtually unchanged at 0.325, alleviating concerns that the results are driven by outliers.

Finally, although our preferred instrument uses geographic connectivity conditions, which more directly capture connection costs and mitigate concerns about non-random backbone-node placement, we assess robustness using the distance-based instrument common in the literature. Specifically, we instrument ICT adoption with the interaction between each firm's distance to the nearest backbone node and a post-2007 indicator. The specification includes the same fixed effects, firm-level controls, and location characteristics as the baseline, except for distance-by-year trends, which are collinear with the instrument.

Columns 3 and 4 of Table A.12 report the results. The estimates are smaller than the baseline but remain broadly consistent with our main findings: 0.189 for Internet-connected PCs per worker, compared with 0.305 in the baseline, and 0.242 for LAN-connected PCs per worker, compared with 0.522.

6 Individual-level Labor Market Outcomes

Empirical Strategy To assess whether the firm-level effects translate into broader improvements in female labor market outcomes, particularly in facilitating women's transition into the formal sector, where most gender gaps arise as documented in Section 3, and to further explore potential mechanisms, we complement the firm-level analysis with household-level data. This allows us to examine a wider set of outcomes, including employment, wages, and sectoral

reallocation.

Specifically, we use data from the Vietnam Household Living Standards Survey (VHLSS) for the years 2002, 2004, 2006, 2008, 2010, and 2012. To construct the ICT treatment measure at the household level, we geo-code each household at the commune level and compute the geographic connectivity conditions between the household's location and the nearest ICT backbone node. We then map these conditions into a treatment intensity measure by applying equation (3) and the estimated β_1 from the firm-level first stage, assigning each household the value $T_{c(i)}$ based on its commune of residence. Intuitively, this measure captures the intensity of ICT exposure through firm adoption in the household's commune.

We then estimate the following specification for the household-level analysis. Conceptually, this extends our firm-level generalized DiD framework in equation (5) by introducing an additional layer of variation across gender using individual-level data, and can be interpreted as a triple-difference (DDD) design:

$$\begin{aligned} Y_{it} = \alpha_0 + \beta_1(T_i \times Post07_t \times Female_i) + \beta_2(T_i \times Post07_t) \\ + \beta_3(Female_i \times T_i) + \kappa_{gt} + \mathbf{X}'_{it}\beta_4 + \gamma_c + \theta_{nt} + \mathbf{Z}'_{it}\beta_5 + \epsilon_{it}. \end{aligned} \tag{6}$$

The dependent variable Y_{it} denotes the labor market outcome of individual i in year t. T_i measures the treatment intensity in commune c where individual i resides. $Post07_t$ is an indicator equal to one for years 2007 and later, and $Female_i$ is an indicator for gender.

The coefficient β_2 captures the effect of ICT expansion on male individuals, while the coefficient on the triple interaction, β_1, is our parameter of interest and measures the differential impact of ICT expansion on women relative to men. Consistent with a triple-difference framework, we include gender-by-year fixed effects, κ_{gt}, to flexibly account for gender-specific time trends. We also include the interaction $Female_i \times T_i$ to allow for baseline differences in treatment intensity across gender.

Unlike the firm-level data, which are structured as a panel, the household data consist of repeated cross-sections.[13] As a result, we cannot include household fixed effects. Instead, we

[13] Although a household panel exists for the 2002–2004, 2004–2006, and 2006–2008 waves of the VHLSS (McCaig, 2025), we are unable to construct a sufficiently long panel that spans the 2007 policy shock while maintaining

include commune fixed effects, γ_c, the most granular geographic unit available, and control for a rich set of individual- and household-level characteristics $\mathbf{X}_{it}$, including age and age squared; marital status; education indicators; household size and its square; household composition (female share, adult share, and college-educated share); and average household education and age (each with quadratic terms). We further include node-by-year fixed effects, θ_{nt}, and the same set of flexible locational trend controls, $\mathbf{Z}_{it}$, as in the firm-level specification in equation (1). Finally, consistent with our main firm-level analysis, we restrict the sample to individuals residing within 20 km of the nearest backbone node.

Results Table 6 reports the results. For each panel, even-numbered columns present estimates from specification (6). Our main parameter of interest is the coefficient on the triple interaction term, β_1, which captures the differential effect of ICT expansion for women. Odd-numbered columns report estimates from a standard difference-in-differences specification without the gender interaction, capturing the overall effect of ICT expansion. Treatment intensity T_i is standardized to facilitate interpretation.

Panel A examines overall labor supply and formal employment. We find no statistically significant effects of ICT expansion on the probability of working or on formal-sector employment, nor any differential effects by gender. Consistent with this, Table A.13 shows no meaningful impacts on self-employment, salaried work, household income, or working hours. While there is some suggestive evidence of reduced engagement in secondary jobs and housework, these effects are small and not gender-specific.[14] Overall, these results indicate that ICT expansion does not operate through increased labor force participation or job creation, consistent with our firm-level evidence (Table 5), which shows no effect on total employment.

Given that ICT adoption occurs primarily within formal firms and is most relevant for formal-sector activities, Panel B focuses on outcomes among formal-sector workers.[15] Columns

adequate variation in ICT exposure.

[14] Worked is defined as whether an individual worked in the previous year. Self-employment refers to work in household production or household business activities. Salaried work refers to employment for wages. Formal employment includes employment in private firms, state-owned enterprises, public institutions, or foreign-invested firms.

[15] Because Panels B and C condition on employment in the formal and informal sectors, respectively, their observation counts sum to the number of employed individuals rather than to the full Panel A sample. Wage outcomes

1-2 examine sectoral reallocation within the formal sector. The positive and statistically significant coefficient on the triple interaction term in column 2 indicates that ICT expansion increases women's likelihood of working in non-agricultural activities relative to men. Quantitatively, after 2007, a one-standard-deviation increase in treatment intensity leads to a roughly 2 percentage point larger increase in non-agricultural employment for women than for men. This reflects a differential shift of women, rather than men, toward non-agricultural activities in response to ICT expansion. This estimate captures a gender differential rather than an absolute increase for women: the implied total effect for women is small and statistically indistinguishable from zero. This pattern is consistent with our firm-level findings that ICT adoption increases the female share of employment in formal firms.

Turning to earnings, columns 3–4 show that ICT expansion significantly increases hourly wages among formal-sector workers. After 2007, workers in communes with a one-standard-deviation higher treatment intensity earn approximately 8–9 percent higher wages relative to those in lower-intensity communes. The triple interaction term in column 4 is small and statistically insignificant, indicating that these wage gains are broadly similar for men and women.

Panel C examines the informal sector and shows little evidence of any effects. ICT expansion does not significantly affect sectoral allocation or wages, either overall or differentially by gender, indicating that its labor market impacts are concentrated in the formal sector where ICT adoption takes place.

We also provide evidence consistent with the parallel trends assumption in Figure A.5. The key parameter of interest is β_1 in specification (6), which captures the differential effect of ICT expansion across gender. To assess pre-trends, we replace the interaction term $(T_i \times Post07_t \times Female_i)$ with a full set of time interactions, $(T_i \times \mathbb{1}\{Year_t = 1\} \times Female_i)$, and plot the year-specific coefficients β_{1t}. Across all outcomes, we find no statistically significant differential pre-trends by gender prior to 2007.

Taken together, these results suggest that ICT expansion affects gender differences in labor market outcomes primarily through compositional changes within the formal sector, rather than through aggregate employment effects. While ICT raises formal-sector wages for men, its

are further restricted to wage earners.

effects along other margins are limited, especially relative to those for women. This pattern is consistent with a growing body of empirical evidence. Our findings align with Juhn et al. (2014), which document that in Mexico, technological change reduces demand for physically demanding tasks and increases women's relative employment and wages in blue-collar work, with men losing ground on the employment margin. Dettling (2017) finds that high-speed home internet in the United States raises labor force participation among married women, with no corresponding effects for single women or men. Similarly, Viollaz and Winkler (2022) show that mobile broadband adoption in Jordan increases female labor force participation without affecting men. In a broader cross-country context, Chiplunkar and Goldberg (2022) find that 3G expansion across low- and middle-income countries substantially raises female labor force participation, with more limited effects for men, and that gender differences in job composition are nuanced rather than uniformly favoring one group.

7 Mechanism

Our results show that ICT expansion increases the female share of firm employment without significantly increasing total employment, firm entry, or transitions into formal-sector employment. These patterns suggest that ICT primarily changes the composition of employment within firms rather than generating a broad increase in labor demand. As discussed in the Introduction, the task-based literature suggests that ICT may alter production processes in ways that increase the relative demand for tasks in which women have a comparative advantage. In this section, we examine this mechanism using firm- and individual-level evidence.

We focus on five dimensions of task content: routine, manual, social, customer-facing, and service-oriented tasks.[16] ICT tends to substitute for routine and manual tasks while complementing non-routine work involving interpersonal interaction, customer engagement, and service provision. Prior research also suggests that women have a comparative advantage in so-

[16] Routine tasks consist of repetitive, rule-based activities; manual tasks require physical effort; social tasks involve interpersonal communication and coordination; customer-facing tasks require direct interaction with clients; and service-oriented tasks emphasize care, assistance, or support.

cial and service-oriented work relative to physically intensive manual tasks.[17] Consistent with this premise, the female labor share in our data is positively correlated across industries with service, customer-facing, and social task intensity, with correlations of 0.24, 0.21, and 0.13, respectively, and negatively correlated with manual task intensity (-0.34). Its correlation with routine task intensity is close to zero (-0.06).

This mechanism yields two testable predictions. First, the effect of ICT adoption on the female employment share should be larger in industries that rely less on routine and manual tasks and more on social, customer-facing, and service-oriented tasks. We test this prediction using firm-level data by examining heterogeneity across industries with different task intensities. Second, ICT exposure should shift women's occupational allocation, relative to men's, away from routine and manual tasks and toward social, customer-facing, and service-oriented tasks. We test this prediction using individual-level data.

To implement these tests, we construct task-intensity measures from U.S. O*NET occupational data and standard task-based classifications (U.S. Department of Labor, Employment and Training Administration, 2025). For the household analysis, we merge task measures directly using workers' reported occupations. Because the enterprise surveys contain industry but not detailed occupation codes, we construct industry-level task intensities for the firm analysis by combining occupational task scores with industry-level occupational employment shares. Appendix Section B provides further details.

Firm Level Evidence We first examine whether the effect of ICT adoption varies with industry-level task intensity by estimating

$$\begin{aligned} FemaleS_{it} = \alpha_0 + \beta_1 ICT_{it} + \beta_2 ICT_{it} \times TaskIntensity_j + \gamma_i + \theta_{nt} + \eta_{mt} \\ + \mathbf{X}'_{it}\beta_3 + \mathbf{Z}'_{it}\beta_4 + \epsilon_{it}. \end{aligned} \tag{7}$$

Relative to equation (1), this specification interacts ICT adoption with standardized industry-level task intensity, $TaskIntensity_j$. We instrument ICT_{it} and its interaction with task intensity

[17] See, among others, Autor et al. (2003); Autor and Dorn (2013); Deming (2017); Acemoglu and Restrepo (2020); Welch (2000); Rendall (2018); Blau and Kahn (2017), and Cortes and Pan (2018).

using $selectGeo_i \times Post07_t$ and $selectGeo_i \times Post07_t \times TaskIntensity_j$, respectively, and include the baseline controls and sector×year fixed effects, η_{mt}. The latter absorb broad sector-specific trends that may be correlated with both task composition and female employment.

Panel A of Table 7 reports the results. Column 1 reproduces the baseline estimate, while columns 2–6 estimate equation (7) by interacting ICT adoption with routine, manual, social, customer-facing, and service-oriented task intensity. The interactions with routine and manual intensity are negative and statistically significant: a one-standard-deviation increase in either measure reduces the estimated effect of ICT adoption by 0.053–0.059, roughly one-fifth of the baseline effect. By contrast, the interactions with social, customer-facing, and service-oriented intensity are positive and significant, increasing the estimated effect by approximately 0.04–0.05. Thus, ICT adoption has larger effects on female employment in industries that rely more heavily on interpersonal and service-oriented tasks and smaller effects in routine- and manual-intensive industries.

Individual Level Evidence Panel B of Table 7 uses VHLSS occupational data to examine whether ICT expansion differentially changes the task content of women's and men's employment. We harmonize occupations across survey waves, link them to the task measures described in Appendix B, and estimate specification (6) with standardized task intensity as the outcome. The coefficient on $T_i \times Post07_t \times Female_i$ is negative and significant for routine tasks, indicating that greater ICT exposure shifts women, relative to men, away from routine occupations. Manual task intensity also declines, but the gender difference is not statistically significant. In contrast, the corresponding coefficients for social, customer-facing, and service-oriented tasks are positive and significant. Quantitatively, a one-standard-deviation increase in treatment intensity reduces women's relative routine-task intensity by about 0.03 standard deviations and raises their relative social, customer-facing, and service-oriented task intensity by 0.03–0.04 standard deviations. These shifts are meaningful relative to baseline occupational sorting: before the expansion, women worked in occupations that were 0.26 and 0.33 standard deviations less routine- and manual-intensive than men's, and 0.18, 0.16, and 0.37 standard deviations more intensive in social, customer-facing, and service tasks. The estimated changes

therefore amount to roughly one-tenth of the baseline gender gap in routine and service tasks and about one-fifth of the gap in social and customer-facing tasks.

Additional Heterogeneity: By Sector and Firm Size We next examine heterogeneity by sector and firm size. These broad firm characteristics may capture differences in production scale, initial ICT adoption, and the scope for organizational adjustment, but they should not be interpreted as direct tests of the task-reallocation mechanism. We extend specification (7) by interacting instrumented ICT adoption with indicators for manufacturing and firm size—small, medium, and large, defined by terciles of initial employment—in place of $TaskIntensity_j$.

Table A.14 reports the results. Column 1 presents the baseline estimate for reference. Column 2 shows stronger effects among larger firms: the interaction coefficients for medium and large firms are 0.33 and 0.24, respectively, and both are statistically significant relative to small firms. Column 3 also shows a larger effect in manufacturing, with an additional effect of 0.28 relative to non-manufacturing firms. Although the Kleibergen–Paap F statistics for these interaction specifications are modest, the relevant effects remain significant under weak-instrument-robust inference.

The larger manufacturing estimate likely reflects a combination of firm characteristics rather than task composition alone. Manufacturing firms in our sample are substantially larger on average, start from lower levels of ICT adoption per worker (0.03 versus 0.12 internet-connected PCs per worker), and have somewhat higher initial female labor shares (0.41 versus 0.37). Their greater scale and larger scope for ICT expansion may therefore make employment adjustments more pronounced. At the same time, manufacturing is more routine- and manual-intensive than services, while the task-intensity results show smaller effects in industries that rely more heavily on these tasks. The sector comparison and the task-intensity analysis thus capture different margins of heterogeneity.

Taken together, the results suggest that the employment effects of ICT depend both on the task content of production and on firms' capacity to adopt and adjust to new technology. The task-intensity estimates provide the clearest evidence for the proposed reallocation mechanism, while the sector and size patterns indicate that firm scale and initial adoption opportunities also

shape the magnitude of the response. We therefore view task reallocation as an important, but not exclusive, channel through which ICT increases the female share of employment.

8 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 upgrade of its ICT infrastructure in 2007. Using detailed firm-level data, we exploit exogenous variation in access to ICT infrastructure, arising from both the timing of the 2007 broadband upgrade 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.

We find that these gains are primarily driven by labor reallocation within the formal sector, with women becoming more likely to transition from agricultural to non-agricultural employment. The effects are strongest in industries with lower routine task intensity and greater reliance on social and service-oriented tasks. This pattern is consistent with a task-based mechanism: ICT complements non-routine, interpersonal, and cognitively intensive activities, domains in which women tend to have comparative advantages. Individual-level evidence further supports this mechanism.

While ICT expansion generates positive wage effects in the formal sector, it does not significantly increase overall employment or transitions into formal work. These findings suggest that the benefits of digital infrastructure depend critically on broader institutional and labor market conditions, particularly barriers to entering the formal sector, which remain substantial in many developing countries (La Porta and Shleifer, 2014; McCaig and Pavcnik, 2018; Ulyssea, 2018). We also find that effects are larger in manufacturing and among larger firms, where adoption is more feasible and task reallocation is more flexible. This implies that reducing barriers to technology adoption may help amplify the gender-equalizing effects of ICT expansion. Although our analysis focuses on Vietnam, the underlying mechanisms, task reallocation and adoption frictions, are likely to apply more broadly to other countries undergoing similar

digital transitions. Future research could examine how rapidly evolving technologies, such as artificial intelligence, further reshape labor market outcomes and gender disparities.

Declaration of generative AI and AI-assisted technologies in the writing process *During the preparation of this work, the authors used ChatGPT to proofread and improve the language and Claude to assist with constructing selected descriptive analyses. The authors reviewed and edited the output as needed and take full responsibility for the content of the published article. .*

References

ABADIE, A., S. ATHEY, G. W. IMBENS, AND J. M. WOOLDRIDGE (2023): "When should you adjust standard errors for clustering?" *The Quarterly Journal of Economics*, 138, 1–35.

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.

ANGRIST, J. D. AND A. B. KRUEGER (1992): "The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples," *Journal of the American statistical Association*, 87, 328–336.

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.

BAHIA, K., P. CASTELLS, G. CRUZ, T. MASAKI, X. PEDRÓS, T. PFUTZE, C. RODRÍGUEZ-CASTELÁN, AND H. WINKLER (2024): "The welfare effects of mobile broadband internet: Evidence from Nigeria," *Journal of Development Economics*, 170, 103314.

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.

BEKKER, P. A. (1994): "Alternative approximations to the distributions of instrumental variable estimators," *Econometrica: Journal of the Econometric Society*, 657–681.

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.

BERAJA, M., A. KAO, D. Y. YANG, AND N. YUCHTMAN (2023): "AI-tocracy," *The Quarterly Journal of Economics*, 138, 1349–1402.

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.

CALLAWAY, B., A. GOODMAN-BACON, AND P. H. C. SANT'ANNA (2024): "Difference-in-Differences with a Continuous Treatment," Tech. Rep. 32117, National Bureau of Economic Research.

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.-H., R. M. GONZALEZ, AND A. S. MOBARIZ (2025): "Mobile Phones and Gender Norms: Evidence from Afghanistan," *Available at SSRN 5399104*.

CIVELLI, A., A. GADUH, A. D. ROTHENBERG, AND Y. WANG (2023): "Urban sprawl and social capital: evidence from Indonesian cities," *The Economic Journal*, 133, 2110–2146.

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.

DETTLING, L. J. (2017): "Broadband in the labor market: The impact of residential high-speed internet on married women's labor force participation," *Ilr Review*, 70, 451–482.

DUFLO, E. (2012): "Women empowerment and economic development," *Journal of Economic literature*, 50, 1051–1079.

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. G., P. A. ROSEN, E. CARO, R. CRIPPEN, R. DUREN, S. HENSLEY, M. KOBRICK, M. PALLER, E. RODRIGUEZ, L. ROTH, ET AL. (2007): "The shuttle radar topography mission," *Reviews of geophysics*, 45.

FIBER BROADBAND ASSOCIATION AND CARTESIAN (2024): "Fiber Deployment Annual Report 2023," Annual report, Fiber Broadband Association.

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.

HANSEN, C., J. HAUSMAN, AND W. NEWEY (2008): "Estimation with many instrumental variables," *Journal of Business & Economic Statistics*, 26, 398–422.

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., J. M. DE JESUS, G. HEUVELINK, M. R. GONZALEZ, M. KILIBARDA, A. BLAGOTIĆ, W. SHANGGUAN, M. N. WRIGHT, X. GENG, B. BAUER-MARSCHALLINGER, ET AL. (2017): "SoilGrids250m: Global gridded soil information based on machine learning," *PLoS one*, 12.

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 (2025): "The economic impact of internet connectivity in developing countries," *Annual Review of Economics*, 17.

HO, L., S. JALOTA, AND A. KARANDIKAR (2024): "Bringing work home: Flexible arrangements as gateway jobs for women in West Bengal," *Structural Transformation and Economic Growth (STEG) Working Paper*, 80.

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

HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2019): "The allocation of talent and us economic growth," *Econometrica*, 87, 1439–1474.

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

INTERNATIONAL STEERING COMMITTEE FOR GLOBAL MAPPING (2007a): "Vietnam Railroads, 2007," [Shapefile].

——— (2007b): "Vietnam Roads, 2007," Shapefile.

——— (2008): "Vietnam Water Course, 2008," Shapefile.

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.

KUHN, P. AND H. MANSOUR (2014): "Is internet job search still ineffective?" *The Economic Journal*, 124, 1213–1233.

KUSUMAWARDHANI, N., R. PRAMANA, N. S. SAPUTRI, AND D. SURYADARMA (2023): "Heterogeneous impact of internet availability on female labor market outcomes in an emerging economy: Evidence from Indonesia," *World Development*, 164, 106182.

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

LA PORTA, R. AND A. SHLEIFER (2014): "Informality and development," *Journal of economic perspectives*, 28, 109–126.

LEE, R. C. (2011): "Telecommunications in Viet Nam," in *The Impacts and Benefits of Structural Reforms in the Transport, Energy and Telecommunications Sectors in APEC Economies*, ed. by C. Findlay, Singapore: Asia-Pacific Economic Cooperation Policy Support Unit, chap. 20, 415–432, aPEC#211-SE-01.1.

LIANG, Y., I. RUDIK, AND E. Y. ZOU (2025): "Economic activity and biodiversity in the United States," *Journal of Public Economics*, 250, 105482.

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. (2025): "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," Tech. rep., National Bureau of Economic Research.

MICHAELS, G., A. NATRAJ, AND J. V. REENEN (2014): "Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years," *The 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.

PARKER, S. W. AND T. VOGL (2023): "Do conditional cash transfers improve economic outcomes in the next generation? Evidence from Mexico," *The Economic Journal*, 133, 2775–2806.

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, T. (2014): *Capital in the twenty-first century*, Harvard University Press.

RENDALL, M. (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., M. MELCHIORRI, R. I. MARI, P. FLORIO, S. FREIRE, C. JACOBS-CRISIONI, L. MAFFENINI, P. POLITIS, J. H. UHL, A. CARIOLI, ET AL. (2025): "GHSL WUP Projections Data Package 2025," .

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.

SURI, T. AND W. JACK (2016): "The long-run poverty and gender impacts of mobile money," *Science*, 354, 1288–1292.

TUAN, T. M. (2011): "Broadband in Vietnam: Forging Its Own Path," Broadband Strategies Toolkit case study 69195, infoDev / World Bank, Washington, D.C.

ULYSSEA, G. (2018): "Firms, informality, and development: Theory and evidence from Brazil," *American Economic Review*, 108, 2015–2047.

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.

VIETNAM INTERNET NETWORK INFORMATION CENTER (2013): "Report on Viet Nam Internet Resources 2013," .

VIOLLAZ, M. AND H. WINKLER (2022): "Does the internet reduce gender gaps? The case of Jordan," *The Journal of Development Studies*, 58, 436–453.

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 BANK (2026): "Gender Data Portal: Vietnam," https://genderdata.worldbank.org/en/economies/vietnam, accessed: 2026-04-19.

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

Table 1: Descriptive Statistics of Main Variables, Firm Level

	Mean	Std. Dev.	Obs.
Outcome			
Female labor share	0.331	0.230	180,025
ICT adoption			
Internet-connected PCs per worker	0.112	0.215	180,025
LAN-connected PCs per worker	0.101	0.220	175,838
Firm characteristics			
Log employment	3.805	1.231	180,025
Log total revenue	9.209	2.058	180,025
Log assets	9.261	1.714	180,025
Foreign-owned (0/1)	0.093	0.291	180,025
State-owned (0/1)	0.100	0.300	180,025

Notes: This table reports means, standard deviations, and the number of firm-year observations for the main variables, computed on the *estimation sample* used in our headline specification—firms within 20 km of the nearest backbone node over 2004–2009 with non-missing ICT adoption and control variables (column 4 of Table 4, $N = 180{,}025$). The LAN measure is available for slightly fewer firm-years. The female labor share is the share of female workers in total firm employment (a 0–1 fraction); ICT adoption is measured as the number of internet- or LAN-connected PCs per worker and enters the regressions in levels. "Foreign-owned" and "State-owned" are indicator variables; the omitted reference category is domestically, privately owned firms.

Table 2: Distance of Firms and Individuals to the Nearest Domestic Backbone Node

Panel A	**Total Admin Units**	
	Firm Level	Individual Level
Number of Provinces	64	63
Number of Districts	665	679
Number of Communes	6,997	3,496

Panel B

	Firm Level		**Individual Level**	
	Total	Percentage	Total	Percentage
# Located within 5 km from the Backbone	34,272	29%	47,451	8%
# Located within 10 km from the Backbone	70,047	60%	144,989	25%
# Located within 15 km from the Backbone	91,746	78%	251,359	43%
# Located within 20 km from the Backbone	102,920	88%	336,533	58%
# Located within 25 km from the Backbone	108,724	93%	413,222	71%
Total	117,333	100%	584,209	100%

Notes: This table summarizes the spatial distribution of firms and individuals relative to the nearest domestic backbone node. Firm-level data are drawn from the Vietnam Enterprise Survey (2004–2009), while individual-level data come from the Vietnam Household Living Standards Survey (VHLSS, 2002–2012). Panel A reports the number of administrative units covered in each dataset at the province, district, and commune levels. Panel B reports the distribution of firms and individuals across distance bins relative to the nearest backbone node, including total counts and cumulative percentages.

Table 3: FIRST STAGE

	# PCs w/ Internet per Worker (level)	# PCs w/ LAN per Worker (level)
	(1)	(2)
Soil Bulk Density (mean) $\times Post07_t$	-0.038***	-0.024***
	(0.010)	(0.009)
Soil Coarse Fragments (min) $\times Post07_t$	0.005**	0.007***
	(0.002)	(0.002)
Observed Water Coverage (mean) $\times Post07_t$	-1.302***	
	(0.215)	
The Frequency Water was Present (std) $\times Post07_t$	0.013***	
	(0.003)	
Ruggedness (mean) $\times Post07_t$	-0.003***	
	(0.001)	
N	180,025	175,838
N Clusters	3,423	3,411
Adj. R^2	0.643	0.636

Notes: This table reports first-stage estimates from equation (2), relating ICT adoption to geographic connectivity conditions interacted with the post-2007 indicator, as selected by the LASSO procedure. For ease of interpretation, all geographic variables are rescaled by dividing by 10. Column 1 uses the number of internet-connected PCs per worker as the dependent variable, while Column 2 uses the number of LAN-connected PCs per worker. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4: MAIN RESULTS, STEPWISE SPECIFICATIONS

Panel A:	**(1) OLS**	**(2) IV**	**(3) IV**	**(4) IV**
# of PCs w/ internet per worker	0.037***	0.227**	0.250**	0.305***
	(0.004)	(0.090)	(0.110)	(0.103)
N	180,025	180,025	180,025	180,025
N Clusters	3,423	3,423	3,423	3,423
Kleibergen-Paap Wald Rank F Stat		16.131	13.672	14.489
Under Id. Test (KP Rank LM Stat)		94.309	50.580	77.456
p-Value		0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)		15.309	12.655	18.940
p-Value		0.018	0.013	0.002
Panel B:	**(1) OLS**	**(2) IV**	**(3) IV**	**(4) IV**
# of PCs w/ LAN per worker	0.023***	0.363***	0.387***	0.522***
	(0.004)	(0.138)	(0.140)	(0.155)
N	175,838	175,838	175,838	175,838
N Clusters	3,411	3,411	3,411	3,411
Kleibergen-Paap Wald Rank F Stat		7.595	8.767	10.210
Under Id. Test (KP Rank LM Stat)		23.749	26.851	23.067
p-Value		0.000	0.000	0.000
AR Wald Test (Weak IV Robust Inf.)		11.327	12.868	17.533
p-Value		0.010	0.005	0.000
Firm Level Control	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Node-Year FE	Yes	Yes	Yes	Yes
Distance to Node × Year	Yes	Yes	Yes	Yes
Population			Yes	Yes
Distance to Urban Center × Year			Yes	Yes
Distance to Infrastructure × Year				Yes
Initial Trade Exposure × Year				Yes

Notes: This table reports OLS and IV estimates of the effect of ICT adoption on female labor share. Each cell reports the estimated coefficient on ICT adoption from a specification of equation (1). Column 1 presents OLS estimates, while Columns 2–4 report IV-LASSO estimates with progressively richer controls, as indicated in the table. Panel A uses the number of internet-connected PCs per worker as the measure of ICT adoption, while Panel B uses the number of LAN-connected PCs per worker. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. Kleibergen–Paap (KP) Wald F-statistics, underidentification tests (KP LM), and Anderson–Rubin (AR) weak-IV robust tests are reported for IV specifications. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 5: ICT AND OTHER FIRM-LEVEL OUTCOMES

Panel A: Employment			
	Log Total Employment (1)	**Log Female Employment (2)**	**Log Male Employment (3)**
# of PCs w/ internet per worker	0.604	0.925**	-0.625***
	(0.555)	(0.376)	(0.207)
N	180,025	180,025	180,025
Kleibergen-Paap Wald Rank *F* Stat	13.015	14.489	14.489
AR Wald Test (Weak IV Robust Inf.)	18.601	9.862	23.777
p-Value	0.002	0.079	0.000
Panel B: Entry and Exit			
	Exit (1 0) (1)	**Entry Rate (Commune Level) (2)**	**Female Share (Balanced Panel) (3)**
# of PCs w/ internet per worker	-0.067	0.225	0.355***
	(0.244)	(1.491)	(0.111)
N	157,539	17,786	104,367
Kleibergen-Paap Wald Rank *F* Stat	10.415	8.648	11.969
AR Wald Test (Weak IV Robust Inf.)	7.032	4.548	16.925
p-Value	0.318	0.603	0.010

Notes: This table reports the effects of ICT adoption on outcomes other than female labor share. Panel A examines total employment using specification (1), with outcomes as indicated in the column headers. Panel B examines firm entry and exit. Column 1 estimates firm exit using specification (8). Column 2 examines firm entry at the commune level, as described in Appendix Section C. Column 3 re-estimates the baseline specification (1) using a balanced panel. ICT adoption is measured using the number of internet-connected PCs per worker. All regressions use the same set of instruments as reported in Column 1 of Table 3. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 6: Individual Labor Market Outcomes and ICT by Gender

Panel A: Overall Employment and Formal Sector	**Worked (1 0)**		**Formal (1 0)**	
	(1)	**(2)**	**(3)**	**(4)**
Treatment Intensity × Post 2007	-0.002	-0.004	0.000	0.003
	(0.006)	(0.006)	(0.007)	(0.007)
Treatment Intensity × Post 2007 × Female		0.004		-0.005
		(0.004)		(0.004)
N	242,374	242,374	193,943	193,943
Panel B: Formal Employment	**Non-Ag Sector (1 0)**		**Hourly Wages (Log)**	
	(1)	**(2)**	**(3)**	**(4)**
Treatment Intensity × Post 2007	-0.015	-0.024	0.081**	0.089***
	(0.020)	(0.020)	(0.032)	(0.032)
Treatment Intensity × Post 2007 × Female		0.021***		-0.016
		(0.007)		(0.012)
N	42,689	42,689	41,713	41,713
Panel C: Informal Employment	**Non-Ag Sector (1 0)**		**Hourly Wages (Log)**	
	(1)	**(2)**	**(3)**	**(4)**
Treatment Intensity × Post 2007	-0.018	-0.012	0.002	0.005
	(0.017)	(0.017)	(0.033)	(0.033)
Treatment Intensity × Post 2007 × Female		-0.016		-0.008
		(0.011)		(0.013)
N	46,083	46,083	41,269	41,269

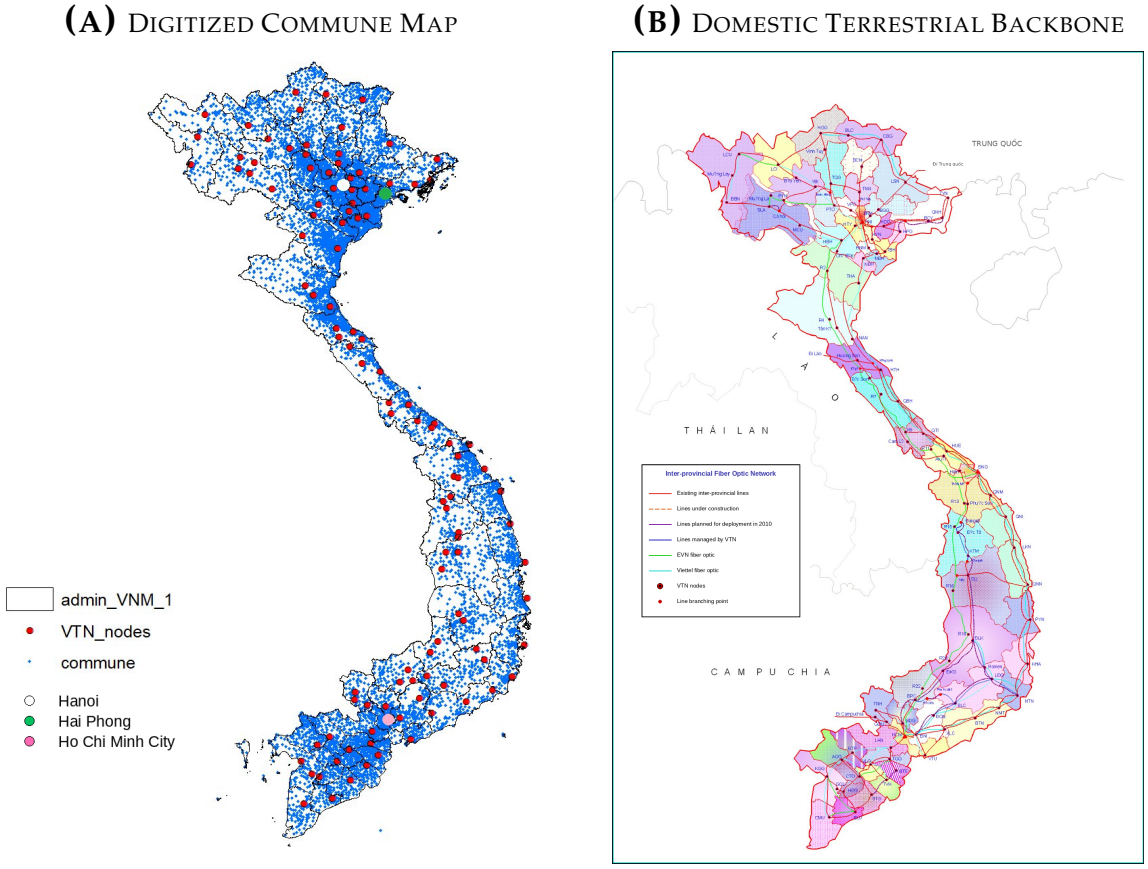
Notes: This table presents the effects of firm-level ICT adoption on individual labor market outcomes. Each cell reports the effect for men and the female differential (the two treatment-interaction coefficients from equation (6)), with dependent variables indicated in the column headings. Panel A reports results for overall employment and formal sector participation. Panel B restricts the sample to workers in the formal sector, while Panel C focuses on workers in the informal sector. Panel A includes all sampled individuals and reports outcomes defined for everyone (whether the individual worked, and whether that work is formal). Panels B and C instead condition on individuals employed in the formal and informal sectors, respectively, so their observation counts sum to employed individuals rather than to the full Panel A sample; the hourly-wage columns are further restricted to wage earners. In each panel, the first row (*Treatment Intensity × Post 2007*) reports the effect for men, the reference category, and the second row (which adds the interaction with *Female*) reports the female differential; the total effect for women is the sum of the two. The sample is limited to individuals located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 7: THE TASK-REALLOCATION MECHANISM: FIRM- AND INDIVIDUAL-LEVEL EVIDENCE

	Baseline	By Task Intensity				
	(1)	Routine (2)	Manual (3)	Social (4)	Customer (5)	Service (6)
Panel A: Firm level						
# of PCs w/ internet per worker	0.305***	0.336**	0.222	0.350***	0.347***	0.257*
	(0.103)	(0.142)	(0.154)	(0.133)	(0.129)	(0.142)
... × Task Intensity		-0.053*	-0.059*	0.036*	0.038**	0.052**
		(0.030)	(0.032)	(0.021)	(0.019)	(0.020)
N	180,025	162,933	162,933	162,933	162,933	162,933
Kleibergen-Paap Wald Rank F Stat	14.489	6.495	6.270	7.470	7.716	6.848
AR Wald Test (Weak IV Robust Inf.)	18.940	113.161	116.480	96.493	94.280	96.464
p-Value	0.002	0.000	0.000	0.000	0.000	0.000
Panel B: Individual level						
Treatment Intensity × $Post07_t$		-0.012	-0.061**	-0.028	-0.025	-0.046*
		(0.030)	(0.031)	(0.028)	(0.026)	(0.028)
... × $Female_i$		-0.026*	-0.007	0.038***	0.035**	0.039**
		(0.015)	(0.015)	(0.014)	(0.015)	(0.016)
N		84,500	84,500	84,500	84,500	84,500

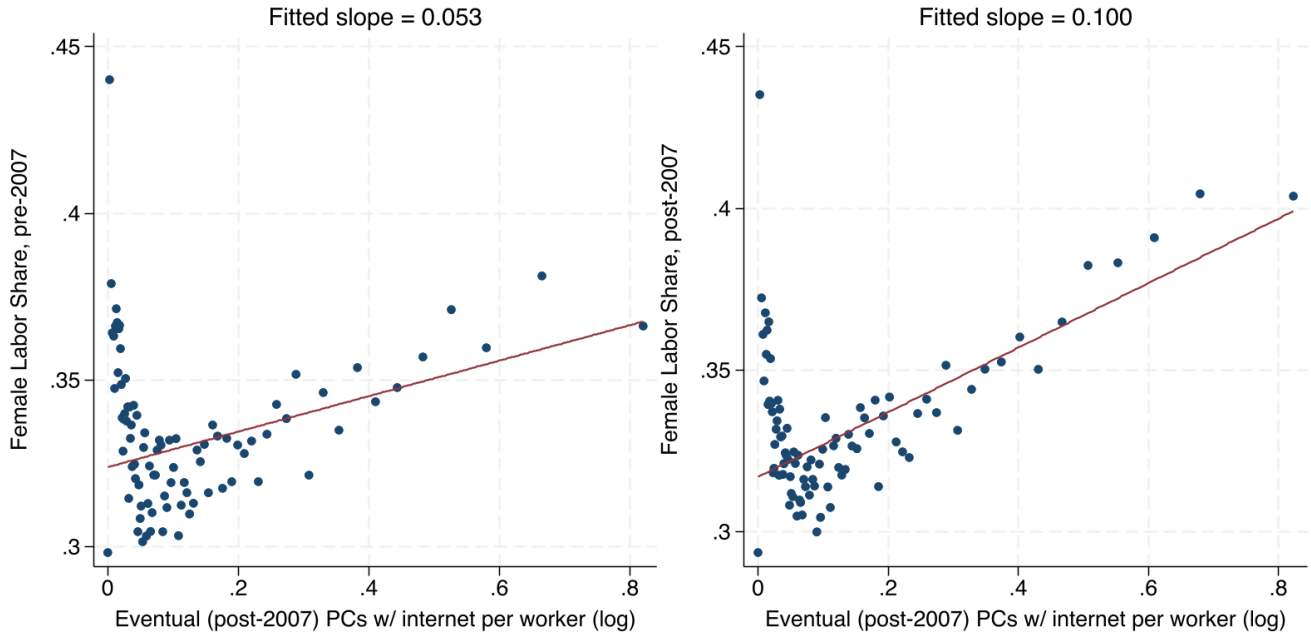
Notes: This table reports the two tests of the task-reallocation channel. *Panel A* presents firm-level 2SLS estimates of the effect of ICT adoption on the female labor share, allowing the effect to vary with industry-level task intensity (equation (7)). Column 1 reproduces the baseline estimate (column 4 of Table 4); columns 2–6 restrict the sample to industries with available O*NET task-intensity measures and add the interaction between ICT adoption and each task-intensity measure, additionally controlling for sector-by-year fixed effects so the interactions are identified net of manufacturing-versus-services trends. Because of the sample restriction and the added controls, the column-1 sample differs from columns 2–6; the instruments (and their task-intensity interactions) are those in column 1 of Table 3. *Panel B* presents individual-level estimates (equation (6)) using occupational task intensity as the outcome, standardized to mean zero and unit standard deviation in the estimation sample; the coefficients are in standard-deviation units. Both panels restrict to units within 20 km of the nearest backbone node. Standard errors are clustered at the commune level. */**/*** denote significance at the 10% / 5% / 1% levels.

Figure 1: Vietnam's Terrestrial Backbone Network and Commune Distribution



Notes: This figure provides an overview of Vietnam's terrestrial backbone network and the spatial distribution of communes. Panel A plots the locations of Vietnam's main terrestrial backbone nodes (red) alongside the centroids of communes (blue). Administrative boundaries correspond to first-level divisions (provinces). The backbone node locations are digitized from ITU transmission network data. The province boundaries and commune centroids used in the digitized map are obtained from the GADM database (version 3.4, April 2018). Panel B presents the original map of Vietnam's terrestrial backbone network and node sites, provided by the International Telecommunication Union (ITU) and Vietnam Telecom National (VTN), obtained from VNPT-I and depicting the network circa 2009 (the map marks certain lines as planned for deployment in 2010). The network consists of fiber-optic transmission lines connecting major cities, including Hanoi (white dot) and Ho Chi Minh City (pink dot), as well as regional networks in the North, Northeast, Central, and Southern regions. As indicated in the map legend, existing inter-provincial backbone lines are shown as solid red lines, lines under construction as dashed orange lines, and lines planned for deployment in 2010 as purple lines; blue, green, and cyan lines denote VTN-managed, EVN, and Viettel fiber-optic lines, respectively. VTN nodes (telecommunication stations) appear as circled dots and line branching points as small red dots. Our instrument relies on the locations of backbone nodes that were operational prior to the 2007 capacity upgrade; the planned and under-construction segments shown in the map play no role in constructing the instrument.

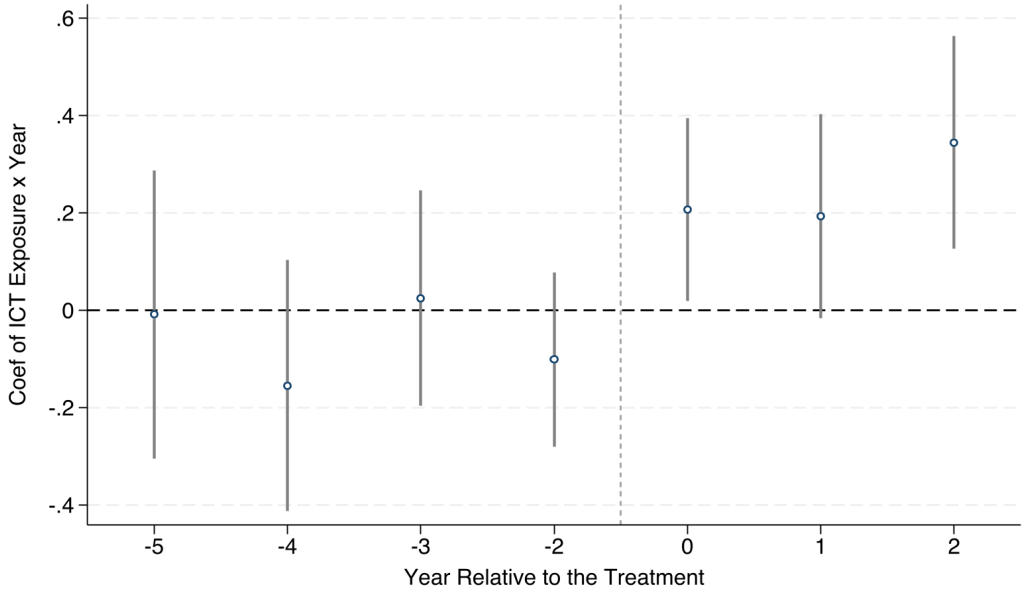
Figure 2: Female Labor Share and Firm ICT Adoption



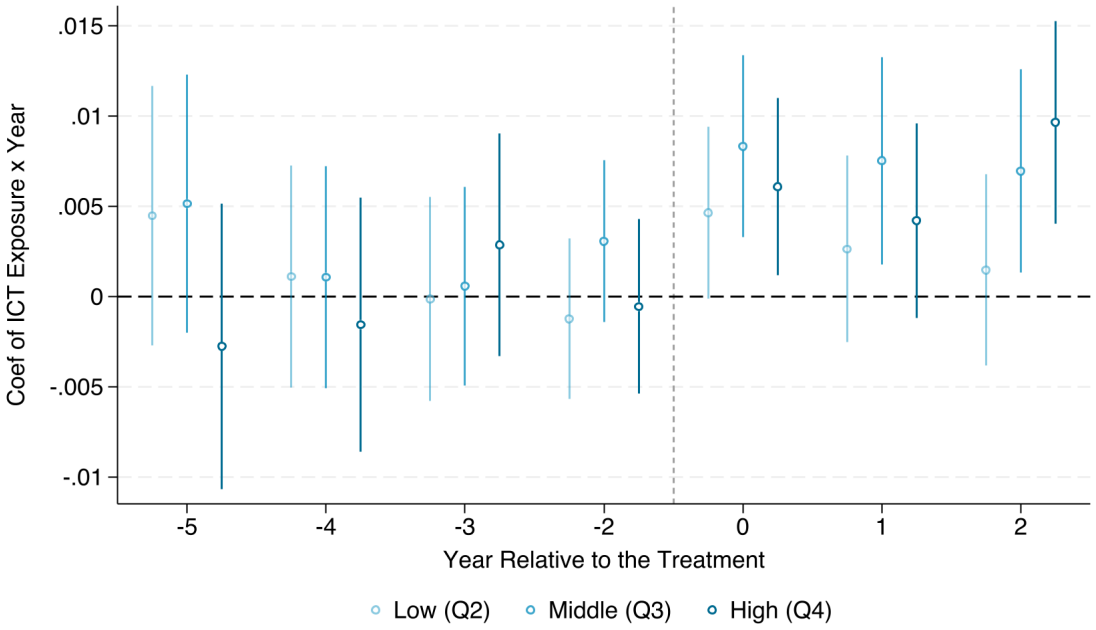
Notes: This figure presents binned scatter plots of firms' female labor share against their eventual (post-2007) ICT adoption, separately for the pre-expansion years (2004–2006, left panel) and the post-expansion years (2007–2009, right panel). ICT adoption is measured as the number of internet-connected PCs per worker and is plotted on the $\log(1+x)$ scale for readability. We repeat the same exercise for the LAN measure in Appendix Figure A.2.

Figure 3: Pre-trend Using Predicted ICT Exposure

(A) CONTINUOUS TREATMENT



(B) BINNED TREATMENT



Notes: This figure presents event-study estimates from specification (4). Panel A plots the estimated coefficients on the continuous treatment intensity T_i, denoted by β_{1t}, along with 95% confidence intervals. The estimation window is 2001–2009, extending the pre-treatment period back to 2001 (see Section 5.3); 2005 is the reference year, and 2006 is excluded because firm-level ICT adoption is not collected in that year. Panel B reports results using a discretized measure of treatment intensity, where firms are grouped into quartiles based on T_i. The darkest markers correspond to high exposure (T_i^h), the medium shade to intermediate exposure (T_i^m), and the lightest shade to low exposure (T_i^l).

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 presents cross-country comparisons of broadband Internet quality, measured by average connection speed and affordability. Internet speed is expressed as bits per second per user, while affordability is captured by the average monthly subscription cost (in U.S. dollars). The data are drawn from the World Bank's World Development Indicators (accessed October 10, 2017).

Table A.2: VIETNAM ANNUAL WAGES FOR WAGE WORKERS BY GENDER AND EDUCATION (THOUSAND VND)

	2004			2012		
	Male	Female	Ratio	Male	Female	Ratio
No Education	4,475	3,203	1.40	21,470	14,890	1.44
Primary	6,198	4,743	1.31	26,990	22,284	1.21
Secondary	7,197	5,909	1.22	31,054	25,835	1.20
High School	11,520	11,043	1.04	44,399	39,486	1.12
College and Above	21,781	17,459	1.25	87,322	66,336	1.32

Notes: This table reports average annual wages (in thousand VND) for wage workers in Vietnam, by gender and educational attainment, for 2004 and 2012, using data from the Vietnam Household Living Standards Survey. The "Ratio" column presents the male-to-female wage ratio within each education category. Statistics are based on individuals aged 15–65.

Table A.3: VIETNAM LABOR FORCE STATISTICS BY GENDER

	2004		2012	
	Male	**Female**	**Male**	**Female**
Employment status of those in labor force				
Work for Salary	42.8	26.1	46.9	30.1
Formal Sector	20.8	15.5	25.3	21.6
Private Company (Main Job)	4.6	2.9	10.0	7.4
State-owned enterprises (Main Job)	11.9	9.3	10.6	9.4
Education of those in labor force				
No Education	18.4	25.8	15.9	21.5
Primary	28.9	27.5	25.9	25.9
Secondary	32.4	29.6	31.2	28.8
High School	16.3	14.7	20.4	18.6
College and Above	3.9	2.3	6.5	5.0

Notes: This table reports labor force statistics in Vietnam by gender for 2004 and 2012, using data from the Vietnam Household Living Standards Survey. It presents employment status and educational attainment for individuals aged 15–65. All values are expressed in percentages.

Table A.4: Summary Statistics on ICT Variables, Firm Level

	2004	2005	2007	2008	2009
Number of PC with Internet per Worker	0.05	0.07	0.11	0.13	0.15
	(0.14)	(0.18)	(0.21)	(0.22)	(0.24)
Number of PC with LAN per Worker	0.05	0.07	0.10	0.12	0.12
	(0.16)	(0.19)	(0.22)	(0.23)	(0.24)
Observations	37,727	46,996	51,619	46,599	51,074

Notes: This table reports summary statistics for the ICT adoption measures used in our analysis, disaggregated by year. The data are drawn from the Vietnam Enterprise Survey (2004–2009). Column headers indicate the survey year; each cell reports the mean, with standard deviations in parentheses.

Table A.5: Geographic Characteristics List

Variable Name	Description	Source	Resolution
Hydrology:			
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:			
bdod	Bulk density of soil.	SoilGrids	250m
cfvo	Coarse fragments in the soil.	SoilGrids	250m
sand	Percentage of sand in the soil.	SoilGrids	250m
Terrain:			
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 summarizes the definitions, data sources, and spatial resolution of the hydrological, soil, and terrain variables used in the LASSO selection procedure.

Table A.6: ROBUSTNESS: EXTENDING THE ESTIMATION WINDOW BACK TO 2001

	(1) Baseline (2004–2009)	(2) Extended (2001–2009)
# of PCs w/ internet per worker	0.305***	0.276***
	(0.103)	(0.079)
N	180,025	237,653
Kleibergen-Paap Wald Rank F Stat	14.489	14.837
AR Wald Test (Weak IV Robust Inf.)	18.940	19.037
p-Value	0.002	0.002
# of PCs w/ LAN per worker	0.522***	0.405***
	(0.155)	(0.114)
N	175,838	233,475
Kleibergen-Paap Wald Rank F Stat	10.210	15.450
AR Wald Test (Weak IV Robust Inf.)	17.533	17.261
p-Value	0.000	0.000

Notes: This table re-estimates our baseline 2SLS specification (column 4 of Table 4) on an extended panel. Column 1 uses the headline 2004–2009 window; column 2 extends the panel back to 2001 by carrying each firm's predetermined, time-invariant node and geographic assignment onto its earlier records. All specifications include firm and node-by-year fixed effects, the full set of firm-level and locational controls, and baseline commune trade-exposure controls, and use the same LASSO-selected instruments as in the main table. Robust standard errors clustered at the commune level are in parentheses. */**/*** denote statistical significance at the 10%/5%/1% levels, respectively.

Table A.7: MEDIUM-TERM EFFECTS

Panel A: Generalized DiD		
	2004-2009 (1)	**2004-2011 (2)**
Treatment Intensity (T_i) $\times$ $Post07_t$	0.010***	0.011***
	(0.003)	(0.003)
N	180,025	319,071
Panel B: Two Sample 2SLS (Internet)		
	2004-2009 (1)	**2004-2011 (2)**
# of PCs w/ Internet per Worker (level)	0.190***	0.236***
	(0.053)	(0.062)
N	429,590	566,341
Panel C: Two Sample 2SLS (LAN)		
	2004-2009 (1)	**2004-2011 (2)**
# of PCs w/ LAN per Worker (level)	0.205***	0.251***
	(0.068)	(0.082)
N	425,524	559,127

Notes: This table reports the medium-term effects of ICT adoption on female labor share. Panel A presents estimates from a generalized DiD specification, as defined in equation (5). Panels B and C report two-sample 2SLS estimates: Panel B uses the number of internet-connected PCs per worker as the measure of ICT adoption, while Panel C uses the number of LAN-connected PCs per worker. Column 1 uses the 2004–2009 sample, while Column 2 uses the extended 2004–2011 sample. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A.8: ICT AND FEMALE LABOR DEMAND, CONNECTIVITY BUFFER

	Baseline		Leave out Immediate Neighbors		Smaller Buffer Size (0.5 km)	
	(1)	(2)	(3)	(4)	(5)	(6)
# of PCs w/ internet per worker	0.305***		0.463***		0.189*	
	(0.103)		(0.136)		(0.100)	
# of PCs w/ LAN per worker		0.522***		0.508***		0.282**
		(0.155)		(0.149)		(0.126)
N	180,025	175,838	180,025	175,838	180,025	175,838
Kleibergen-Paap Wald Rank *F* Stat	14.489	10.210	10.825	8.064	22.707	14.802
AR Wald Test (Weak IV Robust Inf.)	18.940	17.533	18.169	18.550	10.069	10.814
p-Value	0.002	0.000	0.000	0.000	0.073	0.055

Notes: This table presents estimates of the effect of firms' ICT adoption on female labor share under alternative definitions of the *connectivity buffer* used to construct the instrumental variables. Columns 1 and 2 report the baseline results. 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 buffer in instrument construction. All regressions are estimated using a 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 commune level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.9: ICT AND FEMALE LABOR DEMAND, VARYING LOCATION RESTRICTIONS

	Distance: 20 km		Distance: 25 km		Distance: 15 km	
	(1)	(2)	(3)	(4)	(5)	(6)
# of PCs w/ internet per worker	0.305***		0.261**		0.360***	
	(0.103)		(0.107)		(0.107)	
# of PCs w/ LAN per worker		0.522***		0.396**		0.571***
		(0.155)		(0.159)		(0.156)
N	180,025	175,838	190,344	185,805	158,043	154,405
Kleibergen-Paap Wald Rank *F* Stat	14.489	10.210	12.555	9.309	11.695	9.919
AR Wald Test (Weak IV Robust Inf.)	18.940	17.533	9.680	7.862	26.112	23.714
p-Value	0.002	0.000	0.085	0.020	0.000	0.000

Notes: This table presents estimates of the effect of firms' ICT adoption on female labor share under alternative location-based sample restrictions. Columns 1 and 2 report the baseline results, where the sample is restricted to firms located within 20 km of the nearest backbone node. Columns 3 and 4 expand the sample to firms within 25 km, while Columns 5 and 6 restrict the sample to firms within 15 km. All regressions use the same set of instruments as reported in Column (1) of Table 3. Robust standard errors are clustered at the commune level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.10: Firm Female Labor Share and ICT Adoption, Different Ways of Clustering

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of PCs w/ internet per worker	0.305***		0.305***		0.305***		0.305***	
	(0.103)		(0.098)		(0.099)		(0.098)	
# of PCs w/ LAN per worker		0.522***		0.522***		0.522***		0.522***
		(0.155)		(0.166)		(0.135)		(0.135)
N	180,025	175,838	180,025	175,838	180,025	175,838	180,025	175,838
Kleibergen-Paap Wald Rank *F* Stat	14.489	10.210	8.534	3.155	19.716	9.781	19.076	14.371
AR Wald Test (Weak IV Robust Inf.)	18.940	17.533	17.889	15.514	36.014	31.137	41.578	36.098
p-Value	0.002	0.000	0.003	0.000	0.000	0.000	0.000	0.000
Cluster Level:								
Commune	Yes	Yes						
District			Yes	Yes				
Nearest Node					Yes	Yes		
Province							Yes	Yes

Notes: This table presents estimates of the effect of firms' ICT adoption on female labor share under alternative clustering methods for standard errors. Columns 1 and 2 report the baseline results, with standard errors clustered at the commune level. Columns 3-8 implement alternative clustering approaches, as indicated in the bottom panel of the table. All regressions use the same set of instruments as reported in Column 1 of Table 3. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.11: Robustness: Additional Fixed Effects and Sample Restrictions

	(1) Baseline	(2) + Province×Year FE	(3) Within-Province Nodes	(4) Positive Adopters	(5) Winsorized
# of PCs w/ internet per worker	0.305***	0.261**	0.279**	0.360***	0.325***
	(0.103)	(0.133)	(0.127)	(0.130)	(0.112)
N	180,025	180,025	163,012	115,263	178,677
Kleibergen-Paap Wald Rank *F* Stat	14.489	10.859	11.540	5.839	13.845
AR Wald Test (Weak IV Robust Inf.)	18.940	12.082	11.631	17.121	18.283
p-Value	0.002	0.034	0.040	0.004	0.003

Notes: This table assesses the robustness of our baseline estimate of the effect of ICT adoption on the female labor share to additional fixed effects and alternative sample restrictions. All columns start from our headline specification (column 4 of Table 4), holding the LASSO-selected instrument set fixed and including the full firm-level and locational controls, firm and node×year fixed effects, and baseline commune trade-exposure controls. Column 1 reproduces the baseline. Column 2 adds province×year fixed effects. Column 3 restricts the sample to firms whose nearest backbone node lies in their own province (about 92% of firms). Column 4 restricts the sample to firms with positive ICT adoption, and column 5 drops the top 1% of the ICT distribution. Robust standard errors clustered at the commune level are in parentheses. */**/*** denote significance at the 10%/5%/1% levels.

Table A.12: MAIN RESULTS, USING DIFFERENT IVS

	Baseline		Dist. IV	
	(1)	**(2)**	**(3)**	**(4)**
# of PCs w/ internet per worker	0.305***		0.189**	
	(0.103)		(0.077)	
# of PCs w/ LAN per worker		0.522***		0.242**
		(0.155)		(0.097)
N	180,025	175,838	180,025	175,838
Kleibergen-Paap Wald Rank F Stat	14.489	10.210	16.507	15.672
AR Wald Test (Weak IV Robust Inf.)	18.940	17.533	5.849	6.030
p-Value	0.002	0.000	0.016	0.014

Notes: This table presents estimates of the effect of firms' ICT adoption on female labor share using alternative instrumental variables. Columns 1 and 2 report the baseline results. Columns 3 and 4 use an alternative instrument constructed as the interaction between distance to the nearest backbone node and a post-2007 indicator. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.13: Individual Labor Market Outcomes and ICT by Gender

	Self-employed (1 0)		Work for Salary (1 0)		Income (log)		Second Job (1 0)		Working Days (log)		Housework (1 0)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment Intensity × Post 2007	0.007	0.005	-0.009	-0.009	-0.016	-0.014	-0.023*	-0.021*	-0.006	-0.008	-0.021**	-0.022*
	(0.010)	(0.010)	(0.009)	(0.010)	(0.018)	(0.018)	(0.012)	(0.012)	(0.008)	(0.008)	(0.010)	(0.011)
Treatment Intensity × Post 2007 × Female		0.005		-0.001		-0.003		-0.004		0.005		0.000
		(0.005)		(0.004)		(0.003)		(0.004)		(0.003)		(0.007)
N	242,374	242,374	242,374	242,374	242,302	242,302	202,972	202,972	194,162	194,162	152,921	152,921

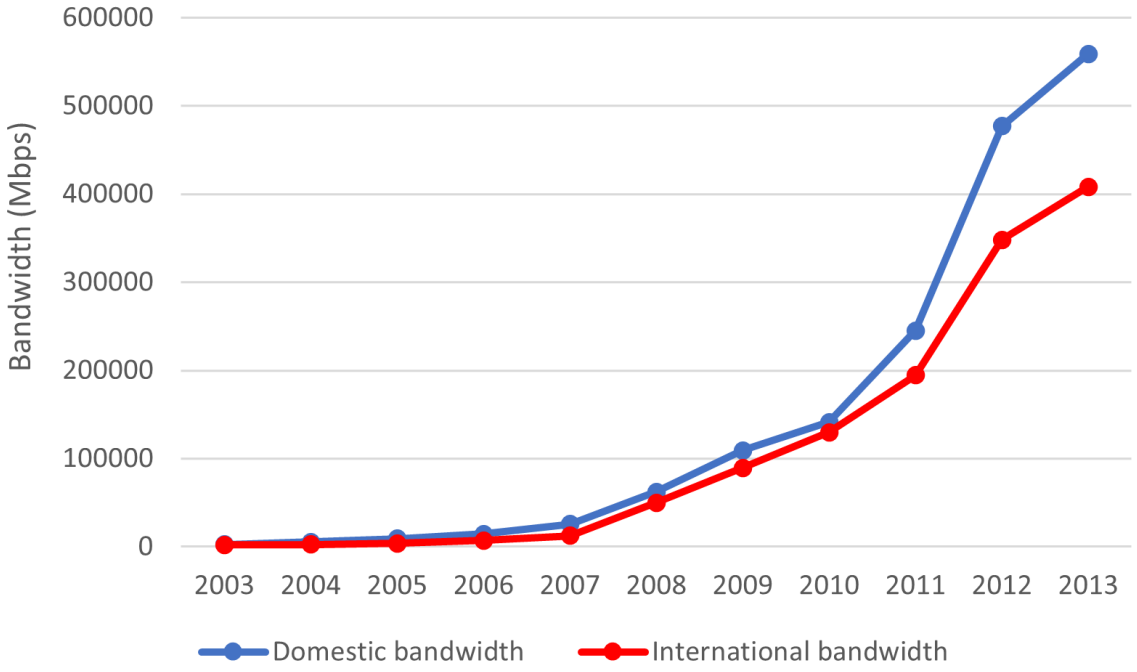
Notes: This table presents the effects of firm-level ICT adoption on other individual labor market outcomes. Each cell reports the effect for men and the female differential (the two treatment-interaction coefficients from equation (6)), with dependent variables indicated in the column headings. The sample is limited to individuals located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.14: HETEROGENEITY BY FIRM SIZE AND SECTOR

	(1)	(2)	(3)
# of PCs w/ internet per worker	0.305***	0.108**	0.310***
	(0.103)	(0.054)	(0.108)
... X Size: Medium		0.330***	
		(0.043)	
... X Size: Large		0.235***	
		(0.072)	
... X Sector: Manufacturing			0.279***
			(0.086)
N	180,025	180,025	180,025
Kleibergen-Paap Wald Rank F Stat	14.489	5.521	5.535
AR Wald Test (Weak IV Robust Inf.)	18.940	209.625	27.114
p-Value	0.002	0.000	0.007

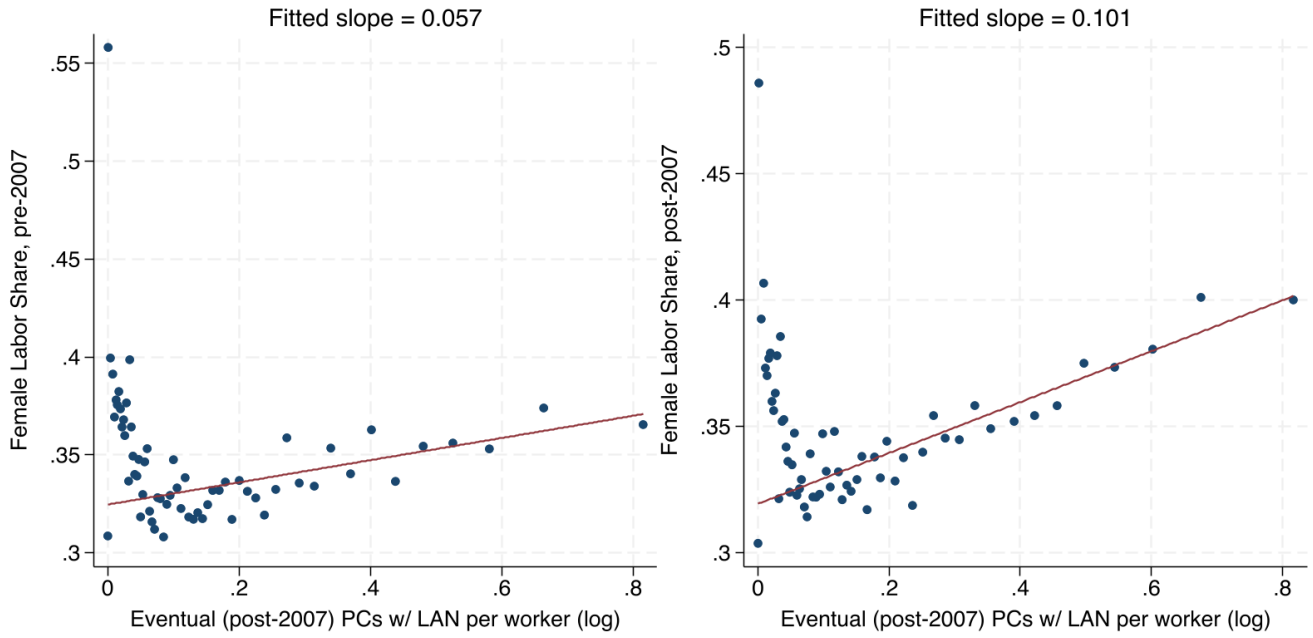
Notes: This table presents heterogeneous effects of ICT adoption on firms' female labor share across firm size and sector. Each cell reports the estimated coefficient on ICT adoption and its interaction with firm size or sector, using the same set of controls as in the baseline specification (1). All specifications use the same set of instruments, along with their corresponding interaction terms, as reported in Column 1 of Table 3. The sample is restricted to firms located within 20 km of the nearest backbone node. Robust standard errors are clustered at the commune level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure A.1: The Expansion of Internet Bandwidth in Vietnam, 2003–2013



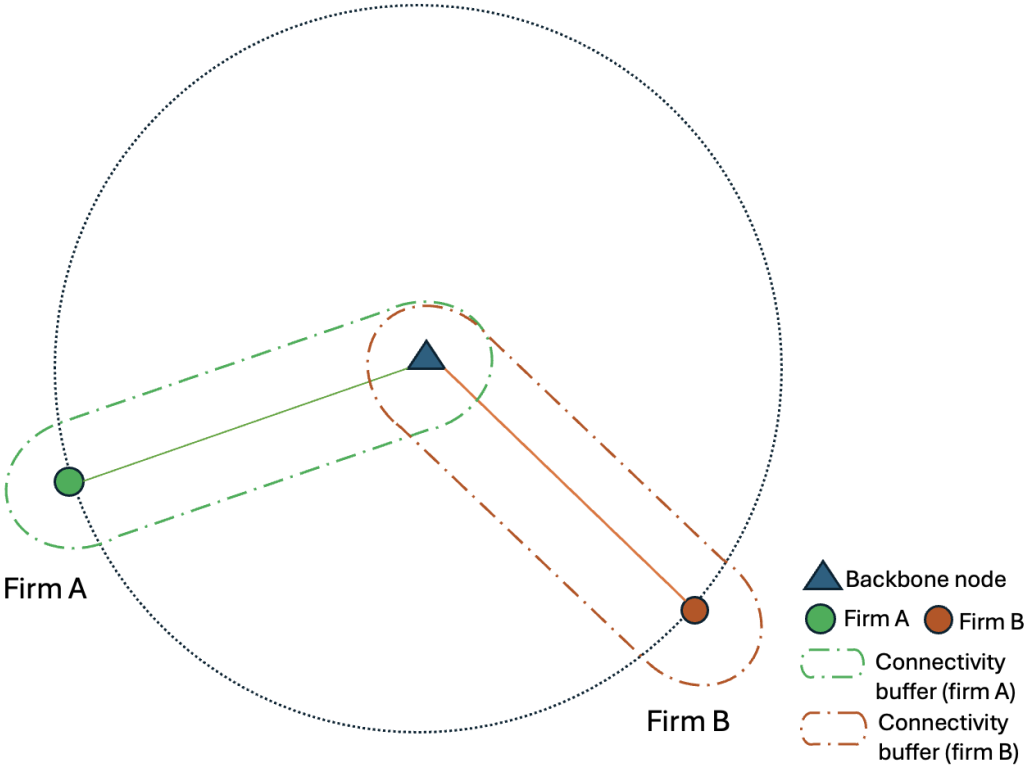
Notes: This figure plots the evolution of domestic and international bandwidth in Vietnam from 2003 to 2013, using data from the Vietnam Internet Network Information Center (Vietnam Internet Network Information Center, 2013). The red line corresponds to international bandwidth, while the blue line represents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

Figure A.2: Female Labor Share and Firm ICT Adoption (LAN measure), by Period



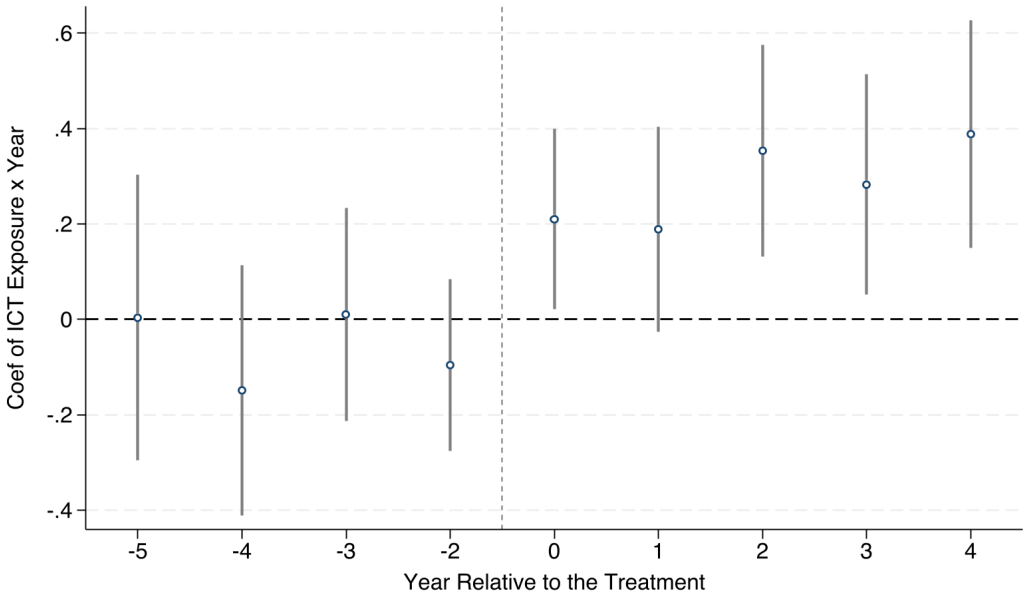
Notes: This figure repeats Figure 2 using LAN-connected PCs per worker as the ICT measure. It presents binned scatter plots of firms' female labor share against their eventual (post-2007) LAN adoption, separately for the pre-expansion years (2004–2006, left panel) and the post-expansion years (2007–2009, right panel); the fitted slope is reported above each panel. ICT adoption is plotted on the $\log(1+x)$ scale for readability.

Figure A.3: Construction of the Geographic Variation in the Instrument



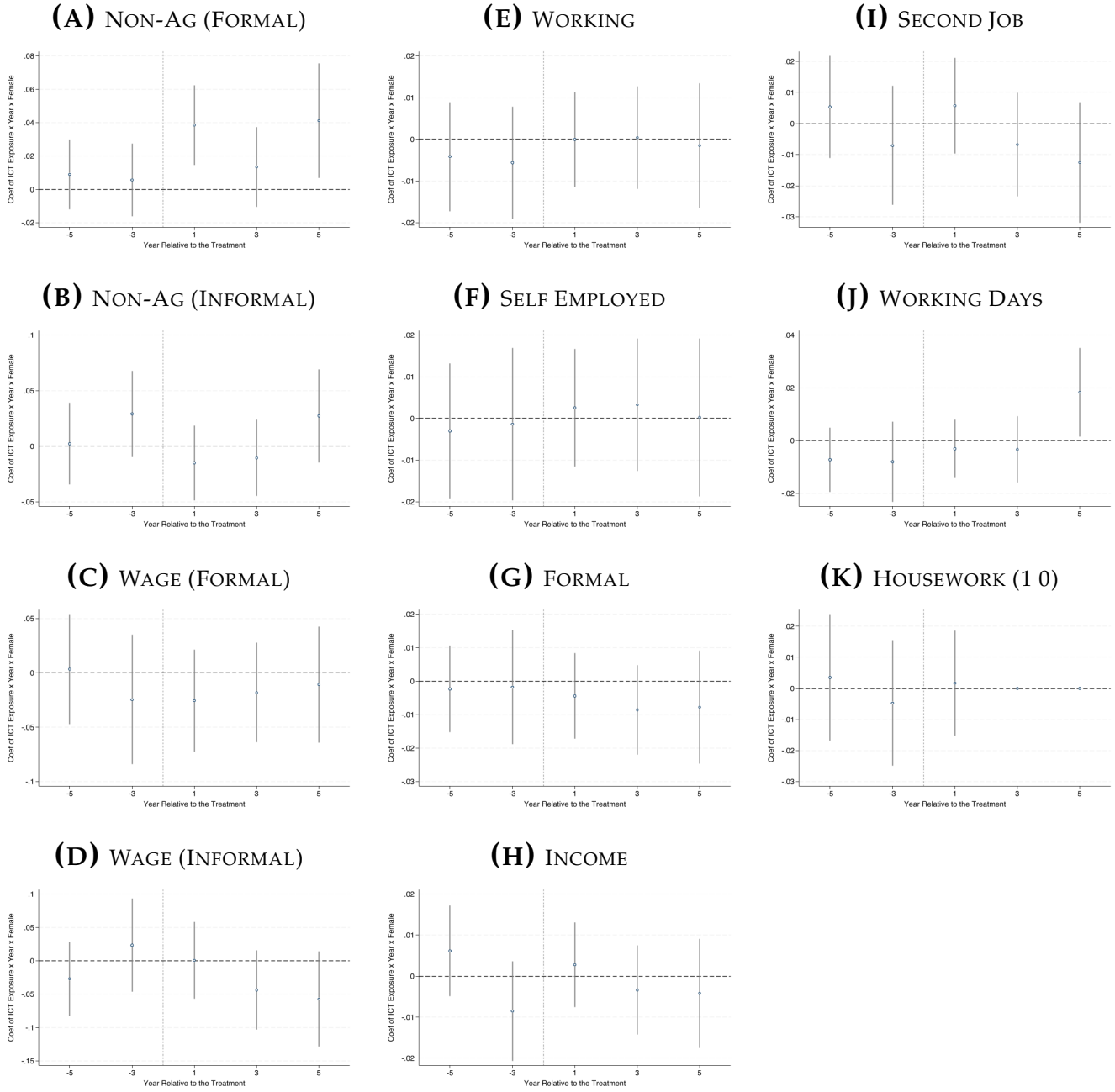
Notes: This figure provides a schematic illustration of the geographic variation underlying our identification strategy. The blue triangle denotes a backbone node. Firm A (green circle) and Firm B (orange circle) both identify this node as their nearest backbone connection and are located at a similar distance from it (illustrated by the large black dotted circle). Solid lines represent the straight-line paths connecting each firm to the node. Around each path, we construct a connectivity buffer (dash-dotted corridors), within which geographic characteristics are aggregated. Although the two firms are equidistant from the same node, differences in geographic conditions within their respective buffers imply heterogeneous connection costs, generating differential increases in ICT adoption following the post-2007 broadband expansion. This variation forms the basis of our identification strategy.

Figure A.4: Pre-trend Using Predicted ICT Exposure



Notes: This figure presents event-study estimates from specification (4) using the extended sample from 2001–2011. It plots the estimated coefficients on the continuous treatment intensity T_i, denoted by β_{1t}, along with 95% confidence intervals.

Figure A.5: Individual Level Analysis, Event Study



Notes: This figure presents event-study estimates based on the event-study version of specification (6), using VHLSS data. The baseline interaction term, $\beta_1(T_i \times Post07_t \times Female_i)$, is replaced with a full set of year-specific interactions, $(T_i \times \mathbb{1}\{Year_t = \tau\} \times Female_i)$. The figure plots the resulting year-specific coefficients β_{1t}, along with 95% confidence intervals. The corresponding dependent variables are indicated in the panel titles. "Formal" is an indicator for holding a formal-sector job, defined as employment in a private firm, state-owned enterprise, public institution, or foreign-invested firm. "Non-Ag (Formal)" further restricts this to *non-agricultural* formal employment, classifying jobs as agricultural or non-agricultural by their industry code; it is therefore the non-agricultural subset of the "Formal" outcome.

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.

Table B.1 reports summary statistics of firm-level task intensity in the Vietnam Enterprise Survey (2004–2009). Tables B.2 and B.3 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 wood products, 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.1: SUMMARY STATISTICS OF TASK INTENSITY, VIETNAM, 2004-2009

Task Type	Mean	Std. dev.	Min	Max
Routine	4.56	1.14	1.77	7.18
Manual	1.77	0.76	0.59	4.11
Social	3.82	1.96	1.66	8.10
Customer	3.63	2.32	1.02	8.55
Service	2.18	0.99	0.65	4.95

Notes: This table reports summary statistics of firm-level task intensity in the Vietnam Enterprise Survey (2004–2009), based on O*NET occupational task measures.

Table B.2: SECTORS BY DIFFERENT TASK INTENSITY, PART I

Rank	Sectors (Code)	Task Intensity
Routine		
1	Manufacture of Tobacco Products (120)	7.18
2	Manufacture of Man-Made Fibres (203)	6.70
3	Services to Buildings and Landscape Activities (81)	6.67
4	Manufacture of Footwear (152)	6.66
5	Scientific Research and Development (72)	6.61
...		
145	Retail Sale in Non-Specialized Stores (471)	2.22
146	Retail Sale of Food, Beverages and Tobacco in Specialized Stores (472)	2.16
147	Wholesale on a Fee or Contract Basis (461)	1.77
148	Manufacture of Paper and Paper Products (17)	0.50
149	Activities of Other Membership Organizations (949)	0.50
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 in Non-Specialized Stores (471)	8.10
2	Retail Sale of Food, Beverages and Tobacco in Specialized Stores (472)	7.93
3	Steam and Air Conditioning Supply (353)	7.78
4	Retail Sale of Cultural and Recreation Goods in Specialized Stores (476)	7.74
5	Retail Sale of Information and Communications Equipment in Specialized Stores (474)	7.74
...		
145	Sawmilling and Planing of Wood (161)	1.70
146	Manufacture of Domestic Appliances (275)	1.66
147	Materials Recovery (383)	1.62
148	Manufacture of Paper and Paper Products (17)	0.62
149	Activities of Other Membership Organizations (949)	0.62

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.3: 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.

C Empirical Methods Appendix

This section provides details on the estimation of ICT's effects on firm entry and exit.

For exit, we construct a firm-level indicator equal to one in the last year a firm appears in the census and zero otherwise. Thus, the exit variable takes the value 1 if a firm exits in the subsequent period and 0 otherwise. We exclude the final sample year to avoid mechanically classifying the end of the panel as an exit. We then estimate a discrete-time hazard model to assess the effect of ICT on firm exit. The probability of firm exit in any period is modeled as a function of elapsed duration of firm survival, τ, the firm's ICT measure ICT_{it}, and other firm-level and regional controls consistent with our main specification (1). Each firm spell is represented as a sequence of (0,1) observations. The model is specified as follows:

$$\begin{aligned} P\big(exit_{ij} = \tau + 1 \mid t_{ij} > \tau,\, \text{ICT}_{it}, \gamma_{c(i)}, \theta_{nt}, X_{it}, \text{industry}_i, \text{age}_{it}, \mathbf{Z}'_{it}\big) \\ = g\big(\tau,\, \text{ICT}_{it}, \gamma_{c(i)}, \theta_{nt}, X_{it}, \text{industry}_i, \text{age}_{it}, \mathbf{Z}'_{it}\big)\,. \end{aligned} \tag{8}$$

where $P(exit_{ij})$ denotes the probability that firm i exits in period j. ICT_{it} captures the firm's ICT adoption. We include a full set of duration dummies τ to flexibly model duration dependence. We use the same set of IVs as in column 1 of Table 3 to instrument for ICT_{it}.

Depending on the hazard specification, we additionally control for commune fixed effects $\gamma_{c(i)}$ and node-by-year fixed effects θ_{nt}.[17] In addition to the firm-level controls included in our main specification, X_{it}, we further include industry dummies and firm-age dummies. We also control for the same set of locational controls as in the main specification, denoted by $\mathbf{Z}'_{it}$. We use a linear probability model to avoid incidental parameter problems.

We apply a different empirical approach in examining firm entry. While firms are observed for multiple years after entry, they are not observed prior to entry, and we do not observe non-entering "potential" firms. Consequently, we cannot construct a firm-level panel with pre-entry ICT exposure. Instead, we analyze entry at the commune–year level. We compute a commune–year entry rate defined as the number of new firms that first appear in a given commune and year divided by the total number of active firms in that commune–year. Using this measure, we estimate IV regressions of entry rates on commune–year ICT exposure, employing the same geographic-based instruments and controlling for commune fixed effects, node-by-year year fixed effects, and the same set of firm-level controls aggregated to the commune level.

[17] We don't use firm fixed effects in most firm exit models because the outcome (exit) is typically a one-time absorbing event per firm, so there is almost no within-firm variation over time to identify the effect.