H. Wicaksono, K. Rzazade, A. Krishnakumar, S. Ganduri (Group 2)

Econometrics II - Spring 2025

Who Benefits from Digitalization?

Task-Biased Technological Change and Gendered Labor Market Polarization:

Evidence from a Province–Industry Panel in Indonesia (2020–2024)

Abstract

This paper examines how digital competitiveness affects labor task composition across Indonesian provinces from 2020 to 2024. Using a task-based framework, we estimate fixed effects regressions linking the Digital Competitiveness Index to the log share of nonroutine versus routine employment. Results reveal that digital advancement increases nonroutine work, but the effects vary by industry tech-intensity and gender composition. Medium-tech industries and female-dominated sectors see weaker gains, suggesting conditional skill-biased technological change. Our findings underscore the need for targeted, inclusive digital skilling policies that account for regional, industrial, and demographic asymmetries in labor market adaptation.

1. Introduction

The question of how technological advancement reshapes labor market outcomes has garnered significant attention over the past two decades. However, most of this work has focused on developed countries, where labor markets are formal, education levels are high, and digital infrastructure is strong. These studies often conclude that digital technologies complement high-skill labor and displace workers performing routine tasks (Autor et al., 2003; Acemoglu & Autor, 2011; Goos & Manning, 2007). Yet, the relevance of these findings to developing economies like Indonesia, where employment is often informal, digital adoption is uneven, and regional disparities are wide, remains underexplored.

Group 2

Econometrics II - Spring 2025

This paper examines how technological advancement affects labor task composition in Indonesia, one of Asia's largest developing economies. Rather than using broad categories like "skilled" and "unskilled," we adopt a task-based framework proposed by Acemoglu and Autor (2011), which distinguishes between nonroutine and routine tasks, and further between cognitive and manual tasks. This approach centers on the nature of work performed, specifically, whether a task is codifiable and susceptible to automation, or inherently flexible and complementable by technology.

We focus on Indonesia for three reasons. First, the country has undergone rapid digitalization in recent years, including through national initiatives such as the Digital Roadmap 2021–2024, with digital competitiveness measured by the East Ventures Digital Competitiveness Index (EV-DCI). Second, the Indonesian labor market displays high spatial and sectoral heterogeneity, making it a compelling setting to study local effects of digital change. Third, despite increasing attention to digital labor impacts globally, very few studies have examined these dynamics in developing Asian contexts at the subnational level.

We analyze how digital competitiveness affects the share of workers in nonroutine versus routine tasks across Indonesian provinces and industries over the 2020–2024 period. Additionally, we investigate how these effects differ by gender and by the technological intensity of the industry. This allows us to capture both demographic and structural dimensions of heterogeneity in Indonesia's digital transition.

2. Literature Review

The task-based approach introduced by Acemoglu and Autor (2011) provides a foundational framework for understanding how technology reshapes labor markets. This model classifies tasks into two main dimensions: routine vs. nonroutine. Routine tasks, whether cognitive (e.g., bookkeeping) or manual (e.g., assembly line work), are repetitive and easily automated. In contrast, nonroutine tasks, both cognitive (e.g., problem-solving, analysis) and manual (e.g., caregiving, maintenance), are less susceptible to automation and may even be augmented by technology.

Empirical studies in developed countries offer strong support for this framework. Autor, Levy, and Murnane (2003) find that "computer capital substitutes for workers doing tasks following explicit rules ('routine tasks') and complements those performing problem-solving and complex communication tasks". Goos and Manning (2007), studying the U.K., report employment growth in both high-paid nonroutine cognitive and low-paid nonroutine manual jobs, alongside a decline in middle-wage routine jobs, a pattern referred to as job polarization.

Building on this theory, Acemoglu and Autor (2011) argue that the impact of technology depends not just on workers' education levels but on the task content of their jobs. This shift in analytical lens allows researchers to distinguish, for example, between low-education workers in routine manufacturing and those in nonroutine personal care, both of whom face different risks and opportunities from technological change.

Applying this logic to a developing context, Almeida, Poole, and Corseuil (2025) analyze the impact of internet expansion in Brazil using task classifications similar to those of Acemoglu and Autor. They find that digital adoption significantly increases the share of nonroutine cognitive tasks, particularly in high-tech industries. This supports the hypothesis that digital technologies reallocate employment toward nonroutine roles, but also highlights the role of industry-level technological intensity in shaping outcomes.

While Almeida et al. offer valuable evidence from Latin America, much of the task-based empirical literature remains focused on high-income countries or firm-level data, with limited attention to subnational variation. In Indonesia, where regional disparities in digital infrastructure and labor market composition are substantial, a province-industry-year panel structure enables us to test the broader applicability of the task-based framework.

Recent literature has also examined heterogeneity in technology's impact across sectors and gender. While some studies emphasize that high-tech sectors tend to benefit more from digital transformation (Almeida et al., 2025), others note that even low-tech industries can gain indirectly, particularly via innovation and organizational upgrades (Blichfeldt & Faullant, 2021). Similarly, research on gendered effects of digitization is mixed. Wu et al. (2024) find that digitization increases female employment and labor share, particularly in services. In contrast, Aum and Shin (2025) report that digital adoption in Korea may worsen outcomes for female workers, particularly in sectors where tasks are easily automated.

Taken together, these studies highlight the importance of examining heterogeneity across industry and demographic groups. Our study builds on this literature by applying the task-based approach in a developing Asian setting, while explicitly accounting for gender and technological intensity. In doing so, we provide new insights into the distributional consequences of digital competitiveness in Indonesia's labor market.

3. Data Sources

Our main dataset is the Indonesian National Labor Force Survey (Indonesian LFS for short), an individual-level survey conducted twice a year (every February and August) since 1976 by Statistics Indonesia. The survey aims to capture the current state of the labor force and employment conditions in Indonesia. Although the Indonesian LFS is based on a sample of over 750,000 individuals each year, it provides population weights that allow us to produce representative population-level estimates. Indonesian LFS includes information on the respondent's province of residence and the industry in which they work, enabling aggregation to the province–industry level using the provided weights. It also records occupational codes, which are key to determining job tasks. Detailed information on the Indonesian LFS is provided in Appendix A

First, we exclude public sector workers and restrict the sample to formal sector workers, as prior studies suggest that these groups are less susceptible to technological substitution. Then, following the frameworks of

Acemoglu and Autor (2011) and Almeida, Poole, and Corseuil (2025), we classify job tasks based on a combination of occupation and industry codes. We then combine the cleaned yearly datasets and construct a province–industry panel dataset by aggregating individual-level records using sampling weights. To ensure accuracy, we verify that the calculated number of formal workers aligns with official statistics published by Statistics Indonesia. Detailed data cleaning procedures are outlined in Appendix B. Our cleaned Indonesian LFS dataset covers 34 provinces, 16 industries, and 5 years, yielding a theoretical maximum of 2,720 observations (34 × 16 × 5; see Appendix B). However, due to missing industry data in certain provinces, the actual number of observations in the final dataset is 2,696.

To measure technological advancement, we use the Digital Competitiveness Index (DCI) published by East Ventures. The index reflects regional digital competitiveness across dimensions such as digital infrastructure, human capital, digital economic activity, and local government policy. However, it is only available from 2019 onward. Further details on the construction of the DCI are provided in Appendix C. For additional control variables—including the province-level GDP, and the Gini coefficient—we draw data from Statistics Indonesia.

We merge the cleaned Indonesian LFS dataset with the DCI and relevant control variables to construct our analytical dataset. Definitions of all variables are provided in Appendix D.

4. Methodology

Almeida, Corseuil, and Poole's (2025) study on Brazil from 1999-2006 reveals that digital technology adoption favors skilled workers by increasing demand for cognitive tasks, a finding highly relevant to our research on Indonesia from 2020-2024. Their key insight—that strict labor regulations can unintentionally amplify this skill-based disparity—provides a valuable lens through which to examine Indonesia's labor market, particularly considering the nation's varying levels of technological adoption and labor policy enforcement during our study period.

Following Acemoglu and Autor (2011), Almeida, Poole, and Corseuil (2025), and adapting their approach to the available data in Indonesia, we specify the following econometric model to test our hypothesis:

$$Y_{pkt} = \beta_1 \cdot DCI_{pt-1} + \beta_2(DCI_{pt-1} \cdot Industry\ tech_k) + \beta_3 \cdot Share\ Female_{pkt} + \beta_4\,(DCI_{pt-1} \cdot Share\ Female_{pkt}) + Z_{pt}\theta + \varphi_p + \delta_{kt} + \epsilon_{pkt}$$

Where:

- Y_{pkt} represents the share of nonroutine workers relative to routine workers in province p, industry k, and year t, in log terms
- DCI_{pt-1} is the measure of technological advancement in province p and year t-1

  We use the East Venture Digital Competitiveness Index (EV-DCI) as a proxy of this
- $Industry\ tech_{pk}$ is the industry's technological intensity in industry k

  Industries are classified into high-tech, medium-tech, and low-tech categories
- $Share\ Female_{pkt}$ is the share of individuals aged over 15 who are female in province p, industry k, and year t
- θ includes province-level control variables such as the share of tertiary-educated population, GDP, the share of urban population, and the Gini coefficient

Industry classification is based on technological intensity, following frameworks developed by the OECD (2011) and Eurostat–OECD (2005). High-tech industries, such as Information and Communication and Financial Services, are highly knowledge- and innovation-driven. Medium-tech industries, like Manufacturing, Utilities, Education, and Health, exhibit moderate technological adoption. Low-tech industries, including Agriculture, Mining, and Retail, are traditionally more labor-intensive with lower digitalization levels.

To account for unobserved heterogeneity, we include two types of fixed effects in our regression model. First, province fixed effects (φ_p) control for time-invariant characteristics at the provincial level—such as geography, long-standing infrastructure, or cultural norms—that may influence labor market composition. Second, we include industry-year fixed effects (δ_{kt}) to capture shocks or trends that are specific to each industry in a given year, such as changes in global demand, sector-specific policies, or industry-wide technological adoption. Together, these fixed effects help isolate the impact of technological advancement from other confounding regional and industry-specific factors.

Our main coefficient of interest, β_1, captures the impact of technological advancement to the share of nonroutine relative to routine occupations. The coefficient β_2 measures whether this effect is more pronounced in high-tech sectors relative to low-tech ones. Building on the framework of Almeida, Poole, and Corseuil (2025), we extend the analysis by examining heterogeneous effects by gender, estimating the regression separately for female and male subsamples.

We hypothesize both β_1 and β_2 are positive, as technological adoption tends to substitute for routine tasks, thereby shifting employment composition toward nonroutine and cognitive occupations, especially in high-tech sectors. Furthermore, we expect these effects to be stronger for women, implying that both coefficients will be larger in the female subsample than in the male subsample. This would support the hypothesis that female workers are more vulnerable to technological displacement than their male counterparts.

5. Descriptive Statistics, Results and Robustness Checks

To empirically test these hypotheses, we now present the descriptive patterns and regression results from our analysis. We begin by examining trends in the share of nonroutine employment across industries and genders, followed by formal econometric estimates of the impact of digital competitiveness on employment composition. We then assess the robustness of these findings through a series of sensitivity analyses.

5.1 Descriptive Statistics

We begin by examining trends in nonroutine employment across industry tech-intensity levels and their association with the Digital Competitiveness Index (DCI). These descriptive patterns motivate the fixed effects regressions that follow, shedding light on how digital transformation intersects with industry structure, gender composition, and regional development.

Figure 5.1 (Appendix E) plots the evolution of nonroutine employment share across three industry tech-intensity categories: High-Tech, Medium-Tech, and Low-Tech. High-Tech sectors, such as Finance and ICT, maintain the highest share of nonroutine jobs across years, hovering around 58–60%, while Low-Tech sectors (e.g., agriculture, trade) display persistently low levels (~20–25%). Notably, Medium-Tech sectors show a sharp rise in 2024, overtaking High-Tech in average nonroutine share. This inversion suggests that technology-driven task restructuring is not confined to digitally advanced industries and could reflect technology diffusion or reclassification of job tasks across sectors.

Figure 5.2 shows scatterplots of Digital Competitiveness Index (DCI) against the nonroutine-to-routine employment share for each year. A weak but positive trend emerges over time, with the slope of the relationship between DCI and nonroutine share appearing to steepen slightly in later years. This visual suggests that more digitally advanced provinces may gradually shift toward nonroutine employment, although the relationship remains noisy, indicating the importance of controlling for fixed effects and covariates in the regression analysis. These descriptive trends preliminarily support our hypothesis that digital competitiveness is associated with an occupational shift toward nonroutine work.

5.2 Results

Table 5.1 reports estimates from the fixed effects regression model using the log of the nonroutine-to-routine employment ratio as the dependent variable. The coefficient on **lagged DCI** is positive and statistically significant ($\beta = 0.0229$, SE = 0.0068), confirming that higher digital competitiveness in the previous year is associated with a greater intensity of nonroutine work today.

The **interaction term between DCI and share of female employment** ($\beta = -0.0348$, SE = 0.0075) reveals a significantly dampened effect in female-dominated industries, consistent with the hypothesis that women may be more exposed to technological displacement, particularly where occupational upgrading is constrained.

The **industry-tech interaction terms** highlight further heterogeneity. The effect of DCI is significantly weaker in both Medium-Tech ($\beta = -0.0213$, SE = 0.0023) and High-Tech industries ($\beta = -0.0130$, SE = 0.0031), relative to Low-Tech sectors. This result, visualized in **Figure 5.3**, may appear counterintuitive at first glance. However, it plausibly reflects a "catch-up" effect: Low-Tech industries may experience more dramatic occupational shifts once digital adoption reaches critical thresholds, whereas High-Tech sectors may have already undergone earlier transformations, thus showing smaller marginal effects. While digital competitiveness is broadly associated with nonroutine job growth, its marginal effect varies across sectors and interacts meaningfully with workforce demographics.

To investigate gender-differentiated patterns in how digital competitiveness shapes labor composition, we disaggregate the analysis by computing separate province-level log ratios of nonroutine to routine employment for males and females.

Table 5.2 presents the fixed-effects regression results for male and female nonroutine employment shares. The **coefficient on lagged DCI** is positive and statistically significant for males (0.0148, $p < 0.1$), but not for females (0.0136, n.s.), suggesting that, on average, provinces with stronger digital competitiveness experience more pronounced increases in nonroutine employment among men. While this finding runs counter to the hypothesis that women are more vulnerable to routine-task displacement, it could reflect the concentration of male workers in occupations more immediately impacted by digital restructuring.

Importantly, the **interaction terms between lagged DCI and industry tech intensity** yield meaningful heterogeneity:

- In **medium-tech industries**, the effect of digital competitiveness on male nonroutine employment is large and negative (-0.0281, $p < 0.01$), and also significantly negative for females (-0.0210, $p < 0.01$). This suggests that, in sectors such as education, healthcare, and basic manufacturing, technological change is associated with a relative contraction of nonroutine roles, possibly due to automation of mid-skill tasks or slower digital absorption in support occupations.
- In **high-tech sectors**, the effect is significantly negative for males (-0.0118, $p < 0.05$) but statistically insignificant for females (-0.0057), hinting that male-dominated roles in these industries may be more exposed to technologically induced task reallocation. The relatively muted effect for women may reflect fewer female workers in high-skill tech-intensive jobs at the provincial level.

Figure 5.4 illustrates these differences visually, showing that while both gender groups benefit modestly from digital transformation in low-tech sectors, the gains dissipate—and in some cases reverse—as we move up the tech-intensity ladder.

Taken together, the gender-based decomposition at the province level highlights that male-dominated occupations in medium- and high-tech sectors are experiencing sharper declines in nonroutine task intensity in response to rising digital competitiveness, indicating a deeper transformation of task structure for men in these industries. However, women in these sectors are not seeing equivalent shifts, which may point to gendered segmentation within industry tasks or lagged occupational mobility for women in digitally evolving contexts. These patterns suggest that the labor market consequences of technological change are not only industry-specific but also gendered.

To explore sectoral heterogeneity, **Table 5.3** presents regressions run separately for low-tech, medium-tech, and high-tech industries using the province-level log ratio of nonroutine to routine employment. This approach allows us to assess how digital competitiveness affects labor structure across industries with varying technological intensities.

In **low-tech industries**, the effect of lagged DCI is small and insignificant, indicating limited digital disruption. However, the strong negative interaction with female share suggests that in provinces with higher female participation, digital transformation may reduce nonroutine job intensity—possibly due to persistent occupational segregation or lower access to upskilling.

In **medium-tech sectors**, lagged DCI remains statistically insignificant, but the negative interaction term with female share is again significant. This points to a vulnerability of women in these transitional sectors, where automation may be displacing routine tasks traditionally held by women without a corresponding increase in nonroutine roles.

Only in **high-tech sectors** do we see a statistically significant positive impact of lagged DCI on nonroutine employment, consistent with task-biased technological change. However, the continued negative gender interaction suggests that these gains may be skewed toward male-dominated roles or skill tracks.

Overall, the results underscore the importance of sectoral context. High-tech industries show the strongest gains in nonroutine employment associated with digital competitiveness, consistent with task-biased technological change. In contrast, medium- and low-tech sectors display weaker or even regressive patterns, especially when accounting for gender composition, suggesting that technological diffusion does not produce uniform benefits across the labor market.

Combined with the gender-based analysis, these patterns indicate a form of conditional skill-biased technological change—driven not only by tech intensity but also by gender segmentation and industry structure. These findings call for targeted policy interventions, such as sector-specific digital skilling programs and inclusive upskilling strategies, particularly for women in medium-tech industries and provinces lagging in digital adaptation.

5.3 Robustness Checks

To ensure the validity and stability of our findings, we conduct a series of robustness checks that account for potential confounding factors. First, in Table 5.4, we explore whether the relationship between digital competitiveness and nonroutine employment is driven disproportionately by more urbanized regions. We do this by interacting the lagged Digital Competitiveness Index (DCI) with a binary indicator for provinces above the median share of urban population. The interaction term is statistically insignificant, suggesting that the impact of DCI on the nonroutine share is not significantly different between high-urban and low-urban provinces. This mitigates concerns that infrastructure advantages or labor market concentration in urban areas are skewing the main results.

Second, we assess the sensitivity of our estimates to the inclusion of DKI Jakarta—a uniquely large and digitally advanced province that could act as an influential outlier. As shown in Table 5.5, excluding Jakarta from the sample does not materially alter the main findings. The coefficient on lagged DCI remains positive and statistically significant, and the interaction terms with industry tech-intensity maintain their direction and magnitude. Notably, the interaction for medium-tech industries remains strongly negative and significant, reinforcing our earlier conclusion that digital transformation is less inclusive in mid-tech sectors.

Together, these robustness checks support the internal validity of our identification strategy. The observed effects are not driven by extreme regional heterogeneity or unbalanced urban development, lending credibility to the conclusion that digital competitiveness influences labor market restructuring in a patterned, sector- and gender-specific manner across Indonesian provinces.

6. Discussion

6.1 Diagnostic Tests

To ensure the internal validity of our estimates, we conducted several diagnostic tests. These include assessments of multicollinearity, serial and cross-sectional dependence, influential observations, and functional form specification. We first examined multicollinearity among our independent variables using Variance Inflation Factors (VIF). While most variables were within acceptable bounds, our interaction terms exhibited moderate multicollinearity, with VIF values of 14.4 for tech_intensity, 17.0 for interaction1, and 19.0 for interaction2. Despite this, the estimated coefficients remained stable and interpretable. Next, we assessed serial and cross-sectional dependence within our province–industry panel. The Wooldridge test confirmed the presence of serial correlation ($F = 357.03$, $p < 0.001$), and the Pesaran CD test revealed strong cross-sectional dependence ($z = 322.62$, $p < 0.001$). These results motivated our use of clustered standard errors at the province level, shown in all regression tables. To detect influential observations, we examined Cook's Distance values by province. All values were well below conventional thresholds, except for Jakarta, which exhibited a moderately high score (~1.5), shown in Figure 6.3 . As our core results remained robust to the inclusion of Jakarta, we retained it in the sample. For functional form misspecification, we implemented the Ramsey RESET test. We first applied the Ramsey RESET test to the model with lagged DCI and found strong evidence of functional form misspecification ($F = 62.805$, $p < 0.001$), as shown in Figure 6.4. In response, we revised the model by including squared terms for lagged_dci and interaction2 to capture potential nonlinearities. However, these quadratic terms were not statistically significant, and the RESET test remained significant, suggesting that the misfit likely stems from more complex interactions or omitted variables rather than simple polynomial effects.(see figure 6.5)

6.2 Discussion of Main Findings and Limitations

Our empirical results confirm that Digital Competitiveness (DCI) plays a key role in shaping employment structure in Indonesia. Specifically, higher DCI levels are associated with a greater share of nonroutine (skilled) employment, but the magnitude and direction of this effect vary by industry type and demographic composition. First, we find that the marginal effect of DCI is strongest in low-tech industries, where baseline digital capabilities are low and even modest improvements trigger significant task upgrading. Surprisingly, high-tech sectors also exhibit notable gains, likely due to cumulative investments in innovation and organizational complexity. However, the positive effect of DCI diminishes in medium-tech industries, which appear more vulnerable to task displacement from automation rather than transition into nonroutine roles. Second, our interaction analysis reveals that gender differences shape the impact of digital transformation. Provinces with a higher share of female workers tend to see weaker gains in nonroutine employment from rising digital competitiveness. This suggests the presence of structural or institutional barriers that limit women's participation in digitally intensive occupations. Together, these findings highlight the heterogeneous nature of digital transition effects across Indonesia's industrial landscape and suggest that digital policy cannot be one-size-fits-all.

However, some **limitations** remain. While we lag DCI to account for delayed labor effects, potential **reverse causality** remains — regions with rising nonroutine employment might attract more digital investment in subsequent years. Moreover, despite our inclusion of quadratic terms, the **RESET test indicates functional form misspecification**, suggesting that unobserved nonlinear dynamics or threshold effects may be at play. Finally, the DCI, while useful as a provincial digital readiness measure, does not capture **within-province firm heterogeneity** or the quality of technology use — a point also raised by Acemoglu and Autor (2011) in the context of task-biased technological change.

6.3 Policy Recommendations

Digitalization is reshaping the Indonesian labor market, but its benefits are not distributed evenly across workers. Our findings show that simply expanding digital infrastructure is not enough to ensure broad-based gains. Workers in routine and mid-skill jobs are particularly vulnerable to being displaced by automation, rather than complemented by it. Therefore, policies must go beyond infrastructure investment and focus on targeted human capital development. Special attention should be given to digital skill programs designed for workers in routine occupations, equipping them with capabilities that technology cannot easily replicate. A key strategy is promoting "task transition" — helping mid-skill routine workers shift into nonroutine cognitive or manual roles, such as caregiving, creative industries, or supervisory work where human judgment, flexibility, and interaction remain crucial. Building pathways for workers to transition into these roles ensures that technology enhances opportunities rather than deepening inequalities, supporting a more inclusive digital economy for Indonesia.

7. References

Acemoglu, D., & Autor, D. H. (2011). *Skills, tasks and technologies: Implications for employment and earnings*. In O. Ashenfelter & D. E. Card (Eds.), *Handbook of Labor Economics* (Vol. 4, pp. 1043–1171). Elsevier. https://doi.org/10.1016/S0169-7218(11)02410-5.

Almeida, R., Poole, J. P., & Corseuil, C. H. (2025). The impact of digital technology on worker tasks: Do labor policies matter? *Journal of Development Economics, 175*, 103475. https://doi.org/10.1016/j.jdeveco.2024.103475

Organisation for Economic Co-operation and Development. (2011). *ISIC Rev. 3 technology intensity definition: Classification of manufacturing industries into categories based on R&D intensities*. OECD. https://www.oecd.org/sti/ind/48350231.pdf

Organisation for Economic Co-operation and Development, & Eurostat. (2005). *Oslo Manual: Guidelines for collecting and interpreting innovation data* (3rd ed.). OECD Publishing. https://doi.org/10.1787/9789264013100-en

David H. Autor, Frank Levy, Richard J. Murnane, The Skill Content of Recent Technological Change: An Empirical Exploration, The Quarterly Journal of Economics, Volume 118, Issue 4, November 2003, Pages 1279–1333, https://doi.org/10.1162/003355303322552801

Maarten Goos, Alan Manning; Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. The Review of Economics and Statistics 2007; 89 (1): 118–133. doi: https://doi.org/10.1162/rest.89.1.118

Blichfeldt, Henrik, and Rita Faullant. 'Performance Effects of Digital Technology Adoption and Product & Service Innovation – A Process-Industry Perspective'. Technovation, vol. 105, 2021, p. 102275, https://doi.org/10.1016/j.technovation.2021.102275

Wu, Yunxia, et al. 'The Impact of Digitization in Manufacturing on Female Employment and Gender Wage Gap'. Journal of Asian Economics, vol. 95, 2024, p. 101821, https://doi.org/10.1016/j.asieco.2024.101821

Sangmin Aum and Yongseok Shin, "The Labor Market Impact of Digital Technologies," Federal Reserve Bank of St. Louis Review, Second Quarter 2025, Vol. 107, No. 5, pp. 1-10. https://doi.org/10.20955/r.2025.05

Group 2

Econometrics II - Spring 2025

A. Appendix: The Indonesian Labor Force Survey Datasets

Figure A.1. Sample View of the Indonesian LFS Data Structure

	TAHUN	URUTAN	SMT	FINAL_WEIG	KODE_PROV	NAMA_PROV	KODE_KAB	NAMA_KAB	ID_NKS	NO_DSRT	KLASIFIKAS	JLH_ART
1	20208	1	2	42	11	ACEH	1	SIMEULUE	9	1	2	3
2	20208	2	2	43	11	ACEH	1	SIMEULUE	9	1	2	3
3	20208	3	2	42	11	ACEH	1	SIMEULUE	9	1	2	3
4	20208	4	2	55	11	ACEH	1	SIMEULUE	9	2	2	5
5	20208	5	2	43	11	ACEH	1	SIMEULUE	9	2	2	5
6	20208	9	2	31	11	ACEH	1	SIMEULUE	9	4	2	4
7	20208	10	2	73	11	ACEH	1	SIMEULUE	9	4	2	4
8	20208	11	2	86	11	ACEH	1	SIMEULUE	9	4	2	4
9	20208	12	2	89	11	ACEH	1	SIMEULUE	9	4	2	4
10	20208	13	2	42	11	ACEH	1	SIMEULUE	9	5	2	3
11	20208	14	2	42	11	ACEH	1	SIMEULUE	9	5	2	3
12	20208	16	2	51	11	ACEH	1	SIMEULUE	9	6	2	3
13	20208	17	2	45	11	ACEH	1	SIMEULUE	9	6	2	3
14	20208	19	2	38	11	ACEH	1	SIMEULUE	9	7	2	4
15	20208	20	2	43	11	ACEH	1	SIMEULUE	9	7	2	4

Figure A.2. Number of Observations and Variables in Each Year

sakernas_2020	793202 obs. of 132 variables
sakernas_2021	777982 obs. of 189 variables
sakernas_2022	752688 obs. of 220 variables
sakernas_2023	758900 obs. of 220 variables
sakernas_2024	768855 obs. of 253 variables

Group 2

Econometrics II - Spring 2025

B. Appendix: Data Cleaning Procedure and the Cleaned Indonesia's LFS Data Structure

1. We define a binary **work** variable equal to 1 if an individual meets any of the following criteria:
    a. Worked for at least one hour in the past week:
    b. Operated a business or engaged in activities intended to generate income;
    c. Assisted in running a family or another person's business or economic activity;
    d. Had a job or business but was temporarily absent from work during the past week.
2. We define a binary **formal work** variable equal to 1 if an individual is employed and their employment status is classified as either: employee or employer assisted by permanent or paid workers.
3. We classify job tasks using a combination of occupation and industry codes as follows:
    a. **Routine cognitive**: Individuals working in clerical, administrative, or sales occupations (occupation code = 4, or 5 **and** industry code = 7);
    b. **Routine manual**: Individuals in production or operative occupations (occupation codes = 6, 7, 8, or 9);
    c. **Nonroutine cognitive**: Individuals in managerial, professional, or technical occupations (occupation codes = 1, 2, or 3);
    d. **Nonroutine manual**: Individuals in service-related occupations that do not fall into the above categories.
4. We retain the following variables for analysis: **year**, **weight**, **province_code**, **urban**, **gender**, **age**, **education**, **industry_code**, **occupation_code**, **work_hour**, **employment_status**, **internet_use**, **computer_use**, and **wage**.
5. We then combine the cleaned yearly datasets and construct a **province–industry level panel dataset** by aggregating individual records using sampling weights, by excluding the public sector workers and restrict the sample to formal sector workers
6. Finally, we verify that the calculated number of formal workers aligns with official statistics published by Statistics Indonesia.

The Cleaned Indonesia's LFS Data Structure

	year	province_code	industry_code	weighted_nonroutine	weighted_routine	weighted_cognitive	weighted_manual	weighted_internet	weighted_computer	total_weight	weighted_nonroutine_f	weighted_routine_f
1	2020	11	1	5632	117414	4421	118625	14899	0	123046	163	181
2	2020	11	2	2916	5978	2398	6496	2605	0	8894	42	
3	2020	11	3	9579	49220	4389	54410	15838	0	58799	1480	180
4	2020	11	4	4179	3580	4200	3559	6136	0	7759	0	4
5	2020	11	5	512	1986	493	2005	847	0	2498	0	
6	2020	11	6	6296	64672	7537	63431	13965	0	70968	0	5
7	2020	11	7	6836	92656	78886	20606	45458	0	99492	1067	218
8	2020	11	8	3293	30895	2736	31452	13546	0	34188	288	2
9	2020	11	9	26913	4716	3063	28566	10474	0	31629	8629	24
10	2020	11	10	4127	1031	3799	1359	5158	0	5158	567	2
11	2020	11	11	5423	5924	9119	2228	10786	0	11347	1228	16
12	2020	11	12	102	128	102	128	102	0	230	0	1
13	2020	11	13	8417	7749	5308	10858	10080	0	16166	1446	34
14	2020	11	15	144398	12825	151850	5373	123951	0	157223	105185	51
15	2020	11	16	55659	9301	55891	9069	50951	0	64960	42453	58
16	2020	11	17	13974	30179	6592	37561	10227	0	44153	6235	159
17	2020	12	1	25800	440558	18863	447495	56920	0	466358	3158	823
18	2020	12	2	1839	16173	2005	16007	3243	0	18012	0	20
19	2020	12	3	45568	364466	33329	376705	83086	0	410034	5286	1064
20	2020	12	4	3965	8269	5427	6807	6825	0	12234	0	23

Group 2

Econometrics II - Spring 2025

C. Appendix: East Venture Digital Competitiveness Index (EV-DCI)

The EV-DCI 2024 measures digital competitiveness in 38 provinces and 157 cities/districts in Indonesia, including four new provinces formed from provincial expansion in 2022. EV-DCI 2024 marks the second year in its index calculation conducted for 38 provinces.

For these newly established provinces, data was gathered from existing cities/districts and aggregated based on EV-DCI indicators, given the unavailability of provincial level data. The same method was applied to the four new provinces in the EV-DCI 2023 report.

This index comprises three sub-indexes: Input, Output, and Support. Each sub-index consists of three pillars, totaling nine pillars that form the EV-DCI. These nine pillars are compiled from 50 indicators, with a distribution of 3-9 indicators for each pillar.

EV-DCI 2024 Indicator List

EV-DCI 2024 indicators list

INPUT

Human Resources

1. Number of Students with Digital Capabilities
2. Growth of Students with Digital Capabilities
3. Number of Lecturers in Digitalization-Related Study Programs
4. Number of Digitalization-Related Study Programs
5. Digital Literacy Index

ICT Usage

1. Ratio of Citizens that Have Cellular Phone
2. Ratio of Households that Have Computer
3. Ratio of Citizens that Have Access to Internet
4. Ratio of Citizens that Have Access Internet from Home
5. Ratio of Citizens that Access Internet from Office
6. Ratio of Citizens that Access Internet from School
7. Ratio of Citizens that Access Internet with Laptop
8. Ratio of Citizens that Access Internet with Cellular Phone

ICT Expenditure

1. Ratio of Households Who Have ICT Expenditure
2. Average Expenditure of Households for ICT
3. Total Renumeration and Wage of Information and Communication Sector Workers
4. Average Renumeration and Wage of Information and Communication Sector Workers

OUTPUT

Economy

1. GRDP of the Information and Communication Sector
2. GRDP Contribution of the Information and Communication Sector
3. GRDP Growth of the Information and Communication Sector
4. GRDP of Warehousing, Transportation Support, Post & Courier Subsector
5. GRDP Contribution of Warehousing, Transportation Support, Post & Courier Subsector
6. GRDP Growth of Warehousing, Transportation Support, Post & Courier Subsector
7. GRDP of the Financial Services Sector
8. GRDP Contribution of the Financial Services Sector
9. GRDP Growth of the Financial Services Sector

Entrepreneurship and Productivity

1. Ratio of Workers Using the Internet in their Main Job
2. Ratio of Workers Using the Internet at Work for Communication
3. Ratio of Workers Using the Internet at Work for Marketing
4. Ratio of Workers Using the Internet at Work for Sales via Social Media
5. Ratio of Workers Using the Internet at Work for Sales via Website/E-commerce
6. Loan Using Fintech

Manpower

1. Number of Workers in Digitalization-Related Sectors
2. Ratio of Workers in Digitalization-Related Sectors
3. Growth of Workers in Digitalization-Related Sectors
4. Number of Workers in Digitalization-Prone Categories (Reverse Indicator)
5. Ratio of Workers in Digitalization-Prone Categories (Reverse Indicator)
6. Growth of Workers in Digitalization-Prone Categories (Reverse Indicator)

SUPPORT

Infrastructure

1. Level of Electricity Disturbance (Reverse Indicator)
2. Ratio of Villages that Get Strong and Very Strong Signal
3. Ratio of Villages that Get 3G Signal
4. Ratio of Villages that Get 4G Signal
5. Ratio of Households with Fixed Line Connection

Finance

1. Financial Inclusion Index
2. Number of Officeless Financial Services Agents (Laku Pandai Agents)
3. E-wallet Adoption as Payment Method

Regulation and Capacity of the Regional Government

1. Gross Enrollment Rate of Senior High Shools/ Vocational Schools
2. Gross Enrollment Rate of Higher Education (Diploma-Bachelor)
3. Life Expectancy Growth
4. Reduction of Poverty Rate (Reverse Indicator)

D. Appendix: Variables Definition

Variable Name	Definition	Level	Source
share_nonroutine_to_routine	Share of individuals working in nonroutine occupations relative to routine occupations	Industry - Province	Indonesian LFS
share_college	Share of individuals aged over 15 with a college degree	Industry - Province	Indonesian LFS
share_urban	Share of individuals aged over 15 who live in urban area	Industry - Province	Indonesian LFS
Digital Competitiveness Index (DCI)	Measure of technological advancement in each province (from 0-100)	Province	East Venture
GDP	The total gross value added of all goods and services produced within a province's economy over one year period	Province	Statistics Indonesia
Gini coefficient	Statistical measure of income inequality in each province (from 0-1)	Province	Statistics Indonesia

E. Appendix: Descriptive Statistics, Results and Robustness Checks

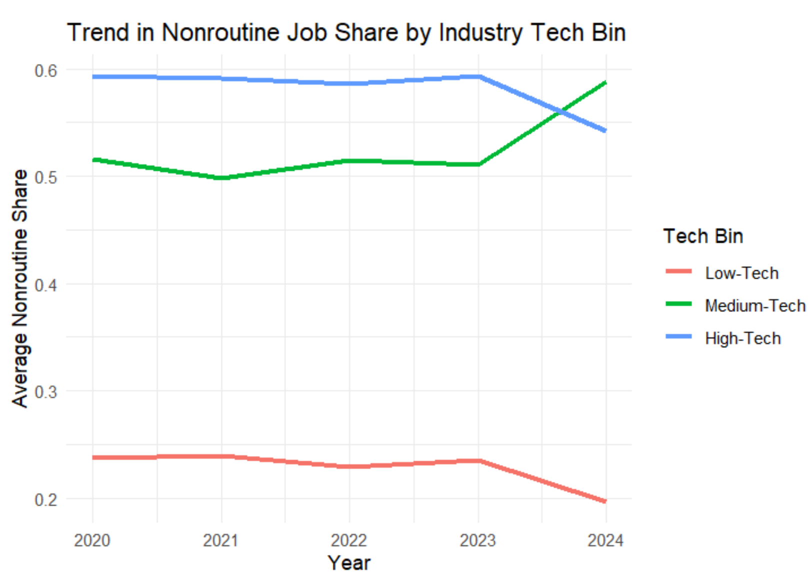

Figure 5.1

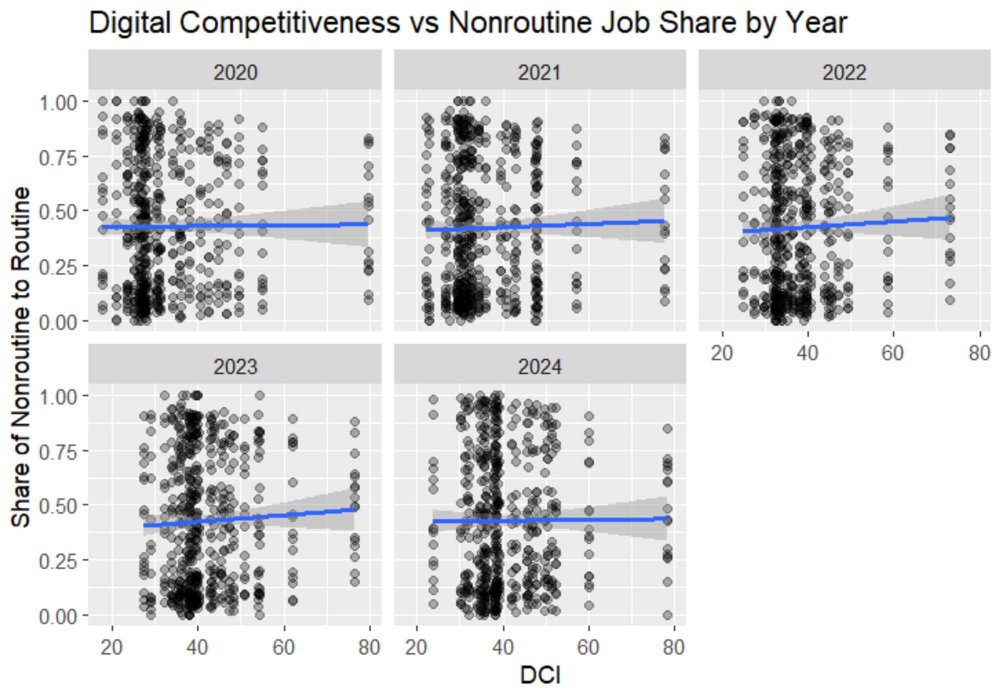

Figure 5.2

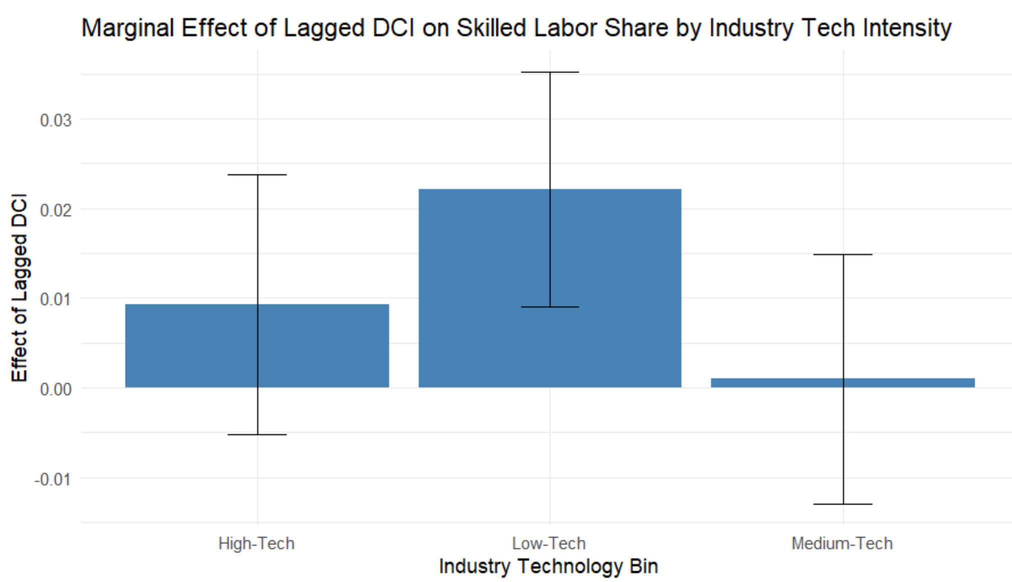

Figure 5.3

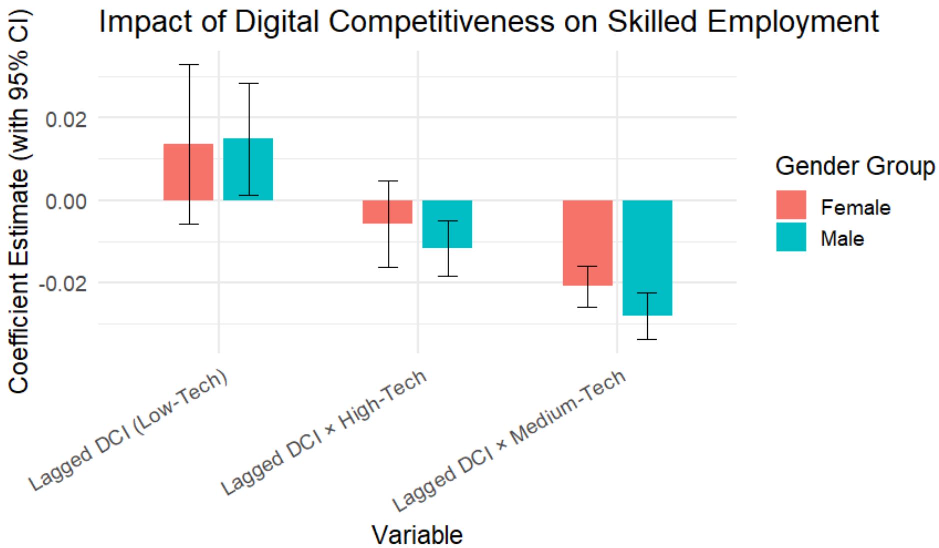

Figure 5.4

	model_nonroutine
Dependent Var.:	log_ratio_nonroutine
lagged_dci	0.0229** (0.0068)
share_female	1.365*** (0.3010)
interaction2	-0.0348*** (0.0075)
share_college	-3.925 (3.881)
GDP	-1.46e-8 (9.32e-8)
share_urban	0.5685* (0.2383)
Gini	0.4660 (1.280)
lagged_dci x industry_tech_binMedium-Tech	-0.0213*** (0.0023)
lagged_dci x industry_tech_binHigh-Tech	-0.0130*** (0.0031)
Fixed-Effects:	--------------------
Province_code	Yes
industry_code-year	Yes
S.E.: Clustered	by: Province_code
Observations	2,575
R2	0.90075
Within R2	0.07505

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 5.1

	model_male	model_female
	Male	Female
Dependent Var.:	log_ratio_nonroutine_male	log_ratio_nonroutine_female
lagged_dci	0.0148* (0.0069)	0.0136 (0.0099)
share_college	-5.408 (4.054)	-0.6550 (6.575)
GDP	-1.02e-7 (1.15e-7)	4.84e-7. (2.41e-7)
share_urban	0.5419* (0.2386)	0.6670. (0.3792)
Gini	-0.2199 (1.345)	0.8674 (2.651)
lagged_dci x industry_tech_binMedium-Tech	-0.0281*** (0.0029)	-0.0210*** (0.0025)
lagged_dci x industry_tech_binHigh-Tech	-0.0118** (0.0034)	-0.0057 (0.0054)
Fixed-Effects:	-------------------------	-------------------------
province_code	Yes	Yes
industry_code-year	Yes	Yes
S.E.: Clustered	by: province_code	by: province_code
Observations	2,569	2,191
R2	0.87042	0.83431
Within R2	0.05236	0.02157

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 5.2

	model_lowtech	model_medtech	model_hightech
	Low-Tech	Medium-Tech	High-Tech
Dependent Var.:	log_ratio_nonroutine	log_ratio_nonroutine	log_ratio_nonroutine
lagged_dci	0.0027 (0.0082)	0.0153. (0.0089)	0.0319* (0.0140)
share_female	1.902*** (0.5168)	0.5345 (0.3987)	0.5894 (1.421)
interaction2	-0.0380** (0.0120)	-0.0247** (0.0078)	-0.0288 (0.0363)
share_college	82.72 (86.81)	-0.2123 (6.441)	-3.379 (2.388)
GDP	-2.09e-7 (2.7e-7)	1.34e-7 (2.3e-7)	8.34e-7** (2.62e-7)
share_urban	1.118* (0.4596)	-0.0969 (0.2713)	0.6795 (0.5169)
Gini	1.998 (2.045)	3.062 (2.248)	-4.133. (2.389)
Fixed-Effects:	--------------------	--------------------	--------------------
Province_code	Yes	Yes	Yes
industry_code-year	Yes	Yes	Yes
S.E.: Clustered	by: Province_code	by: Province_code	by: Province_code
Observations	778	800	470
R2	0.90515	0.90105	0.55493
Within R2	0.08074	0.01621	0.02977

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 5.3

	model_urban_intera..
Dependent Var.:	log_ratio_nonroutine
lagged_dci	0.0228** (0.0076)
urban_high	-0.0333 (0.1786)
share_female	1.662*** (0.3307)
interaction2	-0.0516*** (0.0082)
share_college	-4.268 (2.718)
GDP	1.89e-7 (1.8e-7)
Gini	1.486 (1.236)
lagged_dci x urban_high	0.0011 (0.0051)
Fixed-Effects:	--------------------
Province_code	Yes
industry_code-year	Yes
S.E.: Clustered	by: Province_code
Observations	2,048
R2	0.90146
Within R2	0.04401

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 5.4

	model_excl_jakarta
Dependent Var.:	log_ratio_nonroutine
lagged_dci	0.0288** (0.0086)
share_female	1.222* (0.4762)
interaction2	-0.0388** (0.0136)
share_college	-2.846 (2.688)
GDP	3.25e-8 (1.43e-7)
share_urban	0.3417 (0.2266)
Gini	1.258 (1.266)
lagged_dci x industry_tech_binMedium-Tech	-0.0196*** (0.0039)
lagged_dci x industry_tech_binHigh-Tech	-0.0070 (0.0045)
Fixed-Effects:	--------------------
Province_code	Yes
industry_code-year	Yes
S.E.: Clustered	by: Province_code
Observations	1,985
R2	0.90631
Within R2	0.04634

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 5.5

F. Appendix: Diagnostic Test Tables

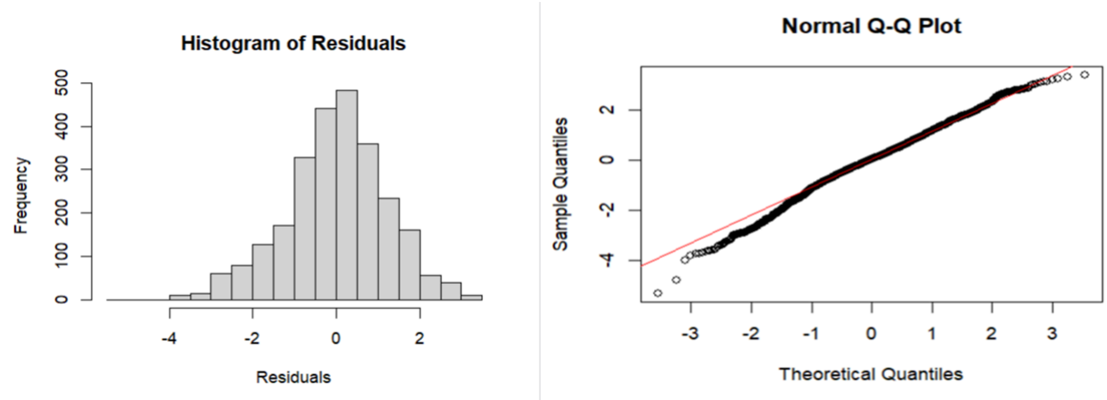

Figure 6.1 and 6.2

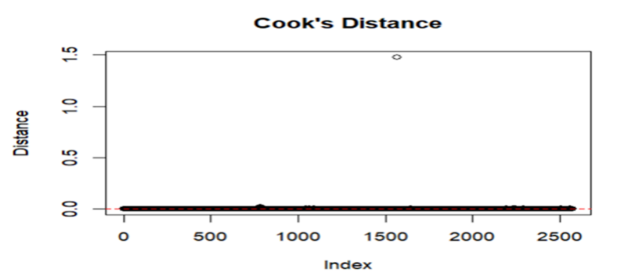

Figure 6.3

```
	RESET test

data:  reset_model
RESET = 62.805, df1 = 2, df2 = 2561, p-value < 2.2e-16
```

Figure 6.4

```
OLS estimation, Dep. Var.: log_ratio_nonroutine
Observations: 2,575
Fixed-effects: Province_code: 34,  industry_code^year: 79
Standard-errors: Clustered (Province_code)
                   Estimate  Std. Error   t value  Pr(>|t|)
lagged_dci       1.664546e-02 0.018843724  0.883342 0.3834416
lagged_dci_sq    7.683990e-06 0.000215917  0.035588 0.9718255
interaction2    -5.107373e-02 0.022051775 -2.316082 0.0269094 *
interaction2_sq -1.272655e-05 0.000282617 -0.045031 0.9643539
share_female     1.898251e+00 0.533230775  3.559905 0.0011506 **
share_college   -5.269394e+00 4.136882058 -1.273760 0.2116472
GDP              2.170000e-09 0.000000114  0.019077 0.9848946
share_urban      5.951338e-01 0.290185703  2.050872 0.0482887 *
Gini             5.238484e-01 1.278798032  0.409641 0.6847162
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.531763     Adj. R2: 0.89299
                 Within R2: 0.049254
```

Figure 6.5

F. Appendix: Contributions

Hendro Wicaksono: Conceptualization, Data Curation, Methodology, Formal Analysis, Writing - Original Draft.

Kamran Rzazade: Integration of industry codes into full panel dataset, Diagnostic testing, Model evaluation, Results discussion, Limitation assessment, Policy recommendations.

Aadhya Krishnakumar: Literature review, Theoretical framework development, Data augmentation for panel construction, Investigation, Resources.

Sri Ganduri: Data merging, Econometric analysis, Implemented multiple regression specifications, and created visualizations to illustrate empirical results, Robustness Checks specifications