Building Trust, Hiring More: The Role of Privacy Policies in Workforce Growth

Linghao Zhang *[a], Qiubo Li[b], Yiwen Ma[c] and Shaoran Li[d]

[a] *Renmin University of China, China*
[b] *University of Jinan, China*
[c] *The University of Hong Kong, Hong Kong, China*
[d] *Peking University, China*

July 2025

Abstract

As AI and big data reshape the global economy, concerns over technological unemployment are mounting. This paper offers a countervailing perspective by examining how strategic adaptation to data privacy risks—specifically, voluntary adoption of Personal Data Protection Systems (PDPS)—affects employment. Using a staggered Difference-in-Differences approach on Chinese A-share firms (2009–2020), we find that PDPS adoption significantly increases employment, particularly in firms with greater market power, lower financial constraints, and higher wages. Mechanism analysis via "instrumental mediation analysis" provides causal evidence that PDPS boosts demand for R&D labor and improves firm performance through enhanced customer trust. A series of robustness checks support the credibility of our results. Our findings offer a counterbalance to the dominant narrative of labor displacement in the AI era: in responding to privacy concerns, firms may actually create new jobs to manage and mitigate technological risks.

JEL Classification: J23, G32, L51
Keywords:Personal Data Protection; Corporate Employment; China

*There are no conflicts of interest in this paper. All errors are ours.

1 Introduction

The global economy is undergoing a profound technological transformation, driven by the exponential growth of big data and related advances in artificial intelligence (Brynjolfsson et al., 2019; Agrawal et al., 2019). As the core engine of this revolution, data is fundamentally reshaping industrial organization and labor market dynamics, raising pressing questions about its employment implications (Acemoglu and Restrepo, 2019). Historical experience suggests that while technological change may displace labor in the short term, it also tends to generate new employment opportunities through the emergence of new industries and tasks (Autor, 2015). However, the current wave of data-driven automation targets not only manual labor but also cognitive and analytical work—tasks once considered uniquely human—thereby intensifying concerns over widespread job displacement across sectors and skill levels (Frey and Osborne, 2017; Acemoglu and Restrepo, 2020). This creates a central theoretical tension in the literature: is technology primarily a force of labor displacement or one of creative destruction that ultimately generates new employment opportunities? The debate remains vigorous, with empirical evidence often pointing in both directions, suggesting that the net effect on employment is contingent on the type of technology, the industry context, and the policy environment.

This study enters the discourse by examining a critical, yet relatively underexplored, facet of the modern technological revolution: the corporate response to data privacy risks. As The Economist points out, "the world's most valuable resource is no longer oil, but data". The proliferation of AI and big data is inextricably linked to the massive collection and analysis of personal information, which has given rise to significant public and regulatory concern over privacy and data security. The prevailing view in much of the economic literature has been to frame privacy protection as a compliance cost—a regulatory burden imposed on firms that can stifle innovation and hinder performance (Peukert et al., 2022). This perspective aligns with the broader narrative of technology-induced challenges, suggesting that the costs of mitigating technology-related risks add another layer of friction to economic activity. In contrast, a growing body of work in strategic management and information

systems proposes that proactive data governance can be a source of competitive advantage. In an environment of heightened information asymmetry and consumer anxiety, a firm's commitment to privacy can serve as a powerful signal of trustworthiness and quality (Acquisti et al., 2016; Martin et al., 2017). By building customer trust, firms may enhance loyalty, command price premiums, and ultimately improve financial performance. This has sparked controversy in this specific technology scenario: is investment in privacy a net drain on firm resources, or is it a strategic investment that yields positive returns?

Our study is motivated by the intersection of these two major debates. We ask: what are the employment consequences when firms strategically respond to the risks inherent in big data technology? Specifically, we move beyond the well-trodden ground of mandatory regulations like the GDPR, which has been the focus of most prior work (Johnson et al., 2023; Demirer et al., 2024). Instead, we investigate the effects of firms' voluntary adoption of Personal Data Protection Systems (PDPS). This focus is crucial for resolving the aforementioned controversies. Unlike compliance with mandatory rules, voluntary adoption represents a proactive, strategic choice, offering a clearer lens through which to test whether privacy-enhancing actions are viewed by firms as a value-creating investment. By linking this strategic choice directly to firm-level employment, we provide a novel test of the competing narratives on technology and labor. If the "compliance cost" view is correct, we would expect PDPS adoption to have a neutral or even negative impact on hiring as firms divert resources. Conversely, if the "strategic advantage" view holds, then the trust and performance benefits of PDPS could fuel business expansion and, consequently, job creation. This study, therefore, aims to resolve this ambiguity by providing causal evidence on whether firms' strategic efforts to mitigate technological risks can, counterintuitively, become a catalyst for workforce growth.

China offers a particularly suitable empirical setting for this analysis. Its comprehensive Personal Information Protection Law (PIPL) did not take effect until November 2021, providing a distinct period during which firms could voluntarily adopt PDPS without government mandate. This institutional feature allows us to employ a quasi-experimental design that compares early voluntary adopters (treatment group) with non-adopters (control

group), helping to mitigate confounding effects from national regulation. This unique institutional context is set against the backdrop of China's explosive digital transformation, where the rapid growth of the digital economy made data collection and privacy concerns increasingly salient for a vast number of corporations. Additionally, mandatory disclosure requirements for Chinese-listed companies provide detailed and reliable employment data.

We use a staggered DID design with two-way fixed effects on a panel of A-share firms listed on the Shanghai and Shenzhen stock exchanges from 2009 to 2020. Our baseline estimates suggest that voluntary PDPS adoption leads to a statistically significant at least **8% increase in total employment** relative to firms that do not adopt such systems. To ensure the credibility of this main finding, we conduct a comprehensive battery of robustness tests. We also document significant heterogeneity in the employment effects. Firms with **greater market power**, **fewer financial constraints**, and **higher employee compensation** exhibit stronger employment growth following PDPS adoption. These results underscore the role of strategic and organizational factors in shaping firms' labor responses to data governance challenges.

To understand how the voluntary adoption of a Personal Data Protection System (PDPS) stimulates hiring, we investigate two primary mechanisms: an R&D investment channel and an operating performance channel. The first channel posits that implementing PDPS necessitates hiring specialized technical and highly-educated personnel, directly increasing labor demand. The second suggests that by enhancing customer trust and mitigating privacy risks, PDPS improves firm performance, which in turn fuels workforce expansion.

Our initial analysis following the Baron and Kenny (Baron and Kenny, 1986) framework, however, uncovered a puzzle. While PDPS adoption is positively associated with a higher proportion of technical (*Tech*) and highly-educated (*High*) employees, these same variables showed a negative correlation with a firm's total employment. This suggests a potential crowding-out effect, where the high cost of skilled labor could constrain a firm's overall hiring budget, complicating a simple mediation story.

To disentangle these countervailing effects, we employ a two-stage "instrumental

mediation analysis [1]." This approach uses the firm's decision to adopt PDPS to isolate the portion of the mediator that is causally driven by the new privacy system. The results from this advanced analysis resolve the initial puzzle: the instrumented measures of skilled labor now have a significant positive effect on total employment. This finding suggests that PDPS adoption not only drives demand for skilled workers but also creates conditions—likely through the performance channel—that alleviate budget constraints, enabling a net increase in the firm's total workforce. Applying this framework across all our proposed mediators provides robust evidence supporting both the R&D and operating performance channels as key drivers of employment growth.

This paper makes three principal contributions to the corporate finance literature.

First, we contribute to the rapidly evolving literature on how digital technologies reshape employment patterns. While extensive research has documented the labor-displacing effects of automation and AI (Acemoglu and Restrepo, 2020; Autor et al., 2024; Frey and Osborne, 2017), we provide a counterbalancing perspective by demonstrating that firms' strategic responses to technology-induced risks can actually create employment. Our findings align with recent work showing that technologies can generate complementary labor demand through new task creation (Acemoglu and Restrepo, 2019; Bessen et al., 2023; Jiang et al., 2025a). We extend this literature by identifying privacy protection as a specific channel through which technological advancement stimulates job creation, particularly for skilled technical workers. This complements recent studies on the heterogeneous employment effects of digitalization across skill levels (Harrigan et al., 2021; Webb, 2019) and adds nuance to the ongoing debate about technological unemployment in the AI era (Brynjolfsson et al., 2025; Hui et al., 2024).

Second, we advance the emerging literature on the economic implications of data privacy regulations and corporate data governance. While prior research has primarily focused on mandatory privacy laws like GDPR (Johnson et al., 2023; Peukert et al., 2022; Goldfarb and Que, 2023), this study pioneers the examination of voluntary privacy protection adoption and its firm-level consequences. Our work extends recent findings on how privacy

[1] This method is adapted from Di Giuli and Laux (2022) and Liu (2025).

regulations affect firm performance (Demirer et al., 2024; Aridor et al., 2020), innovation activities (Chen et al., 2024a; Blind et al., 2024), and market structure (Johnson et al., 2023). By demonstrating that voluntary PDPS adoption enhances firm performance through improved customer trust and operational efficiency, we contribute to the growing body of work on the strategic value of data governance (Martin et al., 2020; Jia et al., 2021). This complements recent studies showing how privacy-enhancing technologies can serve as competitive differentiators (Acquisti et al., 2016).

Third, we enrich the corporate finance literature on the determinants of firm-level employment. Traditional models emphasize factors such as financial performance, investment opportunities, and financing constraints (Nickell, 1999; Caggese and Cuñat, 2008). We introduce a novel, non-financial determinant: strategic investment in digital risk management. Our analysis causally identifies two mechanisms through which this occurs—a direct R&D investment channel driven by the need for technical human capital (Acemoglu and Restrepo, 2020), and an indirect operating performance channel where enhanced customer trust fuels business expansion and subsequent hiring (Martin et al., 2017). Collectively, these contributions offer a more nuanced understanding of how firms' strategic responses to the challenges of the big data era can foster, rather than hinder, employment.

2 Literature review and theoretical development

In recent years, big data technologies and artificial intelligence (AI), along with the modern automation technologies built upon them, have undergone rapid development (Brynjolfsson et al., 2019), fundamentally reshaping global labor market structures and employment patterns (Acemoglu and Restrepo, 2019). The extant literature concerning the impact of technological revolution on labor markets presents a complex duality of characteristics.

On one hand, technological advancement demonstrates pronounced employment-enhancing effects. Within the realm of information and communication technology (ICT), a significant correlation exists between increased ICT adoption rates and skilled labor demand (Jiang et al., 2025a; Gallipoli and Makridis, 2018; Maurin and Thesmar, 2004). The adoption

of new technologies generates novel positions, enhances worker productivity (Vu, 2013), and subsequently stimulates labor demand growth through income effects (Autor, 2015). Regarding innovation activities, product innovation has been shown to stimulate employment growth (Dachs and Peters, 2014; Vivarelli et al., 1996), while simultaneously maintaining lower employee turnover rates (Greenhalgh and Mavrotas, 1996; Pacelli et al., 1998). Furthermore, research by Edler and Ribakova (1994) and Gentili et al. (2020) provides additional evidence that robotics technology increases demand for skilled labor, with related industries exhibiting superior productivity and employment growth.

However, conversely, the technological revolution also engenders considerable labor substitution effects. Specifically, Behaghel and Moschion (2016) and Balsmeier and Woerter (2019) have substantiated that IT diffusion, while increasing the proportion of high-skilled occupations, concurrently diminishes the share of low-skilled positions. Similarly, although robotics technology demonstrates exceptional performance in enhancing labor productivity, it simultaneously reduces the employment share of low-skilled workers (De Vries et al., 2020; Camina et al., 2020; Blanas et al., 2019; Acemoglu and Restrepo, 2019). Moreover, innovation activities manifest pronounced polarization effects, with Cirillo (2014) indicating that innovation may diminish the firm-specific human capital value of non-technical workers, thereby increasing their mobility (Silva and Lima, 2017).

Concomitant with the rapid advancement of big data technologies, privacy breaches and data security vulnerabilities have become increasingly prominent, elevating personal data protection as a critical issue in the digital economy era. Currently, research on privacy protection predominantly concentrates on the impacts of mandatory regulatory policies such as the European Union's GDPR (Demirer et al., 2024; Aridor et al., 2020; Chen et al., 2024a; Blind et al., 2024; Johnson et al., 2023). Distinct from the passive compliance model of mandatory privacy regulations, we focus on enterprises' voluntary adoption of PDPS as a proactive strategic choice, thereby providing a quasi-experimental design and a relatively neglected mechanism for exploring the technological revolution-labor market nexus. PDPS implementation transcends mere compliance mechanisms, representing a strategic transformation in corporate governance that significantly influences enterprise technological

investment, customer trust and labor demand.

Building upon this analytical framework, we propose the following hypotheses:

H1: Voluntary PDPS adoption by firms significantly enhances corporate employment levels.

Drawing upon resource-based theory and signaling theory, with consumer trust as the core linkage, we elucidate in depth how enterprises' voluntary adoption of PDPS promotes employment growth through distinctive mechanisms.

Resource-based theory and the complementary capabilities proposition provide a deeply persuasive analytical framework for understanding how PDPS implementation drives employment growth through R&D investment. In the era of rapidly evolving big data technologies, characterized by exponential expansion in data collection and analytics capacities, privacy breaches and data security vulnerabilities have become increasingly salient, concomitantly elevating consumer apprehensions regarding corporate data handling practices (Martin et al., 2017). Within this contextual milieu, establishing and maintaining stakeholder trust has emerged as a strategic imperative for contemporary enterprises. In the increasingly important context of data protection, firms have strong motivations to build PDPS systems through continuous R&D investments to maintain their differentiated competitive advantages and cultivate robust trust relationships with their clientele. The empirical research by Johnson et al. (2023) provides strong support for this perspective, finding that companies responding to privacy regulations significantly increased their R&D investments, especially in privacy-enhancing technologies. Regarding the relationship between corporate R&D investment and employment levels, when firms proactively implement PDPS, they inevitably need to invest in developing specialized privacy-enhancing technologies, including but not limited to advanced encryption algorithms, data anonymization techniques, and security protocol systems (Acquisti et al., 2016). Leiponen (2005) and Mohnen and Röller (2005) demonstrate that to obtain substantial benefits from innovation activities, firms must cultivate sufficient internal capabilities; among these, employees' technical expertise constitutes the core element of a firm's learning capacity and knowledge absorption capability, effectively enhancing internal R&D activities and the

positive impact of innovation on profitability. As pointed out by Hall and Rosenberg (2010), 50% or more of R&D expenditures are wages for highly educated scientists and engineers, and corporate R&D investment is largely human capital investment. Therefore, firms must expand their R&D capabilities by recruiting technical talent with professional knowledge in privacy technology fields (Tambe et al., 2020). This targeted talent acquisition represents firms' deliberate investment in strategic capabilities, which can significantly enhance their competitive position in increasingly privacy-conscious market environments (Blind et al., 2024).

The signal theory provides another profound theoretical perspective for analyzing how PDPS stimulates the labor market by optimizing corporate operational performance. Initially developed by Spence (1978), this theory elucidates how firms transmit their intrinsic quality characteristics to stakeholders through observable behaviors in environments characterized by information asymmetry. By proactively establishing PDPS, firms demonstrate their firm commitment to responsible data governance and customer-centric value propositions, thereby significantly cultivating and strengthening stakeholder and customer trust in an environment of growing privacy concerns and information vulnerability (Aridor et al., 2021), which undoubtedly generates positive reputational effects (Syed, 2019; Cavusoglu et al., 2004). An exemplary corporate reputation is regarded as a strategic intangible asset that creates sustainable competitive advantages due to its scarcity and inimitability (Barney, 1991). The academic community widely acknowledges corporate reputation as a critical strategic indicator for enhancing corporate value (McGuire et al., 1988; Flammer, 2013), exerting significant positive influence on corporate performance (Miller et al., 2020; Zhu et al., 2014). Additionally, companies empowering consumers and building a foundation of trust with their customers incentivize consumers to pay a premium for privacy protection (Tsai et al. (2011)). From an operational risk dimension, Kamiya et al. (2021) noted that regulatory penalties and market volatility resulting from data breach incidents severely undermine corporate performance, with affected firms experiencing shareholder wealth losses exceeding $500 million per incident. Through the proactive deployment of privacy protection mechanisms, organizations can effectively mitigate these

risks and maintain operational stability. The synergistic effects of revenue growth potential, enhanced customer loyalty, and reduced operational risks directly influence corporate talent recruitment strategies. As multidimensional business performance improves, firms demonstrate a greater propensity to expand recruitment scales to further develop business territories and enhance long-term corporate value (Van Reenen, 1997; Lachenmaier and Rottmann, 2011).

Accordingly, we propose the following hypothesis:

H2: Voluntary PDPS adoption generates employment via R&D investment and operational performance.

3 Data and Research Methodology

3.1 Data and Sample

Since 2007, China has been adopting internationally recognized accounting standards. We use data from 2008 as the starting year to ensure the comparability of financial data. The mandatory national regulation, the Personal Information Protection Law of the People's Republic of China (PIPL), was officially implemented on November 1, 2021. In order to effectively assess the voluntary adoption of personal data protection by firms, we select all A-share listed firms from 2009 to 2020 [2] as the initial sample. Furthermore, we applied the following adjustments: (1) excluding ST/ST* firms; (2) excluding financial sector firms; (3) excluding insolvent firms; (4) excluding observations with missing important variables; and (5) winsorizing continuous variables at the 1st and 99th percentiles. As a result, we obtained a sample of 2,838 listed firms, yielding 22,213 annual observations. The data on personal data protection was obtained through the text mining of corporate annual reports using Python software and large language models (LLMs). Data on corporate employment were sourced from the CSMAR database.

[2]In China, when a new regulation is enacted, firms often implement it at any cost, making it difficult to observe the differential impact of the law at the firm level and to identify an appropriate control group. However, the spontaneous actions of firms prior to such regulations provide valuable insights, enabling us to construct a suitable control group for analyzing the potential impact of the PDPS on firms.

3.2 Variable Definitions

3.2.1 Personal Data Protection System

Our main independent variable is the PDPS. In terms of data collection, commercial entities often acquire highly sensitive user information, including biometric data, personal information, software development kits (SDKs), and other private content. Some internet platforms utilize tracking cookies to store user data, employing technologies to create behavioral profiles of users. By combining individual characteristics, they implement precision marketing strategies, which, in effect, pose a potential risk to citizens' privacy and security. Regarding data processing, firms may implement user privacy protection mechanisms at the technological level. These measures include the application of encryption algorithms, anonymization procedures, de-identification technologies, and the establishment of security risk assessment frameworks. In terms of compliance management, commercial entities can establish legally binding data governance contracts with users by drafting privacy policy statements, signing data security agreements, and refining confidentiality clauses. As for rights protection, there are three primary categories of rights claims by data subjects. First, data collection must adhere to the principle of informed consent, requiring explicit authorization from the user. Second, data subjects have the right to request the deletion of their information (i.e., the right to be forgotten). Third, during the commercial utilization of data, both the property rights and personal rights associated with the data must be safeguarded.

Based on the preceding analysis, we first construct a keyword lexicon for personal data protection, encompassing multidimensional indicators, drawing on existing literature related to personal data protection and the Personal Information Protection Law of the People's Republic of China (PIPL) (see Table A1 for further details) following Chen et al. (2024a). Unlike traditional compliance reviews, which focus on adherence to domestic information protection regulations, we emphasize the voluntary establishment of PDPS by firms. Consequently, the keyword selection incorporates both international legal texts and

industry technical standards rather than solely referencing domestic regulations.

We use Python to search for firms that may have established PDPS based on personal information protection keywords in the annual reports of listed firms. We then combined LLMs and manual cross-checking to accurately confirm the specific year in which the firm established the PDPS. We constructed a binary variable $PDPS$ (the value is 1 in the year in which the firm established the PDPS and in subsequent years, otherwise it is assigned a value of 0) to quantify and evaluate whether the firms voluntarily established the PDPS.

3.2.2 Corporate Employment

Our primary dependent variable is the corporate employment level (lnEmploy). In line with the approaches of Liu et al. (2021) and Popov and Rocholl (2018), we define lnEmploy as the natural logarithm of the number of employees. This transformation allows for a clear interpretation of the firm's total employment at the corporate level.

3.3 Empirical Models

A staggered DID model with TWFE is used to assess the effect of PDPS on corporate employment, similar to Souillard (2022).

$$lnEmploy_{i,t} = \beta_0 + \beta_1 PDPS_{i,t-1} + \beta_2 Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{it} \tag{1}$$

where the subscripts i and t represent firm and year, respectively. The dependent variable is $lnEmoloy$. The key explanatory variable, $PDPS$, indicates whether the firm voluntarily establishes a PDPS. Following the existing literature (Ghaly et al., 2020; Yun et al., 2025; Jiang et al., 2025b; Li et al., 2024), we control for several firm-level factors related to firm employment, including firm size ($Size$), leverage ratio (Lev), return on assets (ROA), monetary fund ratio ($Cash$), sales revenue growth rate ($Growth$), market-to-book ratio (MB), firm age (Age), a dummy variable for whether the firm is a state-owned enterprise (SOE), a dummy variable for whether the chairman and CEO belong to the same person ($Dual$), the natural logarithm of the minimum wage standard in the county where the firm

is located ($lnMWage$), the natural logarithm of per capita GDP at the prefecture-level city ($lnperGDP$), the natural logarithm of the number of higher education institutions at the prefecture-level city ($lnSchool$), and the share of secondary industry in GDP at the prefecture-level city ($Indstru$). All the variables and definitions are presented in Table A2.

To control for firm-specific characteristics that do not vary over time, as well as external factors influencing the firms, we employ a model with both firm and year-fixed effects. Furthermore, to mitigate concerns regarding autocorrelation and heteroscedasticity, robust standard errors clustered at the firm level are used in the analysis.

4 Empirical Results

4.1 Summary Statistics

Table 1, Panel A presents descriptive statistics for firm-level variables. The median value of $lnEmploy$ is reported as 7.702, with a standard deviation of 1.269, ranging from a minimum of 7.646 to a maximum of 11.218. These metrics indicate substantial heterogeneity in employment levels across sampled firms, as evidenced by the pronounced dispersion in logged workforce size. The variable $PDPS$ exhibits a mean of 0.090 and a standard deviation of 0.286. Figure A1 illustrates the number and proportion of firms establishing PDPS each year. This pattern likely reflects the pre-2021 institutional environment, wherein a majority of listed corporations faced limited exigency for digital transformation, coupled with relatively underdeveloped consumer awareness of privacy protections. Furthermore, the establishment of PDPS entailed significant short-term cost burdens, which plausibly constrained voluntary adoption to a modest subset of firms. Panel B of Table 1 reports descriptive statistics for region-level control variables, with all measures aligning closely with established patterns documented in prior literature.

(Insert Table 1 about here)

4.2 PDPS and Corporate Employment

Table 2 presents the results of the staggered DID model. No control variables are included in Column (1), firm-level control variables are included in Column (2), regional-level control variables are further included in Column (3), the mechanism variables used further below are included in Column (4), and the moderating variables are included in Column (5) [3]. The estimated coefficient of $PDPS$ is significantly positive in all regressions, ranging from 0.0797 to 0.1432. Although dampened by adding potential mediators, these results indicate that compared with firms that have not established PDPS, the establishment of PDPS will increase the number of employees in the firms by 14.32% and after controlling other factors, the pure effect is not less than 8.00%. Considering that the average number of employees in the sample firms is 5,943, these estimates suggest that the average increase in headcount is at least 374 employees. This shows that PDPS has a strong economic effect on corporate employment.

(Insert Table 2 about here)

4.3 Robustness Tests

4.3.1 Robustness of Inference to Replacement (RIR) Analysis

To address concerns regarding potential endogeneity arising from omitted variable bias, we conduct a RIR analysis using the konfound command in Stata, following the methodology proposed by Xu et al. (2019). The RIR approach specifically assesses the sensitivity of our key inference to the possible influence of unobserved confounders. It quantifies the proportion of the estimated effect that would need to be replaced with bias from omitted variables in order to invalidate our statistically significant findings (Busenbark et al., 2022). The logic of this method is as follows: if a large portion of the observed effect must be attributable to omitted variables before the result loses statistical significance, then the inference can be

[3]Because listed firms only began disclosing the number of employees with different positions and educational backgrounds in 2011, data for $High$ and $Tech$ are only available after 2011, resulting in missing samples in Columns (4) and (5).

considered robust. Conversely, if only a small portion of the effect needs to be biased away, the result is more vulnerable to omitted variable bias.

Applying this analysis to the main effects model (Column (3) in Table 2), we find that **41.88%** (corresponding to **9,303** cases) of the estimated effect of PDPS would need to be replaced by bias from unobserved variables for the effect to lose statistical significance. This relatively high threshold indicates that our results are robust: it is unlikely that omitted variables alone could account for such a large proportion of the effect. Therefore, the conclusions drawn regarding PDPS remain reliable even in the presence of potential unobserved confounding factors.

4.3.2 Parallel Trend Test

We employ an event study approach to examine parallel trends and explore the evolution of treatment effects. We compute the estimates based on the following specifications:

$$\text{lnEmploy}_{i,t} = \alpha + \sum_{n=-8}^{6} \beta_n \text{PDPS}_{i,t-1}^{n} + \gamma \text{ Control }_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t}, \tag{2}$$

where $lnEmploy_{i,t}$ is the natural logarithm of the number of employees in a firm, $PDPS_{i,t}^{n}$ is the relative year variable of the establishment of PDPS by the firms. If firm i establishes PDPS in year t, it takes the value of 1, otherwise, it takes the value of 0. The year before the establishment of PDPS is taken as the base year for event analysis, β_n is the regression coefficient relative to the base year, and the control variables and fixed effects are consistent with the baseline regression model.

Figure 1 shows the results of the event study using two-way fixed effects (TWFE), indicating that the parallel trend assumption is met. Before the establishment of the PDPS, the estimated coefficient is insignificant, indicating that there is no difference in employment between the treatment group and the control group before the establishment of the PDPS. At the same time, in order to eliminate the interference of unobservable time-varying factors, we re-estimate using the Fixed Effects Counterfactual Estimators proposed by Liu et al. (2024)

[4]. Figure 2 shows that when using counterfactual estimation, our results still meet the parallel trend and establish that PDPS has a significant impact on corporate employment.

(Insert Figure 1 about here)

(Insert Figure 2 about here)

4.3.3 Placebo Test

By employing the methodology of Liu and Lu (2015), we conducted a placebo test to exclude the possibility that unobservable factors drive our results. We randomly selected firms from the sample as part of the fake treatment group and allocated fake treatment time. We repeated the test 500 times to obtain the distribution of the placebo effect. After analyzing the kernel density distribution of the estimated coefficients on the artificially generated treatment groups in Figure 3, a clear pattern emerged. The average value of the pseudo-regression coefficients within the placebo test clusters was close to 0, indicating a significant deviation from the true estimates of the primary explanatory factors. Additionally, the associated p-values were mostly greater than 10%, reinforcing the statistical significance of the results. This indicates that our results are caused by PDPS rather than other random factors.

4.3.4 Matching Technology

To further mitigate endogeneity concerns, we employ both propensity score matching (PSM) and entropy balancing methods. Firms that have adopted PDPS are designated as the treatment group, while the remaining firms serve as the control group.

For PSM, we implement 1:6 nearest neighbor matching based on a set of control variables. As illustrated in Figure A2, there are substantial differences in the kernel density curves between the treatment and control groups prior to matching, reflecting imbalance in covariate distributions. After PSM, the kernel density curves of the two groups overlap considerably, indicating that the matching procedure has effectively improved covariate balance between

[4]This method is also effective in alleviating the problem of negative estimated weights of staggered DID.

groups. The regression results based on the matched sample are presented in Column (1) of Table 3, where the coefficient on the variable $PDPS$ remains significantly positive, suggesting that our main findings are robust to this matching strategy.

In addition, we utilize entropy balancing, which addresses some limitations of PSM—such as loss of sample size and dependence on subjective matching choices—by reweighting observations to achieve exact balance on the means and higher-order moments of the covariates (Hainmueller, 2012). Table A3 demonstrates that, following entropy balancing, the covariate distributions between the treatment and control groups are statistically indistinguishable. The regression results using the entropy-balanced sample are shown in Column (2) of Table 3, and once again, the coefficient on $PDPS$ remains significantly positive.

These results confirm that our main findings are not driven by observable differences between the treatment and control groups, thereby further strengthening the credibility of our empirical conclusions.

(Insert Table 3 about here)

4.3.5 Specification Curve Analysis

Following the method of Gao et al. (2021), we conduct specification curve analysis to verify the robustness of PDPS to various model specifications. We further add industry-year high-dimensional fixed effects and city-year high-dimensional fixed effects to the model to control the interference of time-varying factors at the industry and region levels. At the same time, we perform double clustering of firm-year and triple clustering of industry-city-year on standard errors for more robust estimation (Cameron et al., 2011). Figure 4 shows that our results are still robust in 12 different model settings.

(Insert Figure 4 about here)

4.3.6 Excluding Other Policy Influences

In this section, we exclude the interference of several Chinese government policies that may have affected firm employment during the sample period. First, we consider the

Intelligent Manufacturing (IM) pilot policy. From 2015 to 2018, the Chinese government identified a group of firms as IM pilot firms each year, which means that these firms have set an example in the use of IM technology. The use of IM technology may also mean the need for relevant technical talents and production personnel. We add a dummy variable *IM*, which is assigned a value of 1 when a firm is selected for the IM pilot and in subsequent years, and 0 otherwise, following Zhang et al. (2024); Bai et al. (2025). Second, we consider the government's employment subsidy policy for firms (known as "Wen Gang Bu Tie" in Chinese). In order to cope with the pressure of social unemployment, the Chinese government provides employment subsidies to firms through fiscal subsidies. We add a dummy variable *Sub*, which is assigned a value of 1 in the year and the following years when a firm receives government employment subsidies, and 0 otherwise. Third, we consider the value-added tax (VAT) credit refunds policy. In 2018, China reformed its tax policy and implemented VAT refunds for firms in 18 industries. This tax incentive policy will allow firms to have more resources for employment. We add a dummy variable *Tax*, which is assigned a value of 1 in the year and the following years when a firm is eligible for tax incentives, and 0 otherwise, following Zhou et al. (2022). Finally, we consider China's local government debt cleanup campaign. In 2016, China's central government required local governments to repay funds owed to firms and launched special inspections in seven provinces. When a firm collects what it is owed, it will also make it more able to hire. We add a dummy variable *Gov*, which is assigned a value of 1 in the year and the following years when a firm benefits from the campaign, and 0 otherwise. Table 4 reports the results after excluding the above policy effects. The coefficient of $PDPS$ is always positive, indicating that the results are robust.

(Insert Table 4 about here)

4.3.7 Heckman Two-Stage Model and Instrumental Variable Regression

To address the potential endogeneity arising from certain industries or firms being more likely to establish PDPS due to their intrinsic characteristics, we employ both the Heckman two-stage model and instrumental variable (IV) regression. In both approaches, following Chen et al. (2024a), we use a single instrumental variable: the product of the proportion

of urban landline telephones in the firm's city in 1984 relative to the national total and the lagged number of national Internet ports. This variable is expected to be strongly correlated with PDPS adoption but should not directly influence firms' employment decisions. The historical number of landline telephones reflects the region's communication infrastructure, which is relevant for subsequent digital technology adoption, thus satisfying the relevance criterion. In the Heckman model, we first estimate the probability of PDPS adoption using a Probit model with this instrument, calculate the Inverse Mills Ratio (IMR), and then include the IMR in the outcome regression to correct for selection bias [5]. In the IV regression, we use the same instrument to address endogeneity in PDPS adoption [6]. The results, reported in Columns (1)–(2) of Table 5, demonstrate that our findings remain robust under both approaches.

4.3.8 Augmented DID Model

Given that the establishment of PDPS by firms may be influenced by their own characteristics, leading to non-random assignment of treatment, we further mitigate this concern by employing an augmented DID approach, following Li et al. (2016). Specifically, we incorporate a set of pre-treatment variables from 2007 (Z_{2007}), the year prior to our research window, and interact them with year fixed effects (γ_t) to flexibly control for the time-varying impact of these covariates on the outcome variable. The estimation equation is as follows:

$$\text{lnEmploy}_{i,t} = \beta_0 + \beta_1 \text{PDPS}_{i,t-1} + \eta Z_{2007} \times \gamma_t + \phi \text{Control}_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \tag{3}$$

Z_{2007} is a series of variables that may affect the establishment of PDPS by firms in 2007 (the year before the research window) γ_t is the year dummies. The Z_{2007} has interacted with the year dummies γ_t, which more flexibly control for the time effects of Z_{2007} on the

[5]The estimated coefficient of IV in the first stage Probit model is 0.0865 (P=0.000).

[6]The estimated coefficient of IV in the first stage is 0.0094 (P=0.003). The Kleibergen-Paap rk LM statistic is 8.673 (Chi-sq(1) P-val = 0.0032. The Cragg-Donald Wald F-statistic is 56.476, and the Stock-Yogo weak ID test critical values: 10% maximal IV size is 16.38. This indicates that our IV are effective.

outcome variables. This approach helps to further address potential confounding by baseline characteristics that might influence both PDPS adoption and employment outcomes. The result is reported in Column (3) of Table 5. Our findings remain.

4.3.9 Heterogeneity-robust DID methods

To tackle the bias associated with two-way fixed effects in staggered DID designs, as highlighted by Goodman-Bacon (2021) and De Chaisemartin and d'Haultfoeuille (2020), we re-estimate our baseline model using the csdid command proposed by Callaway and Sant'Anna (2021) and the did-imputation command proposed by Borusyak et al. (2024). These advanced estimation techniques account for treatment timing heterogeneity and provide more accurate causal effect estimates under staggered adoption settings. The result is reported in Columns (4)–(5) of Table 5, showing the robustness of our findings.

4.3.10 Double Machine Learning DID

To further address potential model specification errors and biases in variable selection, we implement the DML-DID approach, as proposed by Bodory et al. (2022) and Chang (2020). DML-DID leverages machine learning methods to flexibly control for high-dimensional covariates and model complexities, providing robust causal inference in panel data settings. The result is reported in Column (6) of Table 5.

(Insert Table 5 about here)

5 Moderation Effect Tests

5.1 Market Power

Market power may have an impact on the employment effects of PDPS implementation. Monopolistic firms have substantial pricing power, and sustained excess profits generate financial slack that enables absorption of the considerable fixed costs associated with comprehensive privacy protection—including encryption technologies, anonymization protocols, and specialized personnel—without compromising other strategic priorities (Zhao and Zhang,

2025). This financial capacity allows monopolistic firms to implement more ambitious privacy initiatives that require substantial workforce expansion. The market dominance of monopolistic firms also attracts heightened attention from regulatory authorities who impose more stringent compliance requirements to prevent exploitation of market advantage at consumers' expense (Depoorter, 1999; Marino, 1988). This regulatory scrutiny creates a quasi-liability regime where monopolistic firms face disproportionately severe penalties for privacy failures compared to firms with limited market power. In contrast, non-monopolistic firms face intense competitive pressure and thin profit margins, which restrict their ability to absorb the substantial fixed costs of comprehensive privacy programs, forcing them toward minimalist implementations that leverage existing personnel rather than expanding headcount. Their limited market visibility reduces both regulatory attention and reputational consequences from privacy failures, weakening incentives for preventive investments in human capital.

We use the Lerner Index to measure the market power of firms, and classify firms in the top third as high market power firms, and the rest as low market power firms, thereby generating a dummy variable LI. Column (1) of Table 6 reports the moderating effect of market power. The interaction between LI and $PDPS$ is significantly positive, indicating that firms with higher market power will hire more employees in the process of establishing PDPS.

5.2 Financing Constraints

The relationship between PDPS implementation and employment outcomes is significantly moderated by a firm's financial constraints, as resource availability fundamentally shapes both the capacity and strategy for privacy-related investments. Financial constraints represent friction in capital markets that limit a firm's ability to fund value-enhancing projects through external financing (Kaplan and Zingales, 1997; Whited and Wu, 2006). For firms facing relatively low financial constraints, they can absorb the substantial upfront investments required for a comprehensive privacy infrastructure—including technical systems, governance frameworks, and specialized personnel, without compromising other

strategic priorities (Chen, 2023). Their superior access to capital markets also enables them to rapidly scale privacy initiatives beyond minimum viable implementations, creating specialized roles that require dedicated staffing rather than additional responsibilities for existing employees (Campello et al., 2010). By creating dedicated privacy teams and integrating privacy considerations into cross-functional processes, these companies can enhance their overall privacy capabilities and complement existing businesses, enhancing operational efficiency and long-term competitiveness (Mastromarco and Ghosh, 2009; Liu and Zhang, 2025; Chen et al., 2024b). Financially unconstrained companies have more resources and motivation to use privacy protection as a competitive advantage and voluntarily establish PDPS. In stark contrast, financially constrained firms face binding capital. When these firms implement PDPS, they tend to redistribute responsibilities among existing technical personnel and invest in automation to minimize human capital requirements without creating substantial new positions (Hadlock and Pierce, 2010).

We use the KZ index to measure the financial constraints of firms, and we also regard the first third of firms as high-financing constraint firms. Column (2) of Table 6 reports the moderating effect of financing constraints. The interaction coefficient of KZ and $PDPS$ is significantly negative, indicating that higher financing constraints will significantly reduce the number of employees employed by firms in the process of establishing PDPS.

5.3 Employee Wages

The moderating effect of the payroll-to-revenue ratio can also be explained through the lens of labor market competition and human capital theory. Firms with a higher payroll-to-revenue ratio often possess stronger capabilities or strategic priorities to offer competitive compensation and benefits, making them more attractive to skilled professionals, especially in specialized domains such as data privacy and compliance (Murphy, 1985).

When implementing PDPS, these firms can leverage their wage advantage to attract and retain the expertise required for privacy governance, technical implementation, and regulatory compliance. The ability to offer higher wages not only expands the pool of potential applicants but also enhances employee motivation and reduces turnover, both

of which are critical when new and complex roles must be filled (Lazear, 2000). As a result, firms with higher payroll-to-revenue ratios are better positioned to increase hiring in response to PDPS requirements and to quickly build dedicated, high-performing privacy teams. Moreover, high payroll-to-revenue ratios often reflect an organizational culture that values and invests in human capital, fostering continuous learning and upskilling opportunities. Such firms are more likely to invest in professional development programs and certifications in data protection, further strengthening their ability to respond effectively to regulatory changes and technological advancements in privacy management (Lazear and Shaw, 2007). This emphasis on human capital development creates a virtuous cycle: as firms become known as desirable employers in the privacy and compliance sector, they attract even more qualified talent, reinforcing their organizational capacity to implement PDPS at a higher standard. In contrast, firms with a lower payroll-to-revenue ratio may face greater difficulty recruiting qualified personnel, either due to budget constraints or less attractive compensation packages. This can limit their ability to respond to PDPS-related labor demands through new hiring, forcing them to rely instead on internal redeployment or automation.

We measure the employee wages of firms by the ratio of employee wages payable to sales revenue, and consider the top third of enterprises as high-wage firms. Column (3) of Table 6 reports the moderating effect of employee wages. The interaction coefficient between $Wage$ and $PDPS$ is significantly positive, indicating that higher employee wages significantly increase the number of employees hired by firms in the process of establishing PDPS.

(Insert Table 6 about here)

6 Mechanism Analysis

We introduce an instrumental mediation analysis, adapted from Di Giuli and Laux (2022). The core idea is that PDPS (Privacy Data Protection Systems) may influence firm employment outcomes through mediating variables. Inspired by standard instrumental

variable (IV) regression frameworks, we employ lagged PDPS ("$PDPS_{i,t-1}$") as an instrument to capture variation in mediators driven specifically by PDPS. Specifically, we examine whether this instrumented variation significantly affects firm employment in the subsequent period, thereby illustrating the chain effect through which PDPS adoption impacts hiring decisions via mediators.

Our approach diverges from Di Giuli and Laux (2022) in an important way. In the first stage, we regress the mediating variable solely on the lagged PDPS adoption variable ("$PDPS_{i,t-1}$") in the cross section, intentionally excluding any control variables or fixed effects. This specification ensures that we isolate the variation in the mediator that is exclusively driven by the PDPS adoption, eliminating the confounding influences of other covariates. In the second stage, we regress firm employment on the fitted values of the mediator (obtained from the first stage) while including relevant control variables and fixed effects. This two-stage approach yields a first-stage estimate of the isolated impact of PDPS on the mediator and a second-stage estimate of the effect of PDPS-driven mediator variation on firm employment outcomes.

The formal specification of the model is as follows.

First Stage:

$$M_{i,t} = \alpha_0 + \alpha_1 PDPS_{i,t-1} + \varepsilon_{it} \tag{4}$$

Second Stage:

$$\ln Employ_{i,t} = \gamma_0 + \gamma_1 \widehat{M_{i,t}} + \gamma_2 Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{it} \tag{5}$$

As discussed above, we focus on the estimates of α_1 and γ_1. Statistically significant estimates of $\hat{\alpha}_1$ and $\hat{\gamma}_1$ provide support for the hypothesized mediation mechanism.

6.1 R&D Investment Channel

The mechanism linking the voluntary adoption of PDPS to employment growth operates through a critical intermediary channel: sustained increases in Research and Development

(R&D) investment. Fundamentally, the adoption of PDPS reflects an firm's strategic commitment to proactively embrace technological advancements in privacy protection. This commitment stems from the necessity to address growing consumer concerns around data privacy and security (Acquisti et al., 2016; Tsai et al., 2011), as well as the need to build trust in increasingly privacy-conscious markets (Martin and Murphy, 2017). However, the adoption and implementation of PDPS are not merely technical compliance exercises; they require substantial and ongoing investments in R&D to develop and refine privacy-enhancing technologies such as advanced encryption methods, data anonymization tools, and secure protocols (Goldfarb and Tucker, 2011). Firms seeking to differentiate themselves in this domain are compelled to allocate significant resources to R&D activities, embedding innovation into their broader competitive strategies (Adner, 2017).

R&D investment, significantly, is not simply a financial allocation but, more critically, a form of human capital investment (Pyatt, 1966). The development of next-generation privacy technologies necessitates the recruitment and retention of highly skilled professionals, including scientists, engineers, and other technical experts specializing in data security, cryptography, and systems integration. As Hall and Rosenberg (2010) highlights, over 50% of R&D expenditures typically consist of wages for highly educated technical personnel. Beyond the direct recruitment of technical personnel, firms also require supporting roles to manage, coordinate, and commercialize their R&D efforts. This includes project managers, data analysts, and market specialists who ensure that privacy-enhancing technologies are not only developed but also aligned with consumer needs and regulatory requirements. For firms that voluntarily adopt PDPS, the need to strengthen their internal capabilities in privacy-related R&D is particularly pronounced, given the technical complexity and rapidly evolving nature of privacy technologies. Consequently, the increase in R&D expenditure directly drives higher demand for skilled labor (Aral et al., 2012), effectively expanding employment opportunities in technical and innovation-oriented roles (Acemoglu and Restrepo, 2019; Jiang et al., 2025a).

Our analysis employs three metrics to gauge corporate R&D investment: the ratio of expensed and capitalized R&D expenditures to total assets ($R\&D$), the proportion of

technical personnel ($Tech$), and the share of employees with bachelor qualifications ($High$). Panel A of Table 7 shows that the lagged indicator of the PDPS adoption does increase R&D investment based on different measures. The Panel B in Table 7 indicates that R&D growth driven solely by PDPS significantly promotes enterprise employment. Among them, the regression coefficients of the fitted values of the two mediating factors - $Tech$ (the proportion of technical personnel) and $High$ (the proportion of employees with a bachelor's degree or above) - are significantly positive, and the direction is opposite to the benchmark regression (Table 2). This suggests that the negative effects of $Tech$ and $High$ identified in the benchmark regression may reflect budget constraints faced by firms that have not adopted PDPS. In contrast, the two-stage estimation method employed here reveals that, following PDPS adoption, the expansion of technical personnel positively contributes to overall employment. This finding underpins the second channel discussed below—specifically, that PDPS adoption enhances firm operational performance, thereby alleviating budget constraints for hiring technical staff.

(Insert Table 7 about here)

6.2 Operating Performance Channel

The voluntary establishment of PDPS by enterprises signals a strategic commitment to responsible data governance and stakeholder-centric value creation, fostering enhanced trust relationships that translate into measurable business performance improvements (Martin and Murphy, 2017). First, PDPS generates significant reputational capital (Lange et al., 2011), which as an intangible strategic asset, directly enhances enterprise valuation and sustainable growth capacity (Roberts and Dowling, 2002). The privacy-protective signaling effect attracts privacy-conscious consumers who demonstrate willingness to pay premium prices for enhanced data protection (Liu et al., 2023; Acquisti et al., 2016), thereby expanding market share and accelerating revenue growth rates. Second, from an operational efficiency perspective, PDPS implementation streamlines compliance processes, reduces information asymmetries, and strengthens internal control systems (Gordon et al., 2009).

This systematic enhancement minimizes operational frictions, shortens business cycles, and creates organizational slack for strategic workforce investments. Third, PDPS serves as a risk management mechanism that substantially reduces the probability and severity of data breach incidents, thereby mitigating regulatory penalties and reputational damages that typically generate significant sales volatility (Nikkhah and Grover, 2022).

These multidimensional performance improvements create both the financial capacity and strategic imperative for employment expansion. Enhanced sustainable growth rates provide the revenue foundation necessary for workforce scaling, while shortened operational cycles release working capital constraints that previously limited hiring decisions (Foster et al., 2001; Brynjolfsson and McElheran, 2016). Simultaneously, reduced sales volatility decreases earnings uncertainty and cash flow variability, enabling firms to make confident long-term employment commitments even during periods of market turbulence (Bachmann et al., 2013).

We evaluate corporate performance through three dimensions: sustainable growth rate (SGR), operating cycle (OC), and sales volatility ($SalesFlu$) [7]. Similarly, we adopt the two-step estimation method for regression. The results of the first stage are shown in Panel A of Table 8. The establishment of PDPS significantly increased the sustainable growth rate of the enterprise and reduced the management rate of the business cycle and the sales volatility. The results of the second stage show that except for the regression coefficient of $\widehat{SGR}$ which is positive but not significant, the mediating fitting values ($\widehat{OC}$, $\widehat{SalesFlu}$) are all in line with the expected direction and significance. This further verifies that PDPS improves the employment level by improving enterprise performance.

(Insert Table 8 about here)

[7] Following Ghosh and Olsen (2009), we measure firm-level sales volatility ($SalesFlu$) by first computing unadjusted sales volatility as the standard deviation of abnormal sales revenue over the preceding three-year period divided by the corresponding three-year average sales revenue; Industry sales volatility is defined as the median of all firms' unadjusted sales volatility within the same industry-year cohort. Firm-specific SalesFlu is then derived by normalizing each firm's unadjusted sales volatility against its industry benchmark.

7 Conclusion

This study investigates the impact of voluntarily established PDPS on corporate employment, a topic that remains underexplored in the literature on privacy regulations and labor economics. While prior research has emphasized the macroeconomic effects of mandatory data protection laws such as the GDPR, little attention has been given to the firm-level consequences of voluntary data governance. Understanding whether and how firms that proactively implement PDPS experience workforce expansion is crucial in an era where digital transformation and data security are central to business strategy.

We find that firms that voluntarily establish PDPS experience a significant increase in employment levels. Robustness tests confirm the validity of our findings. The positive employment effect is more pronounced for firms with strong market power, lower financial constraints, and less corporate diversification. Further analysis identifies two key mechanisms through which PDPS influences employment: increased R&D investment and improved operating performance.

Our findings hold important implications for policymakers, particularly in balancing privacy protection with economic growth. First, given that voluntary PDPS adoption can stimulate employment, regulators should design policies that encourage firms to implement privacy measures beyond minimum legal requirements. This could involve offering tax incentives, government grants, or certification programs that recognize firms with strong data governance practices, thereby fostering a market-driven approach to privacy protection. Second, our study suggests that industry-specific privacy regulations may be more effective than uniform, one-size-fits-all policies. Some industries, particularly those with high consumer data sensitivity (e.g., fintech, healthcare), may require stronger regulatory oversight, while others may benefit from flexible, incentive-based approaches. Policymakers should consider differentiated regulatory frameworks that account for sectoral variations in data usage, compliance costs, and competitive dynamics. Third, our findings highlight the potential for privacy protection policies to contribute to labor market stability. Current regulatory debates often focus on the trade-off between privacy compliance costs

and business efficiency, but our results suggest that strong data governance can also create jobs, particularly in technology and compliance-related roles. Governments should integrate privacy protection into broader employment and industrial policies, ensuring that digital security measures align with workforce development strategies.

Despite its contributions, this study has several limitations that suggest directions for future research. While we establish a causal link between PDPS adoption and employment growth, we do not examine long-term labor market effects, such as changes in wage structures, skill demand, or job quality. Future research could explore whether firms that implement PDPS experience sustained productivity gains or shifts in workforce composition over time. Understanding how privacy protection interacts with broader corporate strategies will be crucial as data governance becomes an increasingly central issue in the digital economy.

References

Acemoglu, D., Restrepo, P., 2019. Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives 33, 3–30.

Acemoglu, D., Restrepo, P., 2020. Robots and jobs: Evidence from us labor markets. Journal of Political Economy 128, 2188–2244.

Acquisti, A., Taylor, C., Wagman, L., 2016. The economics of privacy. Journal of Economic Literature 54, 442–492.

Adner, R., 2017. Ecosystem as structure: An actionable construct for strategy. Journal of Management 43, 39–58.

Agrawal, A., Gans, J., Goldfarb, A., 2019. The economics of artificial intelligence: An agenda. University of Chicago Press.

Aral, S., Brynjolfsson, E., Wu, L., 2012. Three-way complementarities: Performance pay, human resource analytics, and information technology. Management Science 58, 913–931.

Aridor, G., Che, Y.-K., Salz, T., 2020. The economic consequences of data privacy regulation: Empirical evidence from GDPR. National Bureau of Economic Research Cambridge, MA, USA.

Aridor, G., Che, Y.-K., Salz, T., 2021. The effect of privacy regulation on the data industry: Empirical evidence from gdpr. In: *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 93–94.

Autor, D., Chin, C., Salomons, A., Seegmiller, B., 2024. New frontiers: The origins and content of new work, 1940–2018. The Quarterly Journal of Economics 139, 1399–1465.

Autor, D. H., 2015. Why are there still so many jobs? the history and future of workplace automation. Journal of Economic Perspectives 29 3–30.

Bachmann, R., Elstner, S., Sims, E. R., 2013. Uncertainty and economic activity: Evidence from business survey data. American Economic Journal: Macroeconomics 5, 217–249.

Bai, X., Zhou, K. Z., Jiang, W., Li, Y., Chen, X., 2025. Smart manufacturing and supply chain resilience: Evidence from an emerging market. Journal of Business Research 196, 115421.

Balsmeier, B., Woerter, M., 2019. Is this time different? how digitalization influences job creation and destruction. Research Policy 48 103765.

Barney, J., 1991. Firm resources and sustained competitive advantage. Journal of Management 17, 99–120.

Baron, R. M., Kenny, D. A., 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 51, 1173.

Behaghel, L., Moschion, J., 2016. It-based technical change and job instability. The Scandinavian Journal of Economics 118, 79–104.

Bessen, J. E., Impink, S. M., Reichensperger, L., Seamans, R., 2023. The business of ai startups. Boston Univ. School of Law, Law and Economics Research Paper .

Blanas, S., Gancia, G., Lee, S. Y., 2019. Who is afraid of machines? Economic Policy 34 627–690.

Blind, K., Niebel, C., Rammer, C., 2024. The impact of the eu general data protection regulation on product innovation. Industry and Innovation 31, 311–351.

Bodory, H., Huber, M., Lafférs, L., 2022. Evaluating (weighted) dynamic treatment effects by double machine learning. The Econometrics Journal 25, 628–648.

Borusyak, K., Jaravel, X., Spiess, J., 2024. Revisiting event-study designs: robust and efficient estimation. Review of Economic Studies 91, 3253–3285.

Brynjolfsson, E., Li, D., Raymond, L., 2025. Generative ai at work. The Quarterly Journal of Economics p. qjae044.

Brynjolfsson, E., McElheran, K., 2016. The rapid adoption of data-driven decision-making. American Economic Review 106, 133–139.

Brynjolfsson, E., Rock, D., Syverson, C., 2019. Artificial intelligence and the modern productivity paradox. The economics of artificial intelligence: An agenda 23, 23–57.

Busenbark, J. R., Yoon, H., Gamache, D. L., Withers, M. C., 2022. Omitted variable bias: Examining management research with the impact threshold of a confounding variable (itcv). Journal of Management 48, 17–48.

Caggese, A., Cuñat, V., 2008. Financing constraints and fixed-term employment contracts. The Economic Journal 118, 2013–2046.

Callaway, B., Sant'Anna, P. H., 2021. Difference-in-differences with multiple time periods. Journal of Econometrics 225, 200–230.

Cameron, A. C., Gelbach, J. B., Miller, D. L., 2011. Robust inference with multiway clustering. Journal of Business & Economic Statistics 29, 238–249.

Camina, E., Diaz-Chao, A., Torrent-Sellens, J., 2020. Automation technologies: Long-term effects for spanish industrial firms. Technological Forecasting and Social Change 151, 119828.

Campello, M., Graham, J. R., Harvey, C. R., 2010. The real effects of financial constraints: Evidence from a financial crisis. Journal of Financial Economics 97, 470–487.

Cavusoglu, H., Mishra, B., Raghunathan, S., 2004. The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. International Journal of Electronic Commerce 9, 70–104.

Chang, N.-C., 2020. Double/debiased machine learning for difference-in-differences models. The Econometrics Journal 23, 177–191.

Chen, W., 2023. No weal without woe: Implementation of personal data protection systems and corporate value. Revista de Administração de Empresas 63, e2022–0166.

Chen, W., Wang, Y., Wu, D., Yin, X., 2024a. Can the establishment of a personal data protection system promote corporate innovation? Research Policy 53 105080.

Chen, X., Mu, J., Liu, G., Liu, Y., 2024b. Bank liability structure and corporate employment: Evidence from a quasi-natural experiment in China. International Review of Financial Analysis 95, 103366.

Cirillo, V., 2014. Patterns of innovation and wage distribution. do "innovative firms" pay higher wages? evidence from chile. Eurasian Business Review 4 181–206.

Dachs, B., Peters, B., 2014. Innovation, employment growth, and foreign ownership of firms: A european perspective. Research Policy 43, 214–232.

De Chaisemartin, C., d'Haultfoeuille, X., 2020. Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review 110, 2964–2996.

De Vries, G. J., Gentile, E., Miroudot, S., Wacker, K. M., 2020. The rise of robots and the fall of routine jobs. Labour Economics 66, 101885.

Demirer, M., Hernández, D. J. J., Li, D., Peng, S., 2024. Data, privacy laws and firm production: Evidence from the GDPR. Tech. rep., National Bureau of Economic Research.

Depoorter, B. W., 1999. Regulation of natural monopoly. WF Deporter.–1999.–532 p .

Di Giuli, A., Laux, P. A., 2022. The effect of media-linked directors on financing and external governance. Journal of Financial Economics 145, 103–131.

Edler, D., Ribakova, T., 1994. The impact of industrial robots on the level and structure of employment in germany—a simulation study for the period 1980–2000. Technological Forecasting and Social Change 45, 255–274.

Flammer, C., 2013. Corporate social responsibility and shareholder reaction: The environmental awareness of investors. Academy of Management Journal 56, 758–781.

Foster, L., Haltiwanger, J. C., Krizan, C. J., 2001. Aggregate productivity growth: Lessons from microeconomic evidence. In: *New developments in productivity analysis*, University of Chicago Press, pp. 303–372.

Frey, C. B., Osborne, M. A., 2017. The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change 114 254–280.

Gallipoli, G., Makridis, C. A., 2018. Structural transformation and the rise of information technology. Journal of Monetary Economics 97, 91–110.

Gao, M., Leung, H., Qiu, B., 2021. Organization capital and executive performance incentives. Journal of Banking & Finance 123, 106017.

Gentili, A., Compagnucci, F., Gallegati, M., Valentini, E., 2020. Are machines stealing our jobs? Cambridge Journal of Regions Economy and Society 13, 153–173.

Ghaly, M., Dang, V. A., Stathopoulos, K., 2020. Institutional investors' horizons and corporate employment decisions. Journal of Corporate Finance 64, 101634.

Ghosh, D., Olsen, L., 2009. Environmental uncertainty and managers' use of discretionary accruals. Accounting, Organizations and Society 34, 188–205.

Goldfarb, A., Que, V. F., 2023. The economics of digital privacy. Annual Review of Economics 15, 267–286.

Goldfarb, A., Tucker, C. E., 2011. Privacy regulation and online advertising. Management Science 57, 57–71.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. Journal of Econometrics 225, 254–277.

Gordon, L. A., Loeb, M. P., Tseng, C.-Y., 2009. Enterprise risk management and firm performance: A contingency perspective. Journal of Accounting and Public Policy 28, 301–327.

Greenhalgh, C., Mavrotas, G., 1996. Job training, new technology and labour turnover. British Journal of Industrial Relations 34, 131–150.

Hadlock, C. J., Pierce, J. R., 2010. New evidence on measuring financial constraints: Moving beyond the kz index. The Review of Financial Studies 23, 1909–1940.

Hainmueller, J., 2012. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. Political Analysis 20, 25–46.

Hall, B. H., Rosenberg, N., 2010. Handbook of the Economics of Innovation, vol. 1. Elsevier.

Harrigan, J., Reshef, A., Toubal, F., 2021. The march of the techies: Job polarization within and between firms. Research Policy 50, 104008.

Higgins, R. C., 1977. How much growth can a firm afford? Financial Management pp. 7–16.

Hui X., Reshef, O., Zhou, L., 2024. The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. Organization Science 35, 1977–1989.

Jia, J., Jin, G. Z., Wagman, L., 2021. The short-run effects of the general data protection regulation on technology venture investment. Marketing Science 40, 661–684.

Jiang, W., Tang, Y., Xiao, R. J., Yao, V., 2025a. Surviving the fintech disruption. Journal of Financial Economics 171, 104071.

Jiang, Y., Zhang, X., Yao, S., 2025b. On ESG and corporate employment decision: Evidence from Chinese listed firms in 2009–2022. Economic Analysis and Policy 85, 854–869.

Johnson, G. A., Shriver, S. K., Goldberg, S. G., 2023. Privacy and market concentration: intended and unintended consequences of the GDPR. Management Science 69, 5695–5721.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., Stulz, R. M., 2021. Risk management, firm reputation, and the impact of successful cyberattacks on target firms. Journal of Financial Economics 139, 719–749.

Kaplan, S. N., Zingales, L., 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? The Quarterly Journal of Economics 112 169–215.

Lachenmaier, S., Rottmann, H., 2011. Effects of innovation on employment: A dynamic panel analysis. International Journal of Industrial Organization 29, 210–220.

Lange, D., Lee, P. M., Dai, Y., 2011. Organizational reputation: A review. Journal of Management 37, 153–184.

Lazear, E. P., 2000. Performance pay and productivity. American Economic Review 90, 1346–1361.

Lazear, E. P., Shaw, K. L., 2007. Personnel economics: The economist's view of human resources. Journal of Economic Perspectives 21, 91–114.

Leiponen, A., 2005. Skills and innovation. International Journal of Industrial Organization 23, 303–323.

Li, P., Lu, Y., Wang, J., 2016. Does flattening government improve economic performance? evidence from China. Journal of Development Economics 123 18–37.

Li, X., Liu, Z., Ye, Y., 2024. Public data and corporate employment: Evidence from the launch of Chinese public data platform. Economic Analysis and Policy 84, 124–144.

Liu, A. Y., 2025. Trade secrets protection and employment of public firms: Evidence from the uniform trade secrets act. Research Policy 54, 105243.

Liu, G., Liu, Y., Ye, Y., Zhang, C., 2021. Collateral menus and corporate employment: Evidence from China's property law. Journal of Economic Behavior & Organization 189, 686–709.

Liu, H., Li, K., Chen, Y., Luo, X. R., 2023. Is personally identifiable information really more valuable? evidence from consumers' willingness-to-accept valuation of their privacy information. Decision Support Systems 173 114010.

Liu, L., Wang, Y., Xu, Y., 2024. A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data. American Journal of Political Science 68, 160–176.

Liu, Q., Lu, Y., 2015. Firm investment and exporting: Evidence from China's value-added tax reform. Journal of International Economics 97, 392–403.

Liu, Y., Zhang, Z., 2025. Unlocking operational efficiency: How AI human capital investment enhances data processing efficiency? Economics Letters 247 112147.

Marino, A. M., 1988. Monopoly, liability and regulation. Southern Economic Journal pp. 913–927.

Martin, K. D., Borah, A., Palmatier, R. W., 2017. Data privacy: Effects on customer and firm performance. Journal of Marketing 81, 36–58.

Martin, K. D., Kim, J. J., Palmatier, R. W., Steinhoff, L., Stewart, D. W., Walker, B. A., Wang, Y., Weaven, S. K., 2020. Data privacy in retail. Journal of Retailing 96, 474–489.

Martin, K. D., Murphy, P. E., 2017. The role of data privacy in marketing. Journal of the Academy of Marketing Science 45, 135–155.

Mastromarco, C., Ghosh, S., 2009. Foreign capital, human capital, and efficiency: A stochastic frontier analysis for developing countries. World Development 37, 489–502.

Maurin, E., Thesmar, D., 2004. Changes in the functional structure of firms and the demand for skill. Journal of Labor Economics 22, 639–664.

McGuire, J. B., Sundgren, A., Schneeweis, T., 1988. Corporate social responsibility and firm financial performance. Academy of Management Journal 31, 854–872.

Miller, S. R., Eden, L., Li, D., 2020. Csr reputation and firm performance: A dynamic approach. Journal of Business Ethics 163, 619–636.

Mohnen, P., Röller, L.-H., 2005. Complementarities in innovation policy. European Economic Review 49, 1431–1450.

Murphy, K. J., 1985. Corporate performance and managerial remuneration: An empirical analysis. Journal of Accounting and Economics 7, 11–42.

Nickell, S., 1999. Product markets and labour markets. Labour Economics 6, 1–20.

Nikkhah, H. R., Grover, V., 2022. An empirical investigation of company response to data breaches. MIS Quarterly 46, 2163–2196.

Pacelli, L., Rapiti, F., Revelli, R., 1998. Employment and mobility of workers in industries with different intensity of innovation: Evidence on italy from a panel of workers and firms. Economics of Innovation and New Technology 5, 273–300.

Peukert, C., Bechtold, S., Batikas, M., Kretschmer, T., 2022. Regulatory spillovers and data governance: Evidence from the GDPR. Marketing Science 41, 746–768.

Popov, A., Rocholl, J., 2018. Do credit shocks affect labor demand? evidence for employment and wages during the financial crisis. Journal of Financial Intermediation 36 16–27.

Pyatt, G., 1966. Human capital: A theoretical and empirical analysis, with special reference to education the residual factor and economic growth econometric models of education.

Roberts, P. W., Dowling, G. R., 2002. Corporate reputation and sustained superior financial performance. Strategic Management Journal 23, 1077–1093.

Silva, H. C., Lima, F., 2017. Technology, employment and skills: A look into job duration. Research Policy 46, 1519–1530.

Souillard, B., 2022. Corporate tax cuts and firm employment: A match made in haven? Economics Letters 219 110835.

Spence, M., 1978. Job market signaling. In: *Uncertainty in economics*, Elsevier, pp. 281–306.

Syed, R., 2019. Enterprise reputation threats on social media: A case of data breach framing. The Journal of Strategic Information Systems 28, 257–274.

Tambe, P., Ye, X., Cappelli, P., 2020. Paying to program? engineering brand and high-tech wages. Management Science 66 3010–3028.

Tsai, J. Y., Egelman, S., Cranor, L., Acquisti, A., 2011. The effect of online privacy information on purchasing behavior: An experimental study. Information Systems Research 22, 254–268.

Van Reenen, J., 1997. Technological innovation and employment in a panel of british manufacturing firms. Journal of Labor Economics 15, 255–284.

Vivarelli, M., Evangelista, R., Pianta, M., et al., 1996. Innovation and employment in the italian manufacturing industry. Research Policy 25, 1013–1026.

Vu, K. M., 2013. Information and communication technology (ict) and singapore's economic growth. Information Economics and policy 25, 284–300.

Webb, M., 2019. The impact of artificial intelligence on the labor market. Available at SSRN 3482150 .

Whited, T. M., Wu, G., 2006. Financial constraints risk. The review of financial studies 19, 531–559.

Xu, R., Frank, K. A., Maroulis, S. J., Rosenberg, J. M., 2019. konfound: Command to quantify robustness of causal inferences. The Stata Journal 19, 523–550.

Yun, F., Wang, S., Ye, Y., 2025. Effect of corporate holding financial institutions on corporate employment: Evidence from China. International Review of Financial Analysis 97, 103860.

Zhang, L., Jin, X., Wu, K., 2024. How does common institutional investors adapt to technology shocks? Available at SSRN 5009715 .

Zhao Y., Zhang, N., 2025. Energy trade access and market monopoly: Evidence from China's power sector. World Development 185, 106797.

Zhou, Q., Li, T., Gong, L., 2022. The effect of tax incentives on energy intensity: Evidence from china's vat reform. Energy Economics 108, 105887.

Zhu, Y., Sun, L.-Y., Leung, A. S., 2014. Corporate social responsibility, firm reputation, and firm performance: The role of ethical leadership. Asia Pacific Journal of Management 31, 925–947.

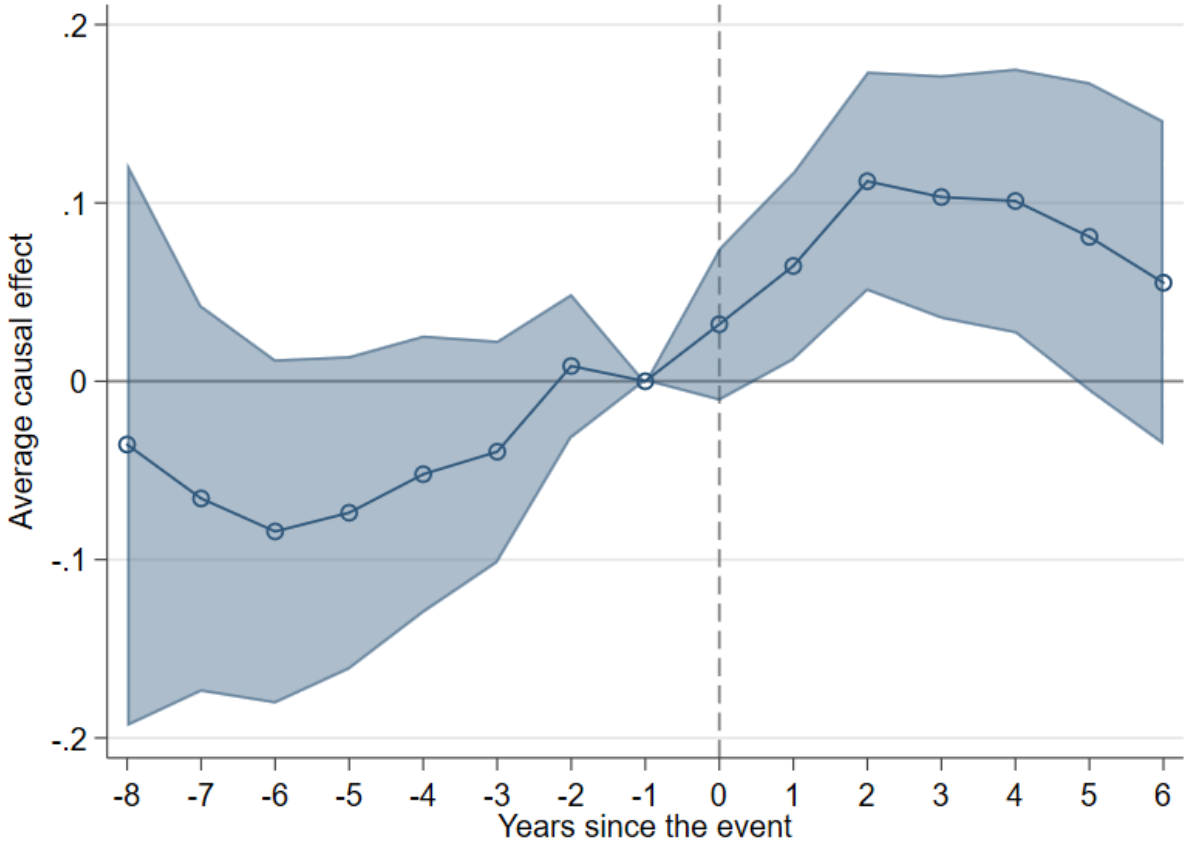


Figure 1: Event Study of Two-Way Fixed Effects

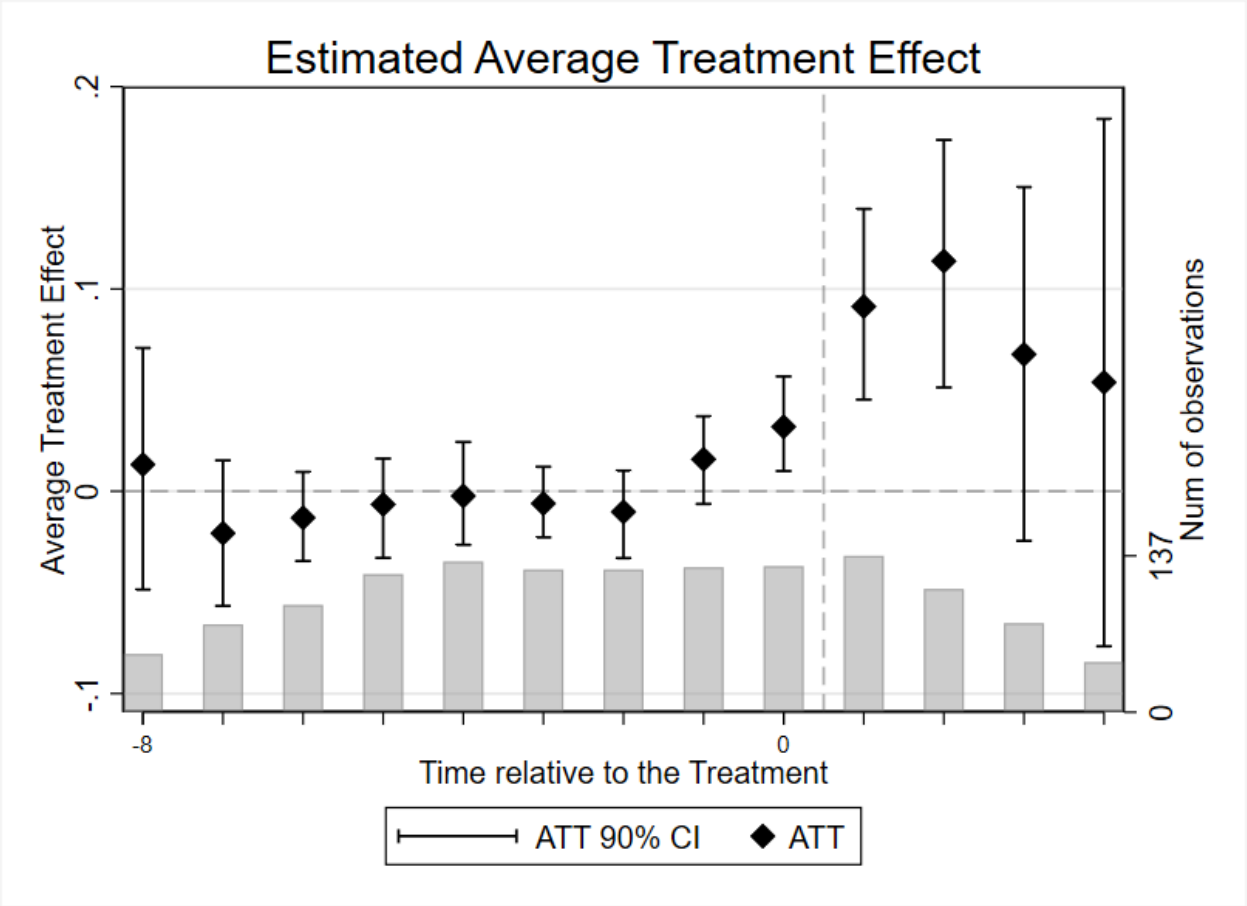


Figure 2: Fixed Effects Counterfactual Estimators

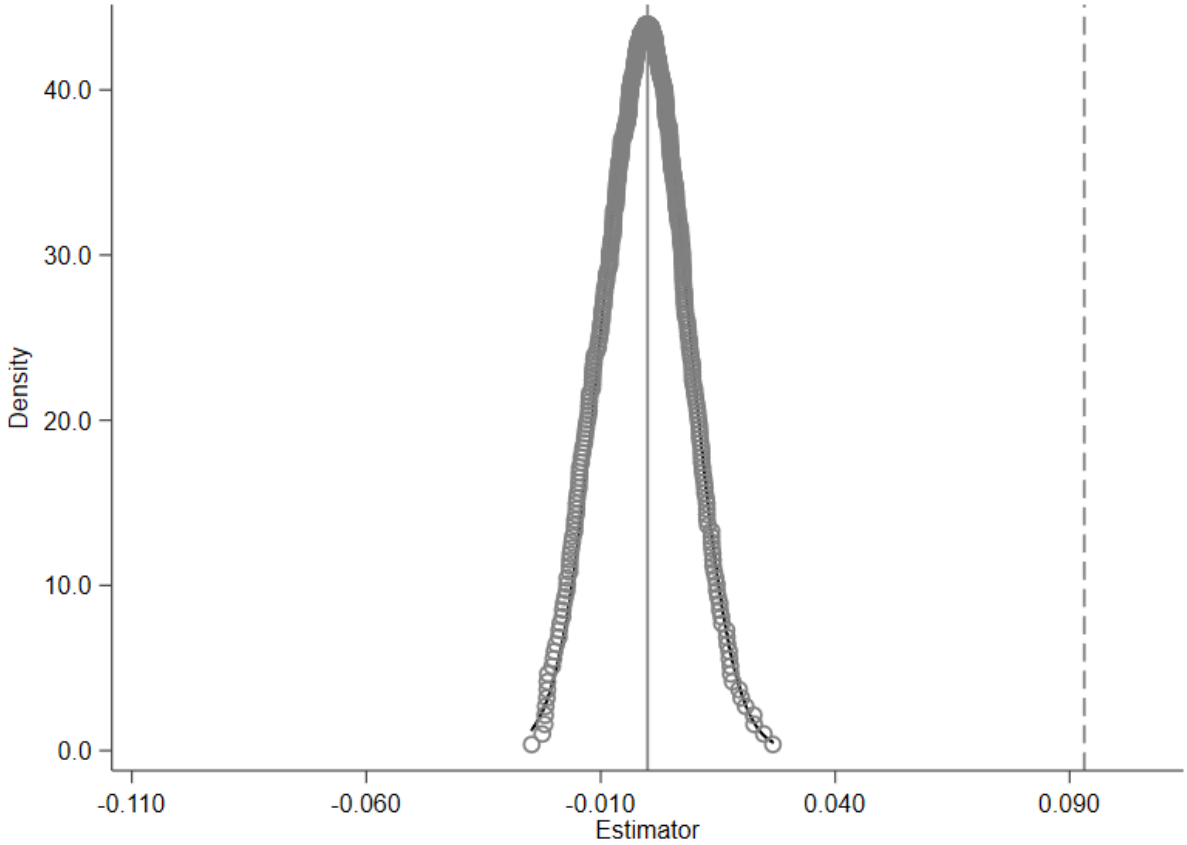


Figure 3: Placebo test

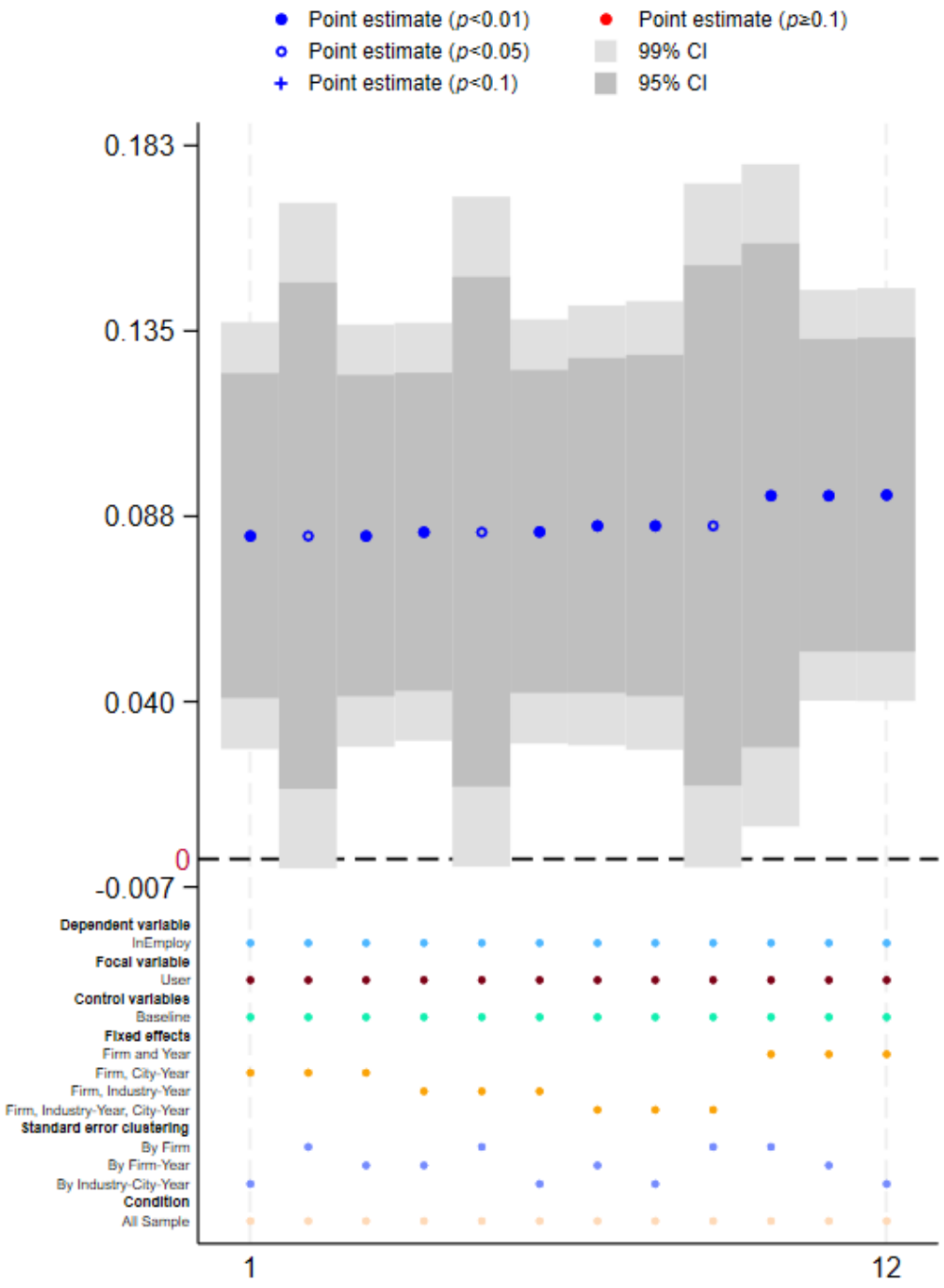


Figure 4: Specification Curve Analysis

Table 1: Summary Statistics

This table presents the summary statistics for the main variables in our study. Panel A contains firm-level control variables, Panel B contains regional-level control variables, Panel C contains moderate variables, and Panel D contains channel variables. *lnEmploy* is the natural logarithm of the number of employees in the firms. *PDPS* is a dummy variable, which is 1 in the year and the following year when a firm establishes PDPS, and 0 otherwise. All continuous variables are winsorized at the 1% and 99% levels to mitigate the effect of outliers. The detailed explanation of other variables can be found in Table A2.

Panel A: Firm-level Variables								
	Mean	S.D.	Min	Q25	Median	Q75	Max	N
lnEmploy	7.702	1.269	7.646	4.511	6.839	8.489	11.218	22213
PDPS	0.090	0.286	0	0	0	0	1	22213
Size	22.209	1.305	22.032	19.803	21.269	22.958	26.237	22213
Lev	0.437	0.207	0.431	0.054	0.272	0.592	0.891	22213
ROA	0.038	0.059	0.036	-0.242	0.014	0.065	0.194	22213
Cash	0.192	0.134	0.155	0.017	0.096	0.250	0.651	22213
Growth	0.437	1.193	0.144	-0.676	-0.020	0.450	8.819	22213
MB	0.613	0.249	0.608	0.114	0.420	0.803	1.161	22213
Age	2.204	0.751	2.303	0.693	1.609	2.833	3.434	22213
SOE	0.438	0.496	0	0	0	1	1	22213
Dual	0.247	0.431	0	0	0	0	1	22213
Panel B: Regional-level Variables								
lnMWage	7.177	0.322	7.237	6.040	7.003	7.415	7.816	6390
lnperGDP	10.791	0.604	10.789	8.599	10.381	11.230	13.056	2422
lnSchool	1.728	1.163	1.609	0	1.099	2.303	4.533	2422
Indstru	0.475	0.100	0.477	0.117	0.413	0.539	0.898	2422
Panel C: Moderate Variables								
Lerner	0.112	0.625	0.107	-86.453	0.048	0.187	0.871	22213
KZ	1.063	2.420	1.294	-11.345	-0.318	2.636	13.662	21794
Salary	0.023	0.073	0.015	-0.002	0.007	0.028	6.355	22181
Panel D: Channel Variables								
RD	0.017	0.035	0	0	0	0.022	0.191	22213
High	0.278	0.225	0.221	0	0.114	0.395	5.961	20041
Tech	0.219	0.183	0.160	0.001	0.099	0.276	0.991	19251
SGR	0.052	0.182	0.050	-9.550	0.018	0.091	12.103	22213
OC	3.050	4.335	1.883	0.103	0.995	3.263	30.539	22199
SalesFlu	1.523	1.811	0.948	0.015	0.410	1.902	10.469	20943

Table 2: PDPS and Corporate Employment

This table reports the results of the staggered DID tests on how PDPS affects corporate employment. The dependent variable is *lnEmploy*. *PDPS* is a dummy variable that takes the value of 1 in the current year and subsequent years if the firm has established user privacy protection in the current year and 0 otherwise. Column (1) includes L.PDPS only; Column (2) controls firm-level control variables; Column (3) further controls regional-level control variables; and Column (4) further controls moderate and channel variables. The detailed explanation of other variables can be found in Table A2. All regressions include the firm-fixed and year-fixed effects. Standard errors are clustered by firm. Robust *t*-statistics clustered at the firm level are displayed in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) lnEmploy	(2) lnEmploy	(3) lnEmploy	(4) lnEmploy	(5) lnEmploy
L.PDPS	0.1432***	0.0945***	0.0931***	0.0797***	0.0800***
	(3.22)	(2.87)	(2.83)	(2.60)	(2.62)
Size		0.6712***	0.6729***	0.6991***	0.6840***
		(28.80)	(28.86)	(32.91)	(31.24)
Lev		0.1456*	0.1462*	0.1943***	0.1944**
		(1.88)	(1.88)	(2.74)	(2.57)
ROA		-0.1538	-0.1541	-0.0603	-0.0626
		(-1.52)	(-1.52)	(-0.51)	(-0.46)
Cash		-0.3590***	-0.3567***	-0.3228***	-0.3091***
		(-6.32)	(-6.27)	(-5.88)	(-4.54)
Growth		-0.0171***	-0.0173***	-0.0105**	-0.0104**
		(-3.57)	(-3.60)	(-2.40)	(-2.47)
MB		-0.0978***	-0.1016***	-0.1033***	-0.1045***
		(-2.58)	(-2.68)	(-3.15)	(-3.08)
Age		-0.0219	-0.0193	-0.0615*	-0.0545
		(-0.74)	(-0.65)	(-1.84)	(-1.63)
SOE		0.1027*	0.1051*	0.0539	0.0458
		(1.88)	(1.92)	(1.07)	(0.89)
Dual		0.0092	0.0089	0.0074	0.0052
		(0.54)	(0.52)	(0.54)	(0.39)
lnMWage			-0.0058	0.0367	0.0337
			(-0.11)	(0.76)	(0.72)
lnperGDP			-0.0454*	-0.0043	0.0061
			(-1.69)	(-0.18)	(0.26)
lnSchool			-0.0737*	-0.0412	-0.0369
			(-1.95)	(-1.21)	(-1.10)
Indstru			-0.3172*	-0.1895	-0.2034
			(-1.69)	(-1.08)	(-1.16)
SGR				-0.0124	0.0073
				(-0.44)	(0.18)
OC				-0.0225***	-0.0230***
				(-6.09)	(-6.07)
MFR				0.2433*	0.2297
				(1.77)	(1.39)
SalesFlu				-0.0035	-0.0022
				(-1.41)	(-0.89)
RD				1.1138***	1.0331***
				(5.35)	(4.87)
Tech				-0.2515**	-0.2669***
				(-2.57)	(-2.69)
High				-0.9985***	-0.9902***
				(-9.02)	(-8.78)
Lerner					0.0363
					(0.58)
KZ					0.0005
					(0.14)
Salary					0.5088
					(1.50)
Constant	7.6888***	-7.1344***	-6.2643***	-7.3318***	-7.1223***
	(1931.14)	(-14.73)	(-9.15)	(-10.99)	(-10.54)
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Observations	22,213	22,213	22,213	18,051	17,725
Number of Firms	2,838	2,838	2,838	2,722	2,713
Adjusted R^2	0.87	0.92	0.92	0.94	0.94

Table 3: Matching Technology

This table reports the staggered DID test results on how user privacy protection affects corporate employment after matching based on PMS and entropy-balance methods. The dependent variable is *lnEmploy*. *PDPS* is a dummy variable that takes the value of 1 in the current year and subsequent years if the firm has established user privacy protection in the current year and 0 otherwise. The PSM method is used in Column (1), and the entropy-balance method is used in Column (2). All regressions include the firm-fixed and year-fixed effects. Standard errors are clustered by firm. Robust *t*-statistics clustered at the firm level are displayed in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) PSM lnEmploy	(2) EB lnEmploy
L.PDPS	0.0596**	0.0780***
	(2.02)	(2.60)
Constant	-6.9852***	-5.6290***
	(-7.38)	(-5.64)
Controls	Yes	Yes
Year FE	Yes	Yes
Firm FE	Yes	Yes
Observations	9,612	22,213
Number of Firms	2,198	2,838
Adjusted R^2	0.94	0.94

Table 4: Excluding Other Policy Influences

This table reports the results of the staggered DID tests on how PDPS affects corporate employment after excluding other policy influences. The dependent variable is *lnEmploy*. *PDPS* is a dummy variable that takes the value of 1 in the current year and subsequent years if the firm has established user privacy protection in the current year and 0 otherwise. *IM* is a dummy variable that takes the value of 1 in the current year and subsequent years when a firm is selected for the IM pilot, and 0 otherwise. *Sub* is a dummy variable that takes the value of 1 in the current year and subsequent years when a firm receives government employment subsidies, and 0 otherwise. *Tax* is a dummy variable that takes the value of 1 in the current year and subsequent years when a firm is eligible for tax incentives, and 0 otherwise. *Gov* is a dummy variable that takes the value of 1 in the current year and subsequent years when a firm benefits from the China's local government debt cleanup campaign, and 0 otherwise. The detailed explanation of other variables can be found in Table A2. All regressions include the firm-fixed and year-fixed effects. Standard errors are clustered by firm. Robust *t*-statistics clustered at the firm level are displayed in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) lnEmploy	(2) lnEmploy	(3) lnEmploy	(4) lnEmploy	(5) lnEmploy
L.PDPS	0.0933***	0.0932***	0.0914***	0.0942***	0.0930***
	(2.84)	(2.83)	(2.77)	(2.88)	(2.83)
IM	0.0465				0.0413
	(1.17)				(1.04)
Sub		0.0225			0.0215
		(1.32)			(1.26)
Tax			0.0336*		0.0301
			(1.71)		(1.53)
Gov				0.0595**	0.0580**
				(2.32)	(2.26)
Constant	-6.2564***	-6.2530***	-6.2169***	-6.1459***	-6.0887***
	(-9.14)	(-9.14)	(-9.06)	(-9.01)	(-8.91)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Observations	22,213	22,213	22,213	22,213	22,213
Number of Firms	2,838	2,838	2,838	2,838	2,838
Adjusted R^2	0.92	0.92	0.92	0.92	0.92

Table 5: Other Robustness Tests

This table reports the results of other robustness tests of how PDPS affects corporate employment. The dependent variable is *lnEmploy*. *PDPS* is a dummy variable that takes the value of 1 in the current year and subsequent years if the firm has established user privacy protection in the current year and 0 otherwise. Column (1) uses a Heckman two-stage regression. Column (1) uses an IV regression. Column (3) uses the Augmented DID regression approach of Li et al. (2016). Column (4) uses the DID regression approach of Callaway and Sant'Anna (2021). Column (5) uses the DID regression approach of Borusyak et al. (2024). Column (6) uses dual machine learning for estimation. Robust t-statistics clustered at the firm level are displayed in parentheses for Columns (1)–(4), and z-statistics are displayed in parentheses for Columns (5)–(6). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Heckman lnEmploy	(2) IV lnEmploy	(3) Augmented DID lnEmploy	(4) csdid lnEmploy	(5) did-imputation lnEmploy	(6) DML-DID lnEmploy
L.PDPS	0.0841**	2.2549**	0.0893*	0.0892*	0.1769***	0.0813***
	(2.51)	(2.01)	(1.73)	(1.87)	(3.37)	(4.54)
IMR	-0.2970**					
	(-2.56)					
Constant	-5.5048***	-6.0313**	-5.8794***			
	(-7.29)	(-7.01)	(-5.64)			
Controls	Yes	Yes	Yes	No	No	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	No	No	No	Yes
Observations	21,248	21,248	11,322	21,086	21,177	22,213
Number of Firms	2,704	2,704	1,068			2,838
Adjusted (Overall) R^2	0.92	0.03	0.90			

Table 6: Moderating Effect Test

This table examines the moderating effect of PDPS on firm employment based on firm market power, financing constraints, and employee wages. The dependent variable is *lnEmploy*. *PDPS* is a dummy variable that takes the value of 1 in the current year and subsequent years if the firm has established user privacy protection in the current year and 0 otherwise. $Lerner_D$ is a grouping dummy variable based on the Lerner index, which takes the value of 1 when the firm's Lerner index is in the top third of the year, and takes the value of 0 otherwise. KZ_D is a grouping dummy variable based on the KZ index, which takes the value of 1 when the firm's KZ index is in the top third of the industry in the year, and takes the value of 0 otherwise. $Salary_D$ is a grouping dummy variable based on the employee's average salary, which takes the value of 1 when the employee's average salary is in the top third of the year, and takes the value of 0 otherwise. The detailed explanation of other variables can be found in Table A2. All regressions include the firm-fixed and year-fixed effects. Standard errors are clustered by firm. Robust *t*-statistics clustered at the firm level are displayed in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) lnEmploy	(2) lnEmploy	(3) lnEmploy	(4) lnEmploy
L.PDPS×$Lerner_D$	0.0732**			0.0499
	(2.12)			(1.47)
L.PDPS×KZ_D		-0.0527*		-0.0391
		(-1.70)		(-1.28)
L.PDPS×$Salary_D$			0.0816**	0.0767**
			(2.16)	(2.13)
L.PDPS	0.0680*	0.1213***	0.0600*	0.0693*
	(1.96)	(3.53)	(1.70)	(1.75)
Lerner	0.0112**			0.0081
	(2.56)			(0.21)
KZ		-0.0014		-0.0036
		(-0.34)		(-0.84)
Salary			-0.0658	-0.0315
			(-1.30)	(-0.39)
Constant	-6.3569***	-5.9066***	-6.2721***	-5.9828***
	(-9.35)	(-8.56)	(-9.15)	(-8.74)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	22,213	21,788	22,180	21,755
Number of Firms	2,838	2,829	2,836	2,827
Adjusted R^2	0.92	0.92	0.92	0.92

Table 7: R&D Investment Channel

This table reports the results of the staggered DID test for the operating performance channel of the PDPS. The dependent variables are the ratio of expensed and capitalized R&D expenditures to total assets ($R\&D$), the share of employees with a bachelor's degree or above ($High$), and the proportion of technical personnel ($Tech$). $PDPS$ is a dummy variable that takes the value of 1 in the current year and subsequent years if the firm has established user privacy protection in the current year and 0 otherwise. Panel A shows the first stage regression, where we fit the $R\&D$, $High$, and $Tech$ by $L.PDPS$ on the year cross-section. While Panel B uses the fitted $\widehat{R\&D}$, $\widehat{High}$, and $\widehat{Tech}$, instead of $L.PDPS$, to explain the dependent variable, $LnEmploy$. All regressions include the firm-fixed and year-fixed effects. Standard errors are clustered by firm. Robust t-statistics clustered at the firm level are displayed in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: PDPS and Firm R&D			
	(1)	(2)	(3)
	R&D	High	Tech
L.PDPS	0.0104***	0.0123*	0.0129**
	(5.06)	(1.71)	(2.25)
Constant	-0.1100***	-0.2061	-0.0379
	(-3.42)	(-1.11)	(-0.30)
Controls	No	No	No
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	22,213	20,038	19,239
Number of Firms	2,838	2,833	2,788
Adjusted R^2	0.67	0.79	0.85
Panel B: Firm R&D and Firm Employment			
	(1)	(2)	(3)
	lnEmploy	lnEmploy	lnEmploy
$\widehat{R\&D}$	2.5083**		
	(2.01)		
$\widehat{High}$		0.5194**	
		(2.49)	
$\widehat{Tech}$			0.5248**
			(2.24)
Constant	-6.3096***	-6.8710***	-6.8448***
	(-9.21)	(-9.71)	(-9.66)
Controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	22,213	20,038	20,038
Number of Firms	2,838	2,833	2,833
Adjusted R^2	0.92	0.93	0.93

Table 8: Operating Performance Channel

This table reports the results of the staggered DID test for the operating performance channel of the PDPS. The dependent variables are sustainable growth rate (SGR), operating cycle (OC), and sales volatility ($SalesFlu$). $PDPS$ is a dummy variable that takes the value of 1 in the current year and subsequent years if the firm has established user privacy protection in the current year and 0 otherwise. Panel A shows the first stage regression, where we fit the SGR, OC, and $SalesFlu$ by $L.PDPS$ on the year cross-section. While Panel B uses the fitted $\widehat{SGR}$, $\widehat{OC}$, and $\widehat{SalesFlu_h}$, instead of $L.PDPS$, to explain the dependent variable, $LnEmploy$. All regressions include the firm-fixed and year-fixed effects. Standard errors are clustered by firm. Robust t-statistics clustered at the firm level are displayed in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: PDPS and Firm Operating Performance			
	(1)	(2)	(3)
	SGR	OC	SalesFlu
L.PDPS	0.0169**	-0.1750*	-0.1792**
	(2.15)	(-1.81)	(-2.10)
Constant	0.2559	-2.3834	-1.4736
	(1.34)	(-0.80)	(-0.66)
Controls	No	No	No
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	22,213	22,199	20,877
Number of Firms	2,838	2,838	2,772
Adjusted R^2	0.34	0.76	0.22
Panel B: Firm Operating Performance and Firm Employment			
	(1)	(2)	(3)
	lnEmploy	lnEmploy	lnEmploy
$\widehat{SGR}$	0.0623		
	(0.09)		
$\widehat{OC}$		-0.1694**	
		(-2.30)	
$\widehat{SalesFlu_h}$			-0.2489***
			(-3.04)
Constant	-6.2699***	-5.7396***	-5.8759***
	(-9.09)	(-7.84)	(-8.38)
Controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	22,213	22,213	22,213
Number of Firms	2,838	2,838	2,838
Adjusted R^2	0.92	0.92	0.92

Appendix for Building Trust, Hiring More: The Role of Privacy Policies in Workforce Growth

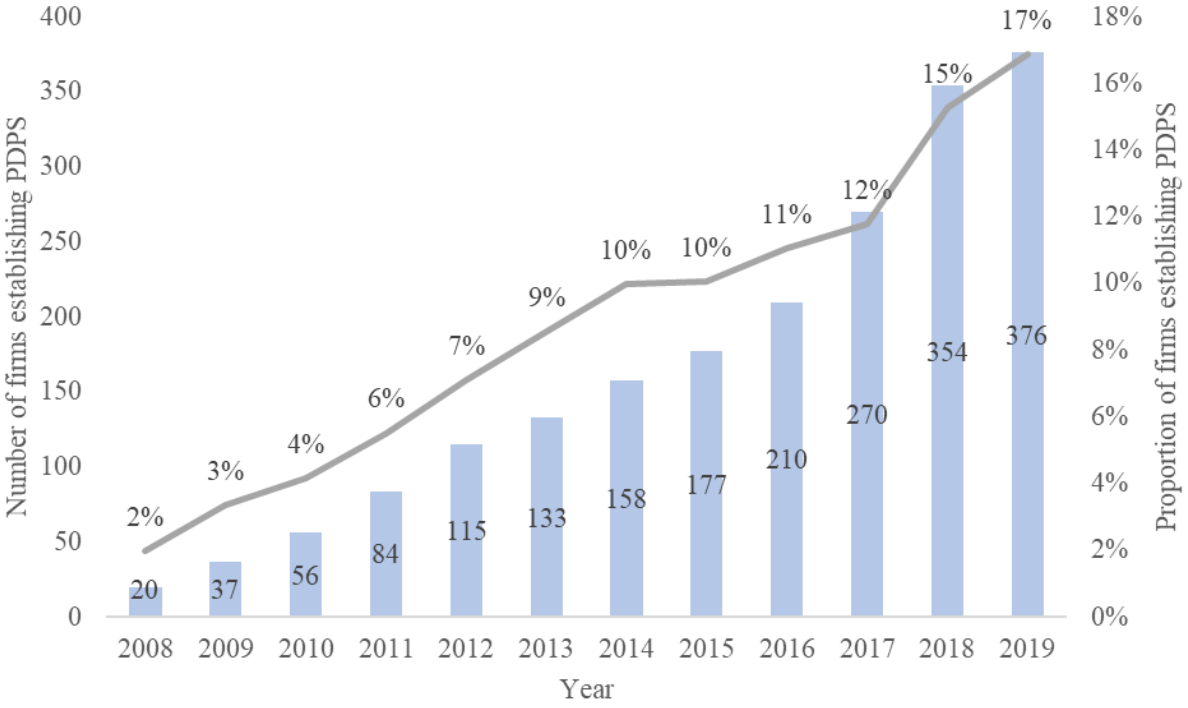
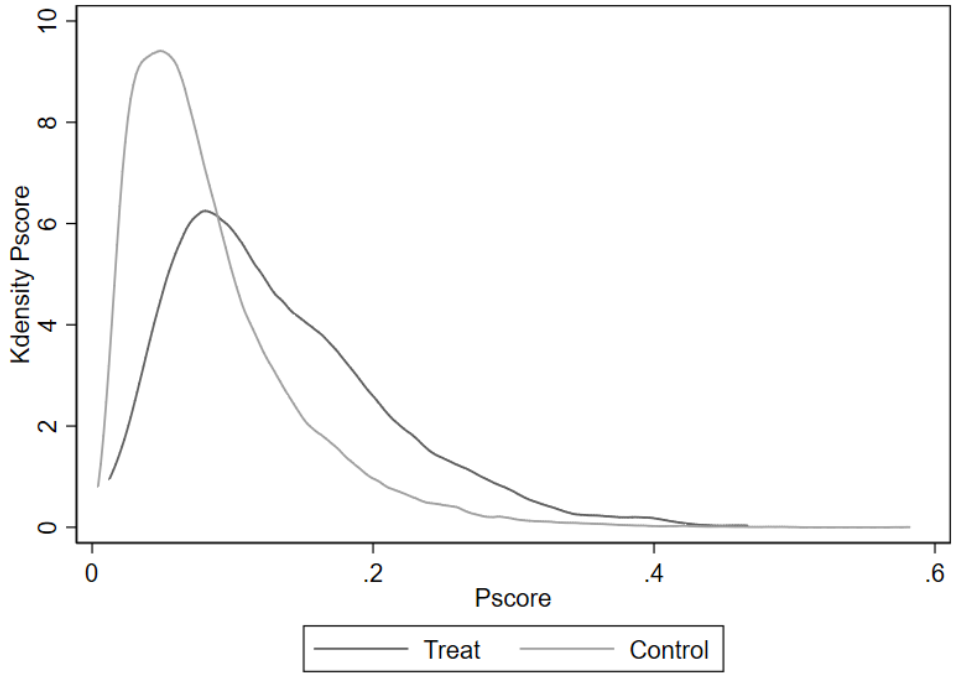
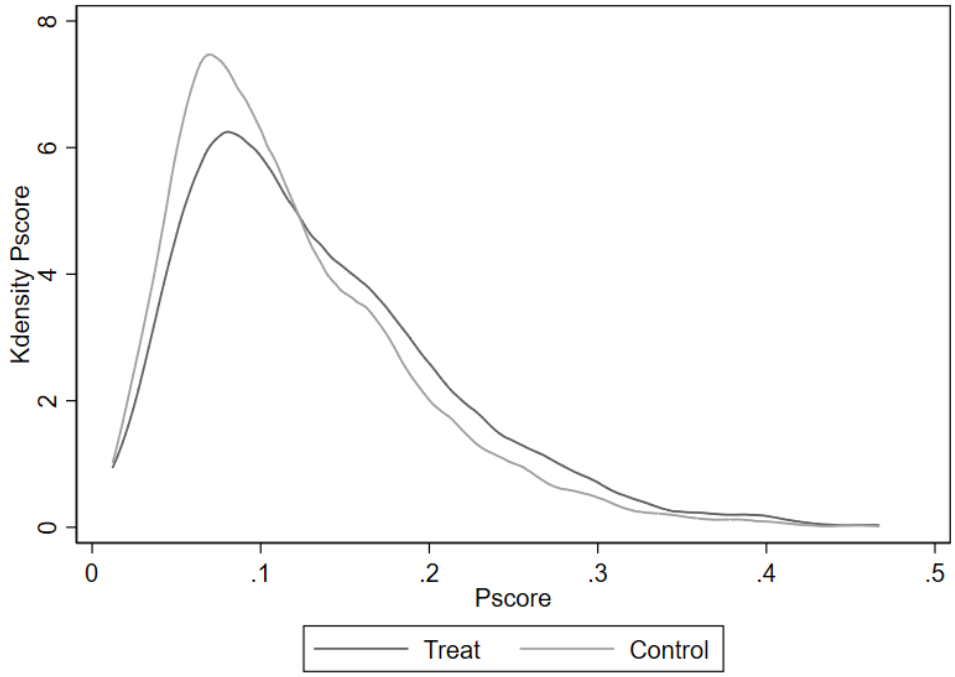


Figure A1: Distribution of Firms Establishing PDPS



(a) Before PSM



(b) After PSM

Figure A2: PSM kernel density curve

Table A1: Keywors of PDPS

This table presents the keywords of PDPS in our study. Based on Chen et al. (2024a), we have systematically constructed a PDPS keyword database that encompasses both directly relevant PDPS terms and extended terminological entries. This taxonomy is organized into five principal dimensions: Information Collection, Information Processing, Related Regulations, Personal Data Protection System, Rights Protection.

Type	**Words**
Information Collection	sensitive personal information, personal biometric data, SDK integration, cookies utilization, user profiling mechanisms, personalized push algorithms
Information Processing	data security capabilities, protection verification protocols, encrypted storage duration, physical/logical partitioning strategies, device tokenization/anonymization techniques, de-identification methodologies
Related Regulations	privacy legitimacy certification, processor compliance statements, confidentiality clauses drafting, internet information sector governance rules, compliance audit frameworks
Personal Data Protection System	data security architecture, information protection supervision mechanisms, dedicated protection departments, data protection officer (DPO) roles and responsibilities
Rights Protection	user's right to erasure ("right to be forgotten"), data portability capabilities, access right implementations, objection to automated processing, personality rights in data ecosystems, property rights in data assets, consent management and withdrawal protocols

Table A2: Variable Definitions

Variable	Definitions
Dependent Variable	
lnEmploy	Natural logarithm of the number of employees.
Independent Variable	
PDPS	A dummy variable that equals 1 if the firm has established PDPS, 0 otherwise.
Control Variables	
Size	Natural log of total assets.
Lev	The ratio of total liabilities to total assets.
ROA	Net income divided by total assets.
Cash	Cash plus cash equivalents divided by total assets.
Growth	Percentage change in sales revenue.
MB	Book-to-Market Ratio.
Age	Natural logarithm of Number of years since establishment plus 1.
SOE	A dummy variable that equals 1 for state-owned firms and 0 otherwise.
Dual	A dummy variable that equals 1 if the chairman and CEO belong to the same person and 0 otherwise.
lnMWage	Natural log of the minimum wage standard in the county where the firm is located.
lnperGDP	Natural log of per capita GDP at prefecture-level city.
lnSchool	Natural log of the number of higher education institutions.
lndstru	Share of secondary industry in prefecture-city-level GDP.
Variables Related to Robustness Tests	
IM	A dummy variable that takes the value of 1 in the current year and subsequent years when a firm is selected for the IM pilot, and 0 otherwise.
Sub	A dummy variable that takes the value of 1 in the current year and subsequent years when a firm receives government employment subsidies, and 0 otherwise.
Tax	A dummy variable that takes the value of 1 in the current year and subsequent years when a firm is eligible for tax incentives, and 0 otherwise.
Gov	A dummy variable that takes the value of 1 in the current year and subsequent years when a firm benefits from China's local government debt cleanup campaign, and 0 otherwise.
Moderator Variables	
Lerner	The Lerner Index measures a company's market power, specifically (operating income - operating costs - sales expenses - administrative expenses) / operating income.
KZ	The KZ index measures the financing constraints of a firm following Kaplan and Zingales (1997).
Salary	Total employee salary/number of employees.
Channel Variables	
R&D	The ratio of expensed and capitalized R&D expenditures to total assets.
High	The percentage of employees with a bachelor's degree or above.
Tech	The percentage of employees engaged in R&D.
SGR	Sustainable growth rate, for specific algorithms, refer to Higgins (1977).
OC	Operating cycle, Days sales of inventory plus Receivables Days/100.
SalesFlu	Fluctuations in abnormal sales performance of firms refer to Ghosh and Olsen (2009).

Table A3: Summary Statistics of Entropy Balancing

This table reports the results of the summary statistics of entropy balancing. Columns (1)-(3) report the mean, variance, and skewness of the treatment group. Columns (4)-(5) report the mean, variance, and skewness of the control group before matching. Columns (6)-(9) report the mean, variance, and skewness of the control group after matching.

	Treat			Control before match			Control after match		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	Variance	Skewness	Mean	Variance	Skewness	Mean	Variance	Skewness
Size	22.480	1.983	0.715	22.180	1.667	0.707	22.480	1.983	0.715
Lev	0.429	0.040	0.253	0.437	0.043	0.122	0.429	0.040	0.253
ROA	0.033	0.004	-1.729	0.038	0.003	-1.254	0.033	0.004	-1.729
Cash	0.211	0.021	1.072	0.190	0.018	1.337	0.211	0.021	1.072
Growth	0.509	1.042	4.424	0.430	1.461	4.947	0.509	1.042	4.424
MB	0.588	0.066	0.221	0.616	0.062	0.057	0.588	0.066	0.221
Age	2.293	0.439	-0.506	2.195	0.576	-0.483	2.293	0.439	-0.506
SOE	0.442	0.247	0.235	0.438	0.246	0.251	0.442	0.247	0.234
Dual	0.297	0.209	0.889	0.242	0.184	1.204	0.297	0.209	0.888
lnMWage	7.432	0.062	-0.899	7.281	0.101	-0.749	7.432	0.062	-0.900
lnperGDP	11.560	0.204	-0.794	11.340	0.303	-0.571	11.560	0.204	-0.795
lnSchool	3.387	1.297	-1.009	3.005	1.481	-0.530	3.387	1.297	-1.009
Indstru	0.361	0.013	-0.100	0.418	0.012	-0.349	0.361	0.013	-0.100