sustainability

MDPI

Article

Compensation–Skills Dual-Dimensional Configuration and Sustainable Upgrading of Export Technological Sophistication: A Re-Explanation of the "Value Paradox" of Skill Mismatch

Zedan Du and Xinxuan Cheng *

School of Economics, Hebei University, Baoding 071000, China; jjxydzd@stumail.hbu.edu.cn
* Correspondence: chengxinxuan@hbu.edu.cn

Abstract

Amid the restructuring of global value chains (GVCs) and the artificial intelligence (AI) revolution, how firms' reallocation of human capital relates to escaping value chain lock-in has become a question of growing theoretical and practical importance for China. Traditional static skill-matching theory treats skill mismatch purely as a distortion in resource allocation. In the course of China's industrial transformation and upgrading, the co-existence of configuration deviation and industrial upgrading is widely observed. Using panel data on Chinese listed firms from 2015 to 2023, this study measures firms' relative configuration deviations along the compensation and skill dimensions against industry-year benchmarks. Different types of configuration deviation are then identified according to the sign combinations of deviations in these two dimensions. The results show that the overall magnitude of configuration deviation is not significantly associated with firms' export technological sophistication (ETS). Instead, the direction and combination of deviations matter: firms with a High-high configuration are associated with significantly higher ETS than those in the reverse configuration. Decomposition by dimension shows that the Compensation-premium is significantly and positively associated with ETS but negatively associated with innovation output, whereas Skill-underutilization is positively associated with innovation yet negatively associated with contemporaneous ETS. Channel tests indicate that the High-high configuration is positively associated with two interconnected prospective channels: innovative-talent accumulation and AI orientation. No robust regional heterogeneity is identified in the association between the High-high and ETS. All coefficients of interest are interpreted as conditional associations estimated within a two-way fixed-effects framework, rather than as identified causal effects. Moving beyond the static job–skill matching paradigm, we interpret these findings through a dynamic capability-building lens and discuss sustainability-oriented policy implications for developing countries seeking to design differentiated talent and digital technology policies.

Keywords: skill mismatch; export technological sophistication; compensation–skills dual-dimensional configuration; innovation output; AI orientation

Check for updates

Academic Editors: António Miguel Rosado da Cruz and Estrela Ferreira Cruz

Received: 11 September 2026
Revised: 11 October 2026
Accepted: 13 October 2026
Published: 16 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The division of labor within global value chains (GVCs) is undergoing systematic restructuring. Against this backdrop, China has progressively built comparative advantages in emerging sectors such as the digital economy and new energy, and its position in GVCs has upgraded in distinct phases. How this upgrading was achieved, and whether it can be sustained, has become a central question for research on trade and industrial development.

Classical economic theory attributes GVC upgrading to the combined outcome of sustained capital accumulation, technological innovation, and improvements in human capital. Within this strand of theoretical reasoning, labor-market outcomes hinge on the alignment between workers' skill endowments and actual job tasks. If employees' skill sets diverge from the operational tasks carried out at the firm level, firm productivity tends to decline, and progress toward high-value-added GVC activities becomes sluggish. From this conventional viewpoint, skill mismatch is often detrimental to the transformation and upgrading of firms.

However, in the practical process of transformation and upgrading among Chinese firms, we observe a phenomenon inconsistent with the classical economic theory of skill matching: some firms whose skill configurations deviate from industry benchmarks have nonetheless achieved sustained growth in export technological sophistication (ETS). According to the authors' calculations based on CSMAR and Wind databases, the mean value of export technological sophistication for Chinese listed firms stood at 10.203, and the mean value of the compensation–skill configuration-deviation index was 0.153 over the period 2015–2023. Data from the World Bank's World Development Indicators (WDI) show that China's share of medium- and high-technology products in manufactured exports rose from 45.49% in 2000 to 61.36% in 2020 [1].

In China, pervasive skill mismatch can co-occur alongside sustained GVC upgrading. This observation motivates our central research question. If the skill-matching theory holds true, why have numerous firms with misaligned skill configurations still achieved improvements in ETS? Behind this paradox, do the wage and skill dimensions show differentiated associations that have been overlooked by static skill-matching theory? Are there mechanisms or contextual factors that classical economic theory overlooks? How can this seemingly paradoxical development path be interpreted?

Sustainable industrial upgrading cannot rely solely on static and perfect job–skill matching; instead, it is associated with a particular pattern of human-capital allocation across the compensation and skill dimensions, as well as the productive empowerment of digital technologies such as artificial intelligence (AI). From a sustainability perspective, it is associated with the trajectory of industrial transformation and with whether industrial evolution follows an inclusive and innovation-driven path or remains confined to low-value and labor-intensive production modes. Against the "value paradox" of skill mismatch and the unique institutional context of China's global value chain restructuring, this study characterizes the association between firm-level two-dimensional compensation–skills configurations and the sophistication of corporate export technology. This paper addresses core sustainability-focused research questions centered on configurational heterogeneity: under what market conditions can specific skill allocation patterns, particularly the High-high, coincide with sustained improvements in firms' ETS without generating inefficient resource misallocation?

2. Literature Review

Within the theoretical framework of global value chains, firms typically pursue export-technology upgrading along four pathways: process upgrading, product upgrading, functional upgrading, and inter-sectoral upgrading [2]. Its main drivers are technological innovation and R&D capacity [3,4], institutional quality and infrastructure [5,6], market scale and industrial ecosystems [7,8], and human capital and organizational agility [9,10]. Mainstream research from the human-capital perspective is rooted in neoclassical human capital theory, emphasizes the importance of job–skill matching, and regards skill mismatch as a form of market friction and inefficient resource allocation [11]. The existing literature documents that skill mismatch reduces firm productivity [12,13], curbs innova-

tion output [14], and impedes technological upgrading [15,16]. The mismatch of skills also hinders the release of the potential benefits from the accumulation of human capital [17]. At the micro level, skill mismatch reduces workers' Compensation-premiums and job satisfaction, thereby weakening their incentives for labor supply and innovative activity [18]. The implied losses are quantitatively substantial. If China and India aligned their factor allocation efficiency with that of the United States, manufacturing total factor productivity (TFP) would rise by 30% to 50% and 40% to 60%, respectively [19]. Within the canonical skill-matching theory, cost efficiency and product heterogeneity constitute the fundamental bases of firms' global competitiveness. By restraining productivity growth and innovative upgrading at the firm level, skill mismatch is therefore expected to erode firms' international competitive advantages.

In recent years, a growing body of scholars has observed real-world phenomena that defy explanations from the static job–skill matching theory. Shao and Bao (2011) [20] find that restructuring among export-oriented firms has no significant effect on employment or wage growth, provided that these firms remain in export markets. Extending this logic to human capital, we argue that stringent requirements from high-end global clients raise the expected value of upgrading, making forward-looking human-capital investment and organizational learning more profitable for firms that approach the skill requirement. Against the backdrop of China's economic transition, some firms build anticipatory talent pools to navigate future technological shifts and market uncertainty [21]. High-skilled labor generates knowledge spillovers that promote intra-firm diffusion of knowledge. Such mechanisms build the human-resource base needed for firms to upgrade operations from production to R&D. Compared with established developed economies, China's GVC advancement trajectory bears "compressed" and "parallel-phase" characteristics. Human-capital allocation needs to move beyond the static optimality framework [16]. Firm-level evidence also accumulates. Existing studies indicate a positive correlation between labor productivity and skill shortages [22], when labor costs climb or specialized skills become scarce, firms tend to scale up their automation investments [23]. The above findings suggest that the economic consequences of skill-configuration deviations are heterogeneous and cannot be uniformly treated as efficiency losses.

Most of the existing literature adopts a static framework and operationalizes skill mismatch as a one-dimensional aggregate magnitude indicator. Few studies decompose the differentiated effects of skill mismatch across distinct dimensions, and the identification of heterogeneous impacts arising from alternative mismatch types remains insufficient. Against this backdrop, this study investigates the "value paradox" of skill mismatch amid the dual context of China's economic transition and global value chain restructuring. This study constructs the compensation–skill configuration-deviation index and decomposes it into four categories. It focuses on whether the High-high configuration is associated with higher ETS above industry benchmarks and a higher share of high-skilled workers relative to industry peers, and examines the differentiated associations of the compensation and skill dimensions. This study further extends the analysis by examining two associative pathways: innovation output and AI orientation. Building on global value chain theory and human capital theory, and specifying the microfoundations of dynamic capability theory for the labor-input margin, this study offers a micro-level theoretical interpretation of Chinese firms' transformation and upgrading from a human-capital perspective.

3. Model and Hypotheses

3.1. Theoretical Positioning

3.1.1. Classical Dynamic Capabilities Theory

Our argument does not displace established theories of the firm; it locates the compensation–skill configuration within them. Dynamic capability theory holds that firms attain sustained performance not through static resource endowments but through the capacity to sense opportunities, seize them, and reconfigure resources accordingly [24,25]. In our setting these three functions map onto distinct, observable margins. Sensing corresponds to a firm's recognition that its existing skill–task alignment is insufficient for higher-value activities. Seizing corresponds to the decision to pay above the industry benchmark so as to improve access to the workforce those activities require, and plausibly to its stability. Reconfiguring corresponds to the accumulation of a high-skilled workforce that permits operations to move from production toward R&D. Our framework is therefore not a substitute for dynamic capability theory but a specification of its microfoundations on the labor-input margin. We treat the compensation–skill configuration as an observable antecedent of capability building rather than as a capability itself. This distinction matters because dynamic capabilities are conventionally defined as high-level routines that are difficult to observe directly [26,27], whereas our configuration indicators provide a measurable firm-year proxy for the input-side conditions under which such routines can be built.

3.1.2. Resource-Based View

The resource-based view explains sustained advantage through resources that are valuable, rare, inimitable, and non-substitutable [28,29]. High-skilled labor satisfies several of these criteria, yet the resource-based view remains a static theory of resource stocks: it explains why heterogeneity persists, not why an allocation that departs from the industry benchmark would be rewarded. Our framework addresses this gap by shifting attention from the level of resources to their configuration. The value-creating unit in our analysis is neither compensation alone nor skills alone, but their joint configuration: compensation above the industry benchmark is expected to improve screening and incentives while a skill surplus raises the return to that compensation. Because this complementarity is firm-specific and difficult to observe from the outside, it is also difficult to imitate, which is precisely the isolating mechanism emphasized by the resource-based view. In this sense, the compensation–skill configuration can be understood as a resource-based advantage operating through the configuration of human capital rather than through its level.

3.1.3. Organizational Learning Theory

Organizational learning theory and the concept of absorptive capacity imply that a firm can exploit external knowledge only when it already possesses relevant internal knowledge [30,31]. Skill-underutilization, commonly treated as waste in static matching frameworks, is precisely the condition that generates this absorptive capacity. It supplies the cognitive slack required to evaluate, absorb, and recombine new technologies, and it is a precondition for the AI-orientation channel we document. The cost–benefit asymmetry we observe is consistent with the exploration–exploitation distinction [32]: the returns to exploratory investment are deferred, whereas its costs are immediate.

Taken together, this study makes three deliberately bounded contributions. Conceptually, it reconceptualizes skill mismatch from a one-dimensional scalar magnitude into a two-dimensional configuration defined by the signs of deviations along the compensation and skill dimensions, thereby distinguishing allocation patterns that aggregate indicators conflate. Mechanistically, it identifies a separability condition between static and dynamic optimality: an allocation that is suboptimal relative to the industry-year benchmark in the

contemporaneous cross-section may nonetheless be associated with superior upgrading outcomes, because the two objectives are evaluated over different horizons. Contextually, it specifies the boundary conditions of this interpretation. The argument applies to firms in transition economies undergoing compressed and overlapping phases of upgrading, in which capability endowments are still being formed. It does not imply that departures from the benchmark are beneficial without limit, nor does it claim that skill mismatch improves efficiency once a configuration exceeds reasonable industry thresholds.

3.2. Definition of Main Variables

Traditional research conceptualizes skill mismatch as static resource misallocation and market failure. Firms select factor inputs to minimize costs or maximize profits, conditional on given production functions and market prices. The primary function of the labor market is to achieve job–skill matching, and skill mismatch is argued to generate direct and persistent productivity losses. Using industry-year averages as the reference benchmark, this study defines two relative positional deviations along the compensation and skill dimensions:

$$r_{it}^{W} = \ln\left(\frac{W_{it}}{\overline{W}_{jt}}\right) \tag{1}$$

$$r_{it}^{S} = \ln\left(\frac{S_{it}}{\overline{S}_{jt}}\right) \tag{2}$$

where r_{it}^{W} denotes the deviation of per capita compensation for firm i from the industry average, W_{it} stands for per capita compensation for firm i, and $\overline{W}_{jt}$ is the industry-year average of per capita compensation in year t. $r_{it}^{W} > 0$ indicates that a firm pays wages above the industry average, while $r_{it}^{W} < 0$ means compensation falls below the industry benchmark.

r_{it}^{S} stands for the deviation of skill level for firm i from the industry average, S_{it} corresponds to the proportion of high-skilled workers among total employees for firm i, and $\overline{S}_{jt}$ is the industry-year average proportion of high-skilled personnel in year t. When $r_{it}^{S} > 0$, high-skilled personnel configuration lies above the industry benchmark; when $r_{it}^{S} < 0$, the share of high-skilled personnel is below the industry benchmark.

Conceptually, the two dimensions are asymmetric. r_{it}^{S} is a quantity-oriented metric that directly characterizes the allocation intensity of high-skilled human capital using the share of R&D personnel. r_{it}^{W} is a price-oriented metric that captures firms' incentive and screening intensity relative to industry-year benchmarks, rather than serving as a proxy for skill endowments. Three conditions must be satisfied for r_{it}^{W} to act as a valid configuration signal: relatively competitive intra-industry labor markets, comparable industry-year benchmarks, and the inability of firms to sustain remuneration payments below marginal output in the long run. It should be noted that compensation incorporates three intertwined components—talent screening, effort incentives, and rent-sharing—which cannot be fully disentangled empirically. Accordingly, this paper only interprets the empirical outcomes as conditional associations and does not claim that compensation deviation is equivalent to skill deviation.

Drawing on the analytical framework of factor misallocation proposed by Hsieh and Klenow (2009) [19], this study constructs firm-level proxy indicators for compensation–skill two-dimensional configuration-deviation index along compensation and skill dimensions, tailored to the characteristics of Chinese data:

$$S_{mis,it} = \left| r_{it}^{W} \times r_{it}^{S} \right| \tag{3}$$

Traditional studies typically define skill mismatch from the perspective of individual job–skill matching. This research incorporates the dimension of relative compen-

sation deviation and characterizes firm-level configurational deviations from the dual compensation–skill perspective; it is conceptually distinct from the vertical/horizontal educational mismatch in the standard literature. Accordingly, the constructed proxy indicator for skill mismatch partially captures price distortion information embedded in human-capital misallocation, which differentiates it from narrow definitions of skill mismatch that focus solely on the supply and demand of capabilities. Note that constructing the skill-mismatch proxy via absolute-value calculation only measures the magnitude of configurational deviation while discarding dimensional and sign information, and thus cannot reflect all types of configurational deviation. This study further identifies four categories of configurational deviation based on sign combinations of the two dimensions, as shown in Table 1.

Table 1. Four types of configuration deviations.

Configuration Type	Notation	Sign Condition	Economic Interpretation
High-high	$H_{H,it}$	$r_{it}^{W} > 0, r_{it}^{S} > 0$	Both compensation and the share of high-skilled personnel exceed the industry benchmark.
Compensation-premium	$W_{H,it}$	$r_{it}^{W} > 0, r_{it}^{S} < 0$	A Compensation-premium is present, yet the share of high-skilled personnel is insufficient.
Skill-underutilization	$R_{H,it}$	$r_{it}^{W} < 0, r_{it}^{S} > 0$	High-skilled personnel are overstaffed, but no Compensation-premium exists.
Low-low	$L_{L,\mathrm{it}}$	$r_{it}^{W} < 0, r_{it}^{S} < 0$	Both compensation and the share of high-skilled personnel fall below the industry benchmark.

3.3. Mathematical Model and Research Hypotheses

This study departs from that view by treating skill mismatch as a dynamic incentive signal and a binding constraint. Consider a representative firm whose average workforce skill level is denoted by S. The firm is initially positioned in the low-end segment of the global value chain and adopts low-technology production (T_l), whose skill requirement is R_l. We assume $S > R_l$, so that employee skills are underutilized: low-technology jobs fail to leverage workers' skill advantages. This configuration reflects a skill surplus relative to the firm's current low-technology activity, it does not imply that skills meet the requirement of the high-technology activity, and it should not be read as over-skilling against that requirement. At this stage, output Y_l is independent of the skill level S and is solely determined by total factor productivity A_l under low-technology conditions. Skill mismatch thus produces inefficient resource allocation: even high-skilled workers cannot raise output in low-technology production activities. The output of low-technology production is therefore given by:

$$Y_l = A_l \tag{4}$$

The representative firm may incur a fixed cost F for upgrading to high-tech production (T_h), associated with a skill requirement R_h where $R_h > R_l$. Post-upgrading, the firm can reduce the degree of skill mismatch by undertaking innovation investment I, which substitutes for part of the skill needs of labor through technological progress. Effective skill requirement after innovative investment is given by:

$$R_{eff} = \frac{R_h}{1 + \nu I} \tag{5}$$

where ν denotes the innovation efficiency parameter with $\nu > 0$, which measures the firm's capacity to alleviate skill mismatch via innovation investment. Within China's institutional context, ν tends to be relatively high owing to policy support, strong absorptive capacity for technological learning, and intense market competition pressure. The innovation cost

function is denoted by $C(I) = 1/2cI^2$, where c represents the innovation cost coefficient and $c > 0$.

Post-upgrading output is determined by the extent of skill matching:

$$Y_h = A_h \times \min\left(1, \frac{S}{R_{eff}}\right) = A_h \times \min\left(1, \frac{S(1+\nu I)}{R_h}\right) \tag{6}$$

where A_h stands for total factor productivity under high-technology production, and $A_h > A_l$. Let $\gamma = S/R_h$ denote the ratio of actual labor skills to the skill requirement for high-technology production. Then:

$$Y_h = A_h \times \min(1, \gamma(1+\nu I)) \tag{7}$$

The enterprise optimizes innovation expenditure I to maximize net output following technological and skill upgrading:

$$\pi_h(I) = Y_h - C(I) \tag{8}$$

The firm faces two decision scenarios:

When $\gamma \geq 1$, we have $\gamma(1+\nu I) \geq 1$, which yields $Y_h = A_h$. The firm has no incentive to conduct innovation investment, hence the optimal innovation investment $I^* = 0$, and profit $\pi_h = A_h$.

When $\gamma < 1$, the high-technology output is expressed as $Y_h = A_h \times \gamma(1+\nu I)$. We solve the optimization problem as follows:

$$\max A_h \times \gamma(1+\nu I) - \frac{1}{2}cI^2 \tag{9}$$

The following discussion is confined to the interior-point region: $\gamma + \frac{A_h\gamma^2\nu^2}{c} \leq 1$.

First-order condition:

$$A_h\ \gamma\nu - cI = 0 \tag{10}$$

That is:

$$I^* = \frac{A_h\ \gamma\nu}{c} \tag{11}$$

Substitution yields:

$$Y_h = A_h\gamma\left(1 + \nu\frac{A_h\ \gamma\nu}{c}\right) = A_h\gamma + \frac{A_h^2\gamma^2\nu^2}{c} \tag{12}$$

$$C_{I^*} = \frac{1}{2}c\left(\frac{A_h\ \gamma\nu}{c}\right)^2 = \frac{1}{2}\frac{A_h^2\gamma^2\nu^2}{c} \tag{13}$$

Calculation results:

$$\pi_h = Y_h - C_{I^*} = A_h\gamma + \frac{1}{2}\frac{A_h^2\gamma^2\nu^2}{c} \tag{14}$$

The condition for firm technological upgrading is as follows:

$$\pi_h - F > Y_l \tag{15}$$

Which implies:

$$A_h\gamma + \frac{1}{2}\frac{A_h^2\gamma^2\nu^2}{c} - F > Y_l \tag{16}$$

Equation (11) shows that optimal innovation investment rises with the degree of skill matching and innovation efficiency and decreases with the increase in innovation costs. Its economic interpretation is as follows: higher skill levels and stronger compensation incentives increase the marginal returns to innovation, which strengthens firms' incentives to raise innovation investment.

By introducing two-dimensional deviations into the model, the skill-dimension deviation affects the degree of skill matching.

$$\gamma = \gamma(r^s), \frac{\partial \gamma}{\partial r^s} > 0 \tag{17}$$

Deviation along the compensation dimension affects innovation efficiency and costs through talent screening and incentives:

$$\nu = \nu(r^w), \frac{\partial v}{\partial r^w} > 0 \tag{18}$$

$$c = c(r^w), \frac{\partial c}{\partial r^w} < 0 \tag{19}$$

It follows from Equation (14) that upgraded net output is higher. Equations (17)–(19) present reduced-form assumptions: a higher skill dimension raises γ; a high compensation dimension may be associated with greater access to relevant talent, enhancing absorptive capacity and skill coordination efficiency, thereby increasing ν and reducing c. These assumptions combine screening, incentive, and rent-sharing components, and the three mechanisms cannot be fully disentangled. Accordingly, this study frames them only as conditional associations rather than causal interpretations.

For firms with a High-high, $r^w > 0$ and $r^s > 0$; hence, γ is relatively high and $\frac{\nu}{c}$ is relatively high. It follows from Equation (11) that their optimal innovation investment is greater:

$$I^* = \frac{A_h \, \gamma(r^s)\nu(r^w)}{c(r^w)} \tag{20}$$

Substituting Equations (17)–(19) into Equation (14) yields the upgraded net output for firms with a High-high.

$$\pi_h(r^w, r^s) = \mathrm{A_h}\gamma(r^s) + \frac{1}{2}\frac{\mathrm{A_h^2}\gamma(r^s)^2\nu(r^w)^2}{c(r^w)} \tag{21}$$

For firms with a High-high, $r^w > 0$ and $r^s > 0$, so $\gamma(r^s)$ and $\frac{\nu(r^w)}{c(r^w)}$ remain relatively high. Accordingly, $\pi_h(r^w, r^s)$ exceeds that of firms with Compensation-premium and Skill-underutilization. Firms with a High-high are more likely to satisfy the upgrading condition and thereby upgrade from low-technology T_l to high-technology T_h.

Let the ETS of a firm be E_{it}. The ETS before upgrading is E_l; after upgrading, it equals E_h, with $E_h > E_l$. Accordingly:

$$E_{it} = E_l + (E_h - E_l) \times 1(\pi_\mathrm{h} - \mathrm{F} > \mathrm{Y_1}) \tag{22}$$

From Equations (21) and (22), the average ETS for firms with a High-high exceeds that of firms with Compensation-premium and Skill-underutilization configurations. Accordingly, this study proposes Research Hypothesis H1:

H1. *Firms with a High-high configuration exhibit higher export technological sophistication than firms with the reverse configuration, which serves as the reference group.*

It follows from Equation (20) that firms with a High-high achieve higher optimal innovation investment. Innovation output is expected to raise the technological content of products and is therefore treated as an associative channel linking the High-high to ETS. On this basis, this study proposes Research Hypothesis H2:

H2. *The High-high skill configuration is positively associated with greater innovation output for firms, and innovation output acts as an associative channel linking the High-high to export technological sophistication.*

AI orientation can be regarded as a special form of innovation with higher efficiency. Let the level of AI orientation be:

$$AI^* = \frac{A_h \gamma(r^s) \nu_{AI}(r^w)}{c(r^w)} \tag{23}$$

where ν_{AI} denotes the AI orientation rate for a firm, with $\nu_{AI} > v$. The compensation and skill dimensions may be associated with ν_{AI} and γ. Under the High-high, $r_{it}^W > 0$ and $r_{it}^S > 0$; accordingly, AI^* is greater. AI orientation is expected to be associated with higher production efficiency and product technological sophistication through automation, intelligent optimization and human–machine collaboration, and thereby with higher ETS. Hence, AI orientation is expected to be associated with higher production efficiency and is therefore treated as another associative channel. On this basis, this study proposes Research Hypothesis H3:

H3. *The High-high skill configuration is positively associated with greater AI orientation—reflected in AI-related disclosure, patenting and AI-subsidiary establishment—and this AI orientation acts as an associative channel linking the High-high to export technological sophistication.*

4. Methodology

4.1. Indicator Selection

4.1.1. Dependent Variables: ETS

This study follows the existing literature [33–35] and adopts industry-level ETS adjusted by firms' relative total factor productivity as the dependent variable. It should be noted that this indicator characterizes the relative position of firms' ETS within the export technology spectrum and is not fully equivalent to functional upgrading or value-added acquisition capacity emphasized in the global value-chain upgrading literature. This distinction is retained throughout the conceptual interpretation and result discussion of this study. The specific calculation procedure is presented below:

Following the measurement framework of Hausmann et al. [33], we estimate the ETS index for product m as:

$$EXP_{m,t} = \sum_k \frac{x_{k,m,t}/X_{k,t}}{\sum_k x_{k,m,t}/X_{k,t}} \times GDP_{k,t} \tag{24}$$

where $EXP_{m,t}$ denotes the ETS of product m in year t. $x_{k,m,t}$ is the export value of product m from country k in year t. $X_{k,t}$ stands for the overall export revenue for economy k, and $GDP_{k,t}$ the per capita GDP of economy k, all measured in year t.

This study uses the ETS of listed firms as the proxy for GVC position. Following the measurement framework of Hausmann et al. (2007) [33], the ETS index for product m is estimated as:

$$EXP_{k,j,t} = \sum_m \frac{x_{k,j,m,t}}{X_{k,j,t}} EXP_{k,m,t} \tag{25}$$

where $EXP_{k,j,t}$ represents the technological sophistication of overseas shipments for sector j for economy k in year t; $x_{k,j,m,t}$ denotes the export value of product m for sector j from economy k in year t. $X_{k,j,t}$ denotes the total export value for sector j for economy k in year t.

Following the empirical framework of Gao and Yuan (2020) [36], we adjust $EXP_{k,j,t}$ with firm-level total factor productivity and obtain the firm-level ETS as follows:

$$EXP_{k,i,t} = \frac{TFP_{k,i,t}}{TFP_{k,j,t}} EXP_{k,j,t} \tag{26}$$

where $EXP_{k,i,t}$ denotes the ETS of firm i for economy k in year t; total factor productivity is calculated according to the Levinsohn–Petrin method. $TFP_{k,j,t}$ stands for the total factor productivity level of individual firm i affiliated with economy k during period t; $TFP_{k,j,t}$ denotes the total factor productivity of industry j for country k in year t.

The dependent variable is an industry-level ETS indicator adjusted using firm-level relative productivity:

$$E_{si,it} = ln\left(EXPY_{jt}\right) \times \frac{TFP_{it}}{TFP_{jt}} \tag{27}$$

The dependent variable essentially captures the position of a firm's productivity relative to its industry within the industry's export technological sophistication spectrum. Therefore, two scope conditions need to be clarified. First, this indicator measures the relative position of a firm's export product portfolio and is not equivalent to functional upgrading or value-added capture capacity. Second, all conclusions of this study hold only within the definitional scope of this indicator; extrapolation beyond this scope should therefore be treated with caution.

4.1.2. Independent Variable

This study constructs the firm-level compensation–skill configuration-deviation index based on the compensation and skill dimensions; see Section 3 for the detailed construction procedure.

4.1.3. Mediating Variables

Guided by the theoretical model, this study identifies two potential channels, corporate innovation (P_{AT}) and artificial intelligence development (A_I), through which configuration-deviation may affect firms' ETS.

4.1.4. Control Variables

Following the firm-level literature on ETS, we control for a set of firm characteristics: firm size, leverage, return on total assets, ownership concentration, the share of fixed assets, and firm age. Year and firm fixed effects are included to absorb common macroeconomic shocks and time-invariant firm heterogeneity.

The definitions and data sources are reported in Table 2, and Supplementary Material Table S1 reports variance inflation factors while Supplementary Material Table S2 reports the pairwise correlation matrix of the main variables.

4.2. Sample Construction and Data Preparation

This study employs panel data of Chinese A-share listed firms. Firm financial indicators and corporate governance variables are obtained from the China Stock Market & Accounting Research (CSMAR) and Wind databases. Data on firm employee compensation and R&D personnel are extracted from the annual reports of listed companies; after sorting, we derive firm-level average employee compensation and the share of R&D personnel. The industry benchmarks are the sample means of firm-level per capita compensation and the

share of R&D personnel, calculated by industry and year within the sample. Industry-level export technological sophistication is calculated using country-product trade flows from the BACI-CEPII database. Specifically, product-level export technological sophistication is first computed following the method proposed by Hausmann et al. [33], and then weighted and aggregated to the two-digit sector level of the national economic classification.

Table 2. Definitions and assignment of variables.

Variable Category	Variable Name	Variable Symbol	Calculation Method	Data Source
Dependent Variable	Export technological sophistication	$E_{si,it}$	Calculated by the authors	BACI database (CEPII),CSMAR database, Wind database
Independent Variable	Proxy indicator of skill mismatch	$S_{mis,it}$	Calculated by the authors	Annual reports of listed firms
	High-high	$H_{H,it}$	Calculated by the authors	Annual reports of listed firms
	Low-low	$L_{L,it}$	Calculated by the authors	Annual reports of listed firms
	Compensation-premium	$W_{H,it}$	Calculated by the authors	Annual reports of listed firms
	Skill-underutilization	$R_{H,it}$	Calculated by the authors	Annual reports of listed firms
Channel variables	Firm innovation	lnP_{AT}	Ln (1 + number of patent applications)	CSMAR Database
	Firm AI orientation	lnA_I	Ln (1 + annual-report AI keyword frequency)	Annual reports of listed firms
Control Variables	Firm Size	S_{ize}	Natural logarithm of total assets	CSMAR Database
	Leverage	L_{ev}	Total liabilities/total assets	CSMAR Database
	Return on assets	R_{OA}	Net profit/total assets	CSMAR Database
	Ownership concentration	T_{op1}	Shareholding ratio of the largest shareholder	CSMAR Database
	Fixed asset ratio	F_{IXED}	Net fixed assets/total assets	CSMAR Database
	Firm age	A_{ge}	Ln (1 + firm age)	CSMAR Database
	Year fixed effects	*Year FE*	Controlling for year-specific effects	-
	Firm fixed effects	*Firm FE*	Controlling for firm-specific effects	-

It should be noted that the BACI-CEPII data are only used to construct the benchmark of industry-level export technological sophistication and are not directly matched with the firm sample. All firm-level variables are obtained within the CSMAR-Wind data framework. The sample processing steps are as follows: (1) exclude ST, *ST, and delisted firms; (2) exclude financial firms; (3) drop observations with missing key variables or implausible values; (4) take the logarithm of selected variables and apply a 1% winsorization to all continuous variables.

The final sample comprises 3693 firms and 21,471 observations. The sample exhibits a distinctly unbalanced distribution across time: 21,317 observations fall within 2015–2023, accounting for 99.28% of the total; due to limited disclosure of listed firms' employee educational attainment and compensation structure information during 2007–2012, only 154 observations (92 firms) are available for this period, representing 0.72% of the total. No valid observations exist for 2013–2014. Accordingly, the empirical tests of this study mainly rely on the 2015–2023 subsample. The baseline analysis reports results for 2015–2023, while the full-sample results are presented in the robustness checks.

Furthermore, the dependent variable is constructed based on industry-level export technological sophistication and adjusted by firm relative productivity relative to its industry (see Equation (26)). Its value does not depend on whether a firm engages in exports; non-exporting firms also have well-defined values for this indicator. Therefore, the sample

covers all A-share listed firms and is not screened by export status, which helps avoid the self-selection concern of exporting firms.

4.3. Model Specification

This study constructs the baseline model to examine the association between the proxy for skill mismatch and export technological sophistication, conditional on controlling for other influencing factors:

$$E_{si,it} = \beta_0 + \beta_1 S_{mis,it} + \sum_j \gamma_j Controls_{jit} + v_t + \mu_i + \varepsilon_{it} \quad (28)$$

where i indexes firms and t indexes years. The dependent variable $E_{si,it}$ measures the ETS of firm i in year t. $S_{mis,it}$ is the explanatory variable, representing the compensation–skill configuration-deviation index of firm i in year t. The vector $Controls_{jit}$ contains the firm-level controls described above: firm size, leverage, return on total assets, ownership concentration, the share of fixed assets, and firm age. Year fixed effects are captured by v_t, firm fixed effects by μ_i, and ε_{it} is the error term. Given the presence of serial correlation within firm-level panel data, standard errors for all regressions are clustered at the firm level. The variables in subsequent models retain the same definitions as those in Equation (28).

$S_{mis,it}$ is the proxy variable for skill mismatch, which reflects the overall deviation in the compensation–skill dimensions relative to industry averages and captures the net value after the offsetting of effects across the two dimensions. The type decomposition in Section 3.2 provides a structural decomposition of this net value. This study defines Compensation-premium and Skill-underutilization with opposite directions of two-dimensional deviation as reverse-configuration mismatch ($M_{ismatch_Negative}$) and constructs a dummy variable (Table 3):

Table 3. Sample composition of the dummy variables.

Type	No. of Observations	Share (%)
High-high ($H_{H,it}$)	4614	21.6%
Low-low ($L_{L,it}$)	9160	43.0%
Opposite-sign (crossed) configuration ($M_{ismatch_Negative}$)	7543	35.4%
Total	21,317	100%

Note: Configurations are classified relative to industry-year mean benchmarks. Opposite-sign (crossed) configuration (firms whose wage and skill deviations point in opposite directions) serves as the reference group in the regressions.

$H_{H,it}$ = 1 if ($r_{it}^W > 0$, $r_{it}^S > 0$), and 0 otherwise.
$L_{L,it}$ = 1 if ($r_{it}^W < 0$, $r_{it}^S < 0$), and 0 otherwise.
$M_{ismatch_Negative}$ = 1 if ($r_{it}^W > 0$, $r_{it}^S < 0$) or ($r_{it}^W < 0$, $r_{it}^S > 0$), and 0 otherwise.

Taking $M_{ismatch_Negative}$ as the reference category, we replace the continuous mismatch index with a set of dummy variables and estimate the following specification:

$$E_{si,it} = \alpha + \beta 1 H_{igh_{High},it} + \beta 2 L_{ow_{Low},it} + \sum_j \gamma_j Controls_{jit} + \mu_i + \varphi_t + \epsilon_{it} \quad (29)$$

To examine the channels through which the High-high relates to export technological sophistication, this study constructs econometric models following the two-step approach proposed by Jiang [37], from two pathways: firm innovation and AI orientation:

$$lnP_{AT} = \beta_0 + \beta_1 H_{H,it} + \sum_j \gamma_j Controls_{jit} + v_t + \mu_i + \varepsilon_{it} \quad (30)$$

$$lnA_I = \beta_0 + \beta_1 H_{H,it} + \sum_j \gamma_j Controls_{jit} + v_t + \mu_i + \varepsilon_{it} \quad (31)$$

5. Empirical Analysis

5.1. The Association Between Compensation–Skills Dual Configuration and Export Technological Sophistication

5.1.1. Descriptive Statistical Analysis

Descriptive statistical analysis enables a preliminary assessment of data distribution, data quality and latent data anomalies. Table 4 presents the descriptive statistics of the main variables. For the dependent variable, the mean value is 10.203, with a standard deviation of 1.068, a minimum of 7.967 and a maximum of 13.187. This indicates substantial heterogeneity in export technological sophistication across firms, which provides sufficient variation for the subsequent empirical tests. For the explanatory variable, the mean equals 0.153, the standard deviation is 0.219, the minimum is 0.000 and the maximum reaches 1.215, suggesting that sample firms generally exhibit a certain degree of configuration deviation, with large cross-firm differences in the magnitude of such deviation. Turning to the mediating variables, Firm innovation has a mean of 3.540 and a standard deviation of 1.588, with a minimum value of 0.000. This implies that some firms have no patent output, and the overall innovation level shows obvious disequilibrium. AI orientation has a mean of 1.070, a standard deviation of 1.253, a minimum of 0.000 and a maximum of 4.509, demonstrating that corporate AI orientation is still at an early developmental stage, and there exists notable divergence in the progress of intelligent transformation across firms. Firm-level control variables include firm size, leverage, profitability, governance, asset structure, firm age and other attributes, which enable us to effectively account for firm heterogeneity. Overall, the distributional characteristics of each variable are broadly consistent with the general profile of Chinese listed firms, and the selected sample exhibits good representativeness.

Table 4. Descriptive statistics of variables.

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
$E_{si,it}$	21,317	10.203	1.068	7.967	13.187
$S_{mis,it}$	21,317	0.153	0.219	0.000	1.215
$ln\ P_{AT}$	21,317	3.540	1.588	0.000	7.206
$ln\ A_I$	21,317	1.070	1.253	0.000	4.509
S_{ize}	21,317	22.080	1.161	20.002	25.662
L_{ev}	21,317	0.378	0.186	0.060	0.850
R_{OA}	21,317	0.046	0.066	−0.198	0.215
T_{op1}	21,317	0.326	0.139	0.089	0.706
F_{IXED}	21,317	0.210	0.126	0.014	0.587
A_{ge}	21,317	2.977	0.291	2.197	3.555

Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

5.1.2. Baseline Regression Analysis

Table 5 reports the baseline regression results. Column (1) only incorporates the continuous proxy for skill mismatch $S_{mis,it}$ as the core explanatory variable. The regression results show that the estimated coefficient of $S_{mis,it}$ fails to pass the significance test. This finding differs from the mainstream view in the existing literature that skill mismatch is inevitably associated with efficiency losses. The reason is as follows: the skill-mismatch proxy $S_{mis,it}$ is constructed as an absolute-value measure, which only captures the overall deviation in the compensation–skill two-dimensional configuration and ignores the heterogeneous effects of different human capital configuration types on export technological sophistication.

Table 5. Baseline regression results for the full sample.

Variable	(1) $E_{si,it}$	(2) $E_{si,it}$
$S_{mis,it}$	−0.043 (−1.230)	
S_{ize}	0.603 *** (33.128)	0.603 *** (33.127)
L_{ev}	0.356 *** (5.967)	0.357 *** (5.933)
R_{OA}	2.144 *** (24.800)	2.147 *** (24.811)
T_{op1}	0.013 (0.130)	0.012 (0.123)
F_{IXED}	−0.774 *** (−9.324)	−0.761 *** (−9.079)
A_{ge}	0.337 *** (2.768)	0.337 *** (2.761)
$H_{H,it}$		0.037 *** (3.184)
$L_{L,it}$		−0.017 * (−1.899)
constant	−4.180 *** (−8.712)	−4.186 *** (−8.719)
year	yes	yes
Firm	yes	yes
N	20,978	20,978
Within R^2	0.432	0.433

Note: *** and * denote significance at the 1% and 10% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Singleton firms are dropped, yielding 20,978 firm-years. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

Column (2) takes opposite-sign (crossed) configuration $M_{ismatch_Negative}$ as the reference group and introduces the dummy variables for $H_{H,it}$ and $L_{L,it}$ for estimation. The results show that the baseline two-way fixed-effects estimate for the High-high configuration is 0.037 (significant at the 1% level), which we interpret as a conditional association rather than an identified causal effect. This indicates that firms with High-high exhibit notably higher export technological sophistication relative to reference-group firms, supporting Hypothesis H1. The coefficient for the Low-low configuration is −0.017 and significant at the 10% level, suggesting that firms with Low-low configuration have lower export technological sophistication compared with reference-group firms. A plausible explanation is that Low-low suffers shortages in both compensation incentives and skill reserves. Such firms do not appear to realize the synergistic benefits associated with the High-high; instead, they face greater disadvantages than opposite-sign (crossed) configuration with one-dimensional advantages.

Regarding the control variables, the coefficients of firm size (S_{ize}), return on total assets (R_{OA}), asset-liability ratio (L_{ev}), and firm age (A_{ge}) are generally positive and significant. This suggests that they are associated with higher export technological sophistication. The coefficient of the fixed asset ratio (F_{IXED}) is significantly negative at the 1% level, indicating that firms overly reliant on traditional heavy assets face relatively constrained growth in technological sophistication. The within R^2 values for Columns (1) and (2) are 0.432 and 0.433, respectively, implying good model fit.

5.1.3. Decomposition of Compensation–Skills Dual-Dimensional Test

To further identify the internal structure through which compensation–skill two-dimensional configuration relates to export technological sophistication, this study decomposes the two-dimensional configuration and separately tests single-dimensional effects and cross-comparisons across four configuration types. Table 6 presents the single-dimensional effects based on dummy-variable specifications, where firms are classified according to whether their corresponding indicators exceed the industry-year mean. Table 7 reports results using the continuous specification, in which the degree of configuration is measured by the log deviation of firm indicators relative to the industry-year mean. Table 8 provides the cross-comparison tests for the four configuration categories.

Table 6. Dimensional decomposition (dummy variables).

Variable	(1) $E_{si,it}$	(2) $E_{si,it}$	(3) $E_{si,it}$
$W_{H,it} > 0$	0.089 *** (9.073)		0.091 *** (9.299)
$R_{H,it} > 0$		−0.029 *** (−2.859)	−0.036 *** (−3.468)
S_{ize}	0.603 *** (33.142)	0.603 *** (33.092)	0.603 *** (33.117)
L_{ev}	0.353 *** (5.874)	0.349 *** (5.828)	0.348 *** (5.814)
R_{OA}	2.136 *** (24.874)	2.142 *** (24.792)	2.131 *** (24.870)
T_{op1}	0.020 (0.210)	0.009 (0.095)	0.020 (0.210)
F_{IXED}	−0.752 *** (−9.026)	−0.779 *** (-9.289)	−0.759 *** (−9.096)
A_{ge}	0.323 *** (2.652)	0.329 *** (2.705)	0.317 *** (2.603)
constant	−4.181 *** (−8.707)	−4.137 *** (−8.602)	−4.141 *** (−8.607)
year	yes	yes	yes
Firm	yes	yes	yes
N	20,978	20,978	20,978
Within R^2	0.437	0.432	0.438

Note: *** denote significance at the 1%. Robust t-statistics clustered at the firm level are reported in parentheses. Columns (1)–(2) separately include the dummy variables for the compensation dimension and the skill dimension, while Column (3) incorporates both dimensions simultaneously. The compensation dimension is defined based on whether the value exceeds the industry-year average compensation per employee, and the skill dimension is classified according to whether the share of R&D personnel surpasses the industry-year mean. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

Table 7. Dimensional decomposition (continuous variables).

Variable	(1) $E_{si,it}$	(2) $E_{si,it}$	(3) $E_{si,it}$
$W_{H,it} > 0$	0.334 *** (13.422)		0.347 *** (13.548)
$R_{H,it} > 0$		−0.023 * (−1.944)	−0.048 *** (−4.090)
S_{ize}	0.597 *** (34.252)	0.603 *** (33.113)	0.597 *** (34.282)
L_{ev}	0.355 *** (6.105)	0.348 *** (5.809)	0.344 *** (5.961)
R_{OA}	2.116 *** (25.196)	2.146 *** (24.787)	2.116 *** (25.195)
T_{op1}	0.042 (0.434)	0.009 (0.096)	0.042 (0.443)
F_{IXED}	−0.726 *** (−8.976)	−0.776 *** (−9.258)	−0.732 *** (−9.043)
A_{ge}	0.311 ** (2.574)	0.328 *** (2.690)	0.297 ** (2.456)
constant	−3.968 *** (−8.482)	−4.153 *** (−8.648)	−3.924 *** (−8.390)
year	yes	yes	yes
Firm	yes	yes	yes
N	20,978	20,978	20,978
Within R^2	0.457	0.432	0.459

Note: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Dummy variables are replaced by continuous values of relative deviation. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

Columns (1) and (2) in Table 6 examine the associations of the compensation dimension and skill dimension with firms' export technological sophistication under the dummy-variable specification, respectively. The results show that the estimated coefficient of Compensation-premium is 0.090 and significant at the 1% level; the estimated coefficient of Skill-underutilization is −0.029 and also significant at the 1% level. After including both dimensions simultaneously in Columns (3), the signs and statistical significance of the coefficients for Compensation-premium and Skill-underutilization remain unchanged. The regression results indicate that Compensation-premium is significantly and positively associated with firms' contemporaneous export technological sophistication, whereas Skill-underutilization exhibits a significant negative association with contemporaneous export technological sophistication.

Table 8. Cross-comparison of the four configuration types.

Variable	(1) $E_{si,it}$
$H_{H,it}$	0.057 *** (4.184)
$W_{H,it}$	0.087 *** (7.281)
$R_{H,it}$	−0.039 *** (−3.398)
S_{ize}	0.603 *** (33.115)
L_{ev}	0.348 *** (5.813)
R_{OA}	2.131 *** (24.864)
T_{op1}	0.020 (0.205)
F_{IXED}	−0.759 *** (−9.095)
A_{ge}	0.317 *** (2.606)
constant	−4.142 *** (−8.609)
year	yes
Firm	yes
N	20,978
Within R^2	0.438

Note: *** denote significance at the 1%. Robust t-statistics clustered at the firm level are reported in parentheses. Low-low is taken as the reference group. The High-high refers to the case where both compensation and the share of skilled personnel exceed their respective industry means. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

Estimates based on the continuous specification further validate the above findings. In Column (1) of Table 7, the estimated coefficient of Compensation-premium is 0.334 and significant at the 1% level. The estimated coefficient of Skill-underutilization in Column (2) is −0.023 and significant at the 10% level. When both dimensions are incorporated simultaneously, Column (3) reports an estimated coefficient of 0.347 for Compensation-premium and −0.048 for Skill-underutilization, both significant at the 1% level. Accordingly, the core finding that Compensation-premium is significantly and positively associated with firms' export technological sophistication while Skill-underutilization is significantly and negatively associated with export technological sophistication does not hinge on the choice between discrete and continuous indicators, demonstrating strong robustness.

The cross-comparison of the four configuration types intuitively illustrates the offsetting relationship between the two dimensions. Firms are classified into four groups: High-high, Compensation-premium, Skill-underutilization, and Low-low. Regressions are estimated using Low-low firms as the reference group, and the results are reported in Table 8. The coefficient of High-high is 0.057, the coefficient of Compensation-premium is 0.087, and the coefficient of Skill-underutilization is −0.039; all are significant at the 1% level. This result indicates that compared with Low-low firms, firms with High-high and Compensation-premium exhibit higher export technological sophistication, whereas firms with Skill-underutilization show lower export technological sophistication.

5.1.4. Construct Validity Test of the Skill Dimension

To test the construct validity of the core explanatory variable, this study conducts robustness checks by changing the measurement specification of the skill dimension. Whereas the baseline regression uses the share of R&D personnel to characterize firms' skill-dimensional configuration, this section adopts two educational attainment indicators—the share of employees with college education or above and the share of employees with postgraduate education or above—to reconstruct the dummy variables for High-high and Low-low. To ensure comparability of results across different specifications, the same model setup is applied to all three sets of indicators. Regressions take opposite-sign (crossed) configuration ($M_{ismatch_Negative}$) as the reference group and simultaneously include the

dummy variables for High-high and Low-low. The regression results are presented in Table 9.

Table 9. Construct validity test results.

Variable	(1) $E_{si,it}$	(2) $E_{si,it}$	(3) $E_{si,it}$
$H_{H,it}$ *(R&D)*	0.037 *** (3.184)		
$L_{L,it}$ *(R&D)*	−0.017 * (−1.899)		
S_{ize}	0.603 *** (33.127)	0.603 *** (33.153)	0.604 *** (33.145)
L_{ev}	0.357 *** (5.933)	0.354 *** (5.887)	0.353 *** (5.887)
R_{OA}	2.147 *** (24.811)	2.142 *** (24.859)	2.140 *** (24.769)
T_{op1}	0.012 (0.123)	0.009 (0.098)	0.015 (0.153)
F_{IXED}	−0.761 *** (−9.079)	−0.762 *** (−9.106)	−0.758 *** (−9.081)
A_{ge}	0.337 *** (2.761)	0.336 *** (2.757)	0.329 *** (2.692)
$H_{H,it}$ *(college degree)*		0.036 *** (3.563)	
$L_{L,it}$ *(college degree)*		−0.040 *** (−4.732)	
$H_{H,it}$ *(graduate student)*			0.042 *** (3.638)
$L_{L,it}$ *(graduate student)*			−0.035 *** (−3.477)
constant	−4.186 *** (−8.719)	−4.170 *** (−8.688)	−4.194 *** (−8.735)
year	yes	yes	yes
Firm	yes	yes	yes
N	20,978	20,978	20,978
Within R^2	0.433	0.434	0.434

Source: *** and * denote significance at the 1% and 10% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

Column (1) reports the baseline regression. Column (2) measures the skill dimension using the share of employees with college education or above. The results show that the coefficient of High-high is 0.036 and significant at the 1% level, while the coefficient of Low-low is −0.040 and also significant at the 1% level. Column (3) further adopts the share of employees with postgraduate education or above to measure the skill dimension. The coefficient of High-high is 0.042 and significant at the 1% level; the coefficient of Low-low is −0.035 and significant at the 1% level.

Overall, after controlling for firm and year fixed effects, the coefficients of High-high under the three skill measures are consistently positive and significant at the 1% level, ranging from 0.036 to 0.042. The coefficients of Low-low are all negative, though their significance levels vary. These regression results suggest that the core findings remain insensitive to alternative measures of the skill dimension, providing support for the convergent validity of the core construct and the robustness of our baseline conclusions.

5.1.5. Robustness Test

We subject the baseline association to a series of robustness checks, reported in Table 10.

Extended sample period: Column (2) re-estimates the model using the full sample spanning 2007–2023. The coefficient of High-high is 0.038 and significant at the 1% level. Its sign and significance are consistent with the baseline regression, indicating that the estimate is essentially unchanged when the sample is extended to 2007–2023.

Exclusion of outliers: Considering that firms with an extremely high share of R&D personnel may exhibit idiosyncratic observations, Column (3) drops outliers where the share of R&D personnel exceeds 60%. The estimated coefficient of High-high is 0.038 and significant at the 1% level, leaving the estimate essentially unchanged.

Replacement of the dependent variable: The baseline dependent variable is industry-level export technological sophistication adjusted by firm-level relative total factor productivity. The TFP estimation relies on labor input, which shares raw labor-related data with the explanatory variables. This may induce a spurious correlation between the dependent

variable and the core explanatory variable due to shared data sources. To address this concern, we replace the dependent variable with firm-level log value added, a measure of value capture that, unlike the baseline measure, uses no labor input. In Column (4), the coefficient of High-high is 0.038 and significant at the 1% level. Its sign and statistical significance are consistent with the baseline regression results.

Table 10. Regression results of robustness test.

Variable	(1) Baseline Regression	(2) Extended Sample	(3) Excluding Firms with R&D-Staff Share Above 60%	(4) Replace the Dependent Variable	(5) Firm-Specific Trends	(6) Industry-Specific Trends
$H_{H,it}$	0.037 ***	0.038 ***	0.038 ***	0.038 ***	0.041 ***	0.044 ***
	(3.184)	(3.252)	(3.240)	(2.898)	(3.723)	(3.886)
$L_{L,it}$	−0.017 *	−0.016 *	−0.017 *	−0.026 ***	−0.027 ***	−0.027 ***
	(−1.899)	(−1.743)	(−1.862)	(−2.894)	(−3.194)	(−3.208)
S_{ize}	0.603 ***	0.608 ***	0.599 ***	0.884 ***	0.510 ***	0.606 ***
	(33.127)	(34.025)	(32.962)	(47.349)	(16.830)	(33.557)
L_{ev}	0.357 ***	0.365 ***	0.357 ***	−0.160 ***	0.333 ***	0.360 ***
	(5.933)	(6.126)	(5.894)	(−2.615)	(5.069)	(6.235)
R_{OA}	2.147 ***	2.154 ***	2.142 ***	4.481 ***	1.711 ***	2.164 ***
	(24.811)	(25.018)	(24.551)	(37.142)	(18.346)	(25.236)
T_{op1}	0.012	0.035	0.013	0.082	0.066	−0.023
	(0.123)	(0.368)	(0.132)	(0.787)	(0.582)	(−0.249)
F_{IXED}	−0.761 ***	−0.745 ***	−0.749 ***	0.570 ***	−0.731 ***	−0.767 ***
	(−9.079)	(−9.035)	(−8.924)	(6.842)	(−8.608)	(−9.809)
A_{ge}	0.337 ***	0.334 ***	0.331 ***	0.300 ***	0.776 **	0.360 ***
	(2.761)	(2.896)	(2.707)	(2.661)	(2.030)	(3.040)
constant	−4.186 ***	−4.296 ***	−4.094 ***	−9.723 ***	−3.434 ***	−4.308 ***
	(−8.719)	(−9.192)	(−8.606)	(−19.563)	(−2.586)	(−8.853)
year	yes	yes	yes	yes	yes	yes
Firm	yes	yes	yes	yes	yes	yes
N	20,978	21,132	20,781	20,811	20,978	20,978
Within R^2	0.433	0.436	0.430	0.566	0.300	0.451

Note: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Column (1) presents the baseline regression using observations from 2015 to 2023 as the main sample. Column (2) reports the full-sample regression based on observations spanning 2007–2023. Column (3) excludes outliers where the share of R&D personnel exceeds 60%. Column (4) reports replacement of the dependent variable. Columns (5) and (6) control for firm-specific and industry-specific linear trends, respectively. Source: Author's calculations based on panel data of Chinese listed companies from 2007 to 2023, using Stata/MP 18.

Controlling for firm-specific linear trends: Column (5) incorporates firm-specific linear trends. The coefficient of High-high is 0.041 and significant at the 1% level. With firm-specific linear trends included, the estimate remains close to the baseline.

Controlling for industry-specific linear trends: Column (6) further incorporates industry-specific linear trends. The coefficient of High-high is 0.044 and significant at the 1% level. With industry-specific linear trends included, the estimate remains positive and significant.

Executive-pay contamination: Firm-level average compensation may be inflated by executive remuneration, which would mechanically raise the probability of being classified as a Compensation-premium firm. Using the CSMAR-derived control-variable database, we separately excluded the executive compensation and the number of executives from the numerator and denominator, and then restructured the per capita compensation. Complete information is available for all 21,317 main-sample firm-years. Re-estimating the baseline with this adjusted measure yields 0.034 (t = 3.03) against 0.041 (t = 3.50) under the identical construction without the adjustment. The coefficient has decreased but remains significant at the 1% level (Supplementary Material Table S3). The four types of configurational-deviation are preserved (High-high 0.031, t = 2.72; Low-low −0.017, t = −1.83), and the continuous compensation deviation is essentially unchanged (0.333, t = 13.37, versus 0.334 before the adjustment). As an additional, more aggressive check, we also removed total

management compensation (board, supervisory board and executives), for which the coefficient is 0.048 (t = 4.27); the Compensation-premium results are therefore not driven by executive pay.

Supplementary Material Table S4 also reports additional robustness checks (two-way clustering, excluding top and bottom 1% outliers of the dependent variable, controlling for the quadratic term of firm size, industry-plus-year joint fixed effect and industry-plus-year fixed effects). The coefficients retain the same signs as in the baseline specification, and the corresponding conclusions regarding statistical significance remain unchanged.

Overall, the coefficient of High-high remains positive and statistically significant across the above multiple sets of tests, which we read as evidence that the association is not an artifact of a single specification.

5.1.6. Endogeneity Test

To mitigate estimation bias arising from endogeneity issues such as bidirectional causality and omitted variables, this study further adopts the instrumental variable (IV) approach for supplementary tests, and the regression results are reported in Table 11.

Table 11. Results of the endogeneity assessment (IV estimates).

Setting	Second-Stage Coefficient	Second-Stage t-Statistic	K-P rk Wald F
Panel A: endogenous variable ($S_{mis,it}$)			
$L.S_{mis,it}$ IV (With Firm FE)	0.078	(0.709)	227.18
$S_{misO,it}$ IV (With Firm FE)	0.430	(1.025)	32.37
Panel B: Dummy Variables ($H_{H,it}$)			
$L.H_{H,it}$ IV (With Firm FE)	−0.114	(−1.597)	102.32
$H_{HO,it}$ IV (With Firm FE)	−0.828	(−1.296)	5.11
Panel C: endogenous variable(r_{it}^{S} including r_{it}^{W})			
Policy Exposure IV (Baseline skill deviation × Post-2018)	−0.048	(−1.307)	285.82

Note: Firm and year fixed effects are absorbed within the estimation sample using ivreghdfe (equivalent to ivreg2 combined with reghdfe). Consistent with other tables, singleton observations are dropped. The leave-one-out High-high instrument has a first-stage F of 5.11 and is therefore weak; the corresponding estimate is reported for transparency only and is not used for inference. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

Panel A first treats the $S_{mis,it}$ as the endogenous explanatory variable. With firm fixed effects controlled, when its one-period lag $L.S_{mis,it}$ is adopted as the instrumental variable, the 2SLS estimated coefficient is 0.078 (t = 0.709), and is no longer statistically significant. The Kleibergen–Paap rk Wald F statistic is 227.18. This study further uses the average value of configuration-deviation index among other firms within the same industry and year, denoted by $S_{misO,it}$, as an alternative instrument. The estimated coefficient is 0.430 (t = 1.025), likewise statistically insignificant. The corresponding Kleibergen–Paap rk Wald F statistic is 32.37. Controlling for firm fixed effects, the second-stage estimates for both instrumental variables are statistically insignificant. Accordingly, the instrumental variable approach yields no robust statistical evidence in models that primarily rely on within-firm variation.

Panel B further treats the $H_{H,it}$ as the endogenous variable. When its one-period lag ($L.H_{H,it}$) is used as the instrumental variable, the 2SLS estimated coefficient is −0.114 (t = −1.597), with a Kleibergen–Paap rk Wald F statistic of 102.32, indicating strong first-stage relevance, whereas the second-stage estimate is statistically insignificant. When the leave-one-out mean of High-high within the same industry and year ($H_{HO,it}$) is adopted as the instrumental variable, the estimated coefficient is −0.828 (t = −1.296), and the corresponding F statistic is merely 5.11, which points to a severe weak-instrument problem.

Neither estimation delivers robust causal evidence for the association between the High-high configuration and ETS. In particular, the second estimation suffers from both weak instruments and insufficient estimation precision; thus, we cannot infer the direction of the underlying relationship from this estimate.

It is important to emphasize that a large first-stage F statistic only demonstrates a statistical correlation between the instrumental variable and the endogenous explanatory variable; it cannot validate the exclusion restriction. Configuration deviation exhibits substantial temporal persistence. The one-period lagged indicator may directly affect current export technological sophistication through the persistent effects of configuration adjustment, rather than operating solely via the contemporaneous mismatch status. Although the leave-one-out industry mean eliminates the mechanical influence of the focal firm on the industry average, it may still be affected by common industry shocks, peer effects, and inter-firm knowledge and labor spillovers. The exogeneity of the aforementioned instrumental variables cannot be fully guaranteed. Accordingly, we treat the IV estimates as supplementary evidence rather than a complete solution to endogeneity concerns. The limitations regarding causal inference are further elaborated in the concluding section on study limitations.

To further explore the issue of reverse causality, this study has supplemented three sets of evidence (Tables 11 and 12).

Table 12. Reverse causality and temporal evidence.

Group	Timing	$H_{H,it}$	r_{it}^{W}	r_{it}^{S}
Panel A:Temporal precedence—joint specification (DV: $E_{si,it}$) Single-dim	current (t)	0.032 *** (2.823)	0.317 *** (11.802)	−0.002 (−0.187)
	lag (t − 1)	−0.028 ** (−2.342)	0.016 (0.720)	−0.024 ** (−2.186)
	lead (t + 1)	0.017 (1.506)	0.023 (0.949)	−0.014 (−1.222)
	N	13,334	13,334	13,334
Panel A:Temporal precedence—joint specification (DV: $E_{si,it}$) Joint specification	current (t)	–	0.328 *** (12.092)	−0.032 *** (−2.886)
	lag (t − 1)	–	0.022 (0.946)	−0.021 ** (−1.991)
	lead (t + 1)	–	0.027 (1.097)	−0.017 (−1.420)
	N	13,334	13,334	13,334
Panel B: Two-year-ahead extension(DV: $E_{si,it}$)	current (t)	0.035 *** (2.885)	0.341 *** (12.714)	−0.015 (−1.075)
	lead (t + 1)	−0.001 (−0.121)	0.008 (0.348)	−0.022 ** (−2.111)
	lead (t + 2)	0.008 (0.678)	0.014 (0.678)	−0.009 (−0.795)
	N	13,334	13,334	13,334
Panel C: Reverse-causality test (DV: current configuration)	current $E_{si,it}$	0.057 *** (3.925)	0.182 *** (11.409)	−0.016 (−0.705)
	$E_{si,it}$ lead (t + 1)	−0.043 *** (−3.542)	−0.056 *** (−5.565)	−0.040 *** (−2.652)
	$E_{si,it}$ lead (t + 2)	−0.010 (−0.859)	0.005 (0.574)	0.009 (0.541)
	N	10,381	10,381	10,381

Note: *** and ** denote significance at the 1% and 5% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Panels A–C retain only firm-years for which the required lagged and leading terms of the regressors are available, so that all columns within a panel share an identical sample: Panel A uses 2016–2022, Panel B 2015–2021 and Panel C 2016–2021; Panel C additionally requires the lagged term and is therefore estimated on the intersection of the Panel A and Panel B samples. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

First, we test the timing of the configuration. In the specification that includes the one-period lead alone, the coefficient on the lead of the High-high configuration is 0.017 (t = 1.506) and is statistically insignificant. When the one-period and two-period leads are included simultaneously, the corresponding coefficients are −0.001 (t = −0.121) and 0.008 (t = 0.678), respectively, and neither is statistically significant. The contemporaneous

coefficient remains positive and significant (0.035, t = 2.885), so the association is concentrated in the current period. The same pattern emerges for the continuous deviations: neither the one-period nor the two-period lead of the compensation deviation is significant (0.008, t = 0.348; 0.014, t = 0.678), whereas the one-period lead of the skill deviation is negative and significant (−0.022, t = −2.111), which is the opposite of the sign implied by a reverse-causality story. These results provide no temporal evidence that firms adopt the configuration in advance and achieve upgrading subsequently. A dedicated lead-lag test of anticipatory recruitment is reported in Supplementary Material Table S5.

Second, we directly test the reverse direction. Taking the configuration as the dependent variable, the coefficient on one-period-ahead export technological sophistication is −0.043 (t = −3.542) for the High-high configuration, −0.056 (t = −5.565) for the compensation deviation, and −0.040 (t = −2.652) for the skill deviation, whereas the two-period-ahead coefficients are insignificant (−0.010, t = −0.859; 0.005, t = 0.574; 0.009, t = 0.541). For comparability we also report the contemporaneous associations: 0.057 (t = 3.93) for High-high, 0.182 (t = 11.41) for the compensation deviation, and −0.016 (t = −0.71) for the skill deviation. Future export technological sophistication therefore carries no positive predictive power for the current configuration in any specification; if anything, the one-year-ahead coefficients are negative and significant, the opposite of the reverse prediction that firms with high export technological sophistication are better able to pay high wages and hire high-skilled workers and thereby form a High-high configuration. Because these negative coefficients may partly reflect mean reversion, we report them as the absence of positive predictive power rather than as evidence of a negative effect.

Third, we extend the instrumental-variable analysis. We extend the instrumental-variable analysis by constructing a policy-exposure instrument (Z_{it}), as the interaction between each firm's average skill deviation during the baseline period (2015–2017) and an indicator for 2018 and subsequent years. This design exploits the increase in the R&D expense super-deduction rate from 50% to 75% under Caishui (2018) No. 99. The instrument strongly predicts contemporaneous skill deviation in the first stage: the coefficient on Z_{it} is −0.325 (t = −16.906), and the Kleibergen–Paap rk Wald F-statistic equals 285.82. In the two-stage least squares regression, the coefficient of skill deviation is −0.048 (t = −1.307), which is broadly consistent with the ordinary least squares estimate (−0.048) yet lacks sufficient precision. Nevertheless, the event-study estimates underlying this design show that the association between baseline exposure and contemporaneous skill deviation declines monotonically from 0.544 in 2015 to 0.045 in 2022, without a discernible discontinuity around the policy change in 2018. This smooth convergence suggests that the instrument may partly capture mean reversion rather than policy-induced variation, thereby raising concerns about the exclusion restriction. We therefore treat this specification only as supplementary sensitivity evidence rather than for definitive causal identification.

Taken together, these results do not support the reverse-causality explanation: the configuration exhibits no anticipatory variation, future export technological sophistication does not positively predict it, and the instrumental-variable exercise, although imprecise, does not reverse the sign of the association. What remains unresolved is precision and the exclusion restriction rather than the direction of the relationship. Accordingly, we interpret the baseline coefficient on the High-high configuration (0.037) as a conditional association that is not overturned by the reverse-causality tests, rather than as an identified causal effect, and we do not claim that the endogeneity problem has been fully resolved.

5.2. Association Channel Tests

This study does not conduct rigorous causal mediation identification. Instead, it only examines the transmission relationships of firm innovation and AI orientation between High-high and export technological sophistication from an associational perspective.

5.2.1. Baseline Association Channel Results

The results of the association channel tests are presented in Table 13. First, we examine the association between High-high and innovation output. The regression coefficient is 0.049 and significant at the 10% level. Column (3) further decomposes the differentiated associations of the compensation dimension and skill dimension with firm innovation. The coefficient of the compensation dimension is −0.060 and that of the skill dimension is 0.060; both are significant at the 5% level. The results indicate that Skill-underutilization has a significant positive association with firm innovation, while Compensation-premium exhibits a significant negative association with firm innovation. Second, we explore the association between High-high and AI orientation. The coefficient is 0.073 and significant at the 1% level, suggesting that High-high is positively associated with firms' AI orientation.

Table 13. Regression results of association channel tests.

Variable	(1) lnP_{AT}	(2) lnA_I	(3) lnP_{AT} by Sub-Dimensions
$H_{H,it}$	0.049 * (1.731)	0.073 *** (3.553)	
$W_{H,it} > 0$			−0.060 ** (−2.356)
$R_{H,it} > 0$			0.060 ** (2.374)
S_{ize}	0.603 *** (15.775)	0.166 *** (6.527)	0.603 *** (15.759)
L_{ev}	−0.037 (−0.318)	0.013 (0.166)	−0.031 (−0.269)
R_{OA}	−0.083 (−0.465)	−0.172 (−1.500)	−0.069 (−0.390)
T_{op1}	0.522 ** (2.198)	−0.117 (−0.767)	0.515 ** (2.161)
F_{IXED}	0.048 (0.294)	−0.102 (−0.997)	0.040 (0.243)
A_{ge}	−0.072 (−0.229)	0.272 (1.211)	−0.058 (−0.185)
constant	−9.718 *** (−8.370)	−3.357 *** (−4.134)	−9.750 *** (−8.391)
year	yes	yes	yes
Firm	yes	yes	yes
N	20,978	20,978	20,978
Within R^2	0.046	0.010	0.047

Note: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Firm and year fixed effects are controlled for in all columns. Columns (1) and (2) examine the associations of High-high with firm innovation and AI orientation, respectively. Column (3) further decomposes the differentiated associations of the compensation dimension and skill dimension with firm innovation. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

This implies that the overall association between High-high and innovation output is relatively weak, consistent with an offsetting pattern: the negative association linked to relatively high wage compensation partly offsets the positive association from a higher share of skilled personnel.

5.2.2. Channel Overlap Tests

Table 14 further explores the association between firm innovation and AI orientation. The estimated coefficient of AI orientation on firm innovation is 0.076, while the coefficient of firm innovation on AI orientation is 0.035; both are significant at the 1% level. This indicates that the two channels exhibit strong contemporaneous correlation and their boundaries are not fully separated. Nevertheless, these regressions cannot identify the causal direction, so we cannot conclude that there exists a bidirectional causal relationship between AI orientation and firm innovation.

Table 14. Channel overlap tests and alternative AI measure.

Variable	(1) $A_I \rightarrow P_{AT}$ lnP_{AT}	(2) $P_{AT} \rightarrow A_I$ lnA_I	(3) MDA $E_{si,it}$	(4) MDA $E_{si,it}$
lnA_I	0.076 *** (5.863)			
S_{ize}	0.590 *** (15.416)	0.145 *** (5.567)	0.146 *** (6.155)	0.146 *** (6.148)
L_{ev}	−0.041 (−0.350)	0.010 (0.128)	−0.032 (−0.424)	−0.034 (−0.452)
R_{OA}	−0.071 (−0.398)	−0.171 (−1.494)	−0.172 (−1.587)	−0.173 (−1.603)
T_{op1}	0.530 ** (2.220)	−0.136 (−0.880)	−0.291 ** (−2.036)	−0.291 ** (−2.031)
F_{IXED}	0.049 (0.297)	−0.116 (−1.128)	−0.022 (−0.229)	−0.029 (−0.310)
A_{ge}	−0.096 (−0.305)	0.269 (1.196)	0.554 ** (2.555)	0.550 ** (2.536)
lnP_{AT}		0.035 *** (5.874)		
$H_{H,it}$			0.045 ** (2.258)	
$S_{mis,it}$				−0.006 (−0.156)
constant	−9.439 *** (−8.104)	−2.979 *** (−3.635)	−3.977 *** (−5.081)	−3.952 *** (−5.041)
year	yes	yes	yes	yes
Firm	yes	yes	yes	yes
N	20,978	20,978	20,836	20,836
Within R^2	0.048	0.012	0.010	0.010

Note: *** and ** denote significance at the 1% and 5% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Columns (1) and (2) reveal a significant bidirectional overlap between AI orientation and innovation. Column (3) measures AI orientation using the more restrictive keyword frequency extracted from the Management Discussion and Analysis (MDA) section of annual reports, and the coefficient of High-high remains significant. Column (4) serves as a benchmark specification for the continuous mismatch index under the same measure and yields an insignificant estimate. the type-based measure is more strongly associated with the outcome than the intensity-based measure. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

To mitigate the concern that full annual report word counts may be affected by general information disclosure, this study further adopts the frequency of artificial intelligence keywords extracted from the Management Discussion and Analysis (MDA) section as an alternative measure. Under this specification, the coefficient of High-high is 0.045 and significant at the 5% level, while the coefficient of the continuous skill mismatch index is −0.006 and statistically insignificant. This result shows that the positive association of High-high remains robust after narrowing the textual scope of the AI indicator, whereas the continuous skill-mismatch proxy remains insignificantly associated with the outcome. This further supports the argument that configuration type carries greater explanatory power than the degree of skill deviation.

5.2.3. Behavioral Validation and Measurement Robustness of the AI-Orientation Channel

Considering that textual keyword frequencies may merely reflect firms' rhetorical attention, this study constructs an AI patent indicator as a behavioral proxy based on the Patent Classification System for Key Digital Technologies (2023) issued by the National Intellectual Property Administration. Multi-dimensional robustness checks are performed for the AI-orientation channel, and the results are shown in Table 15. The regression results reveal that the coefficients of High-high are consistently positive and statistically significant when using AI patent applications, AI patent stock, and AI invention patents as dependent variables. For comparison, when full annual-report keyword frequency and MDA keyword frequency are adopted as dependent variables, the coefficients of High-high are 0.078 and 0.046, respectively, and remain positively significant. These findings support the robustness of the estimated results for the AI-orientation channel and are interpreted as associations rather than causal effects.

Table 15. Behavioral validation of the AI-orientation channel: aggregate measure.

Variable	(1) AI Patents	(2) AI Patents Inventory	(3) Only Invention Patents	(4) AI Word Frequency	(5) MDA Word Frequency
$H_{H,it}$	0.086 ***	0.054 **	0.101 ***	0.078 ***	0.046 **
	(3.422)	(2.387)	(3.944)	(3.767)	(2.326)
$L_{L,it}$	−0.007	−0.018	−0.017	0.023	0.006
	(−0.320)	(−0.951)	(−0.848)	(1.375)	(0.374)
S_{ize}	0.395 ***	0.385 ***	0.366 ***	0.166 ***	0.146 ***
	(13.536)	(14.252)	(12.638)	(6.521)	(6.154)
L_{ev}	−0.250 ***	−0.325 ***	−0.215 **	0.012	−0.033
	(−2.673)	(−3.676)	(−2.376)	(0.150)	(−0.428)
R_{OA}	−0.105	−0.805 ***	−0.109	−0.172	−0.172
	(−0.764)	(−6.602)	(−0.790)	(−1.500)	(−1.587)
T_{op1}	−0.085	−0.003	0.038	−0.120	−0.292 **
	(−0.416)	(−0.015)	(0.183)	(−0.786)	(−2.041)
F_{IXED}	0.090	0.255 **	0.122	−0.108	−0.023
	(0.689)	(1.961)	(0.983)	(−1.054)	(−0.244)
A_{ge}	−0.323	0.308	−0.356	0.273	0.554 **
	(−1.222)	(1.160)	(−1.350)	(1.215)	(2.557)
constant	−6.234 ***	−7.119 ***	−5.758 ***	−3.366 ***	−3.979 ***
	(−6.770)	(−7.611)	(−6.286)	(−4.144)	(−5.084)
year	yes	yes	yes	yes	yes
Firm	yes	yes	yes	yes	yes
N	20,978	20,978	20,978	20,978	20,836
Within R^2	0.029	0.046	0.025	0.010	0.010

Note: *** and ** denote significance at the 1% and 5% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Table 15 reports the action-based validation results. The dependent variables are, in sequence, the log-transformed AI patent applications (plus one), log-transformed cumulative patent stock (plus one), invention patents, and the two annual-report keyword frequency measures for comparison. AI patent data are identified following the Patent Classification System for Key Digital Technologies (2023) issued by the National Intellectual Property Administration, covering listed firms themselves as well as their subsidiaries, associated enterprises, and joint ventures. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

5.2.4. Examination by Sub-Dimensions

Table 16 presents the sub-dimensional test results. The coefficients of Skill-underutilization on AI patent applications and patent stock are 0.094 and 0.079, respectively, both significant at the 1% level. The coefficients of Compensation-premiumon AI patent applications and patent stock are −0.017 and −0.015, respectively, and both are statistically insignificant. Meanwhile, the skill dimension exhibits a significant positive association with firm innovation, while the compensation dimension shows a significant negative association with firm innovation. Accordingly, the positive associations of High-high with AI patenting activities and firm innovation are associated mainly with the skill dimension rather than with Compensation-premium alone. This finding further clarifies the association pattern underlying the High-high.

The theoretical discussion argues that High-high firms use their compensation advantage to attract and accumulate high-skilled talent, a mechanism that has so far remained inferential. We therefore construct three observables: net growth of R&D staff, net growth of technical staff, and the change in the high-skill share (Supplementary Material Table S6). Because Chinese listed firms disclose period-end headcounts by staff category but not hires or separations, true turnover rates are not observable; headcount-based growth measures therefore capture net talent accumulation rather than retention of incumbent staff per se. Descriptively (Panel A), High-high firms combine rapid net growth of R&D staff (0.100, versus 0.071 for Low-low and 0.026 for Compensation-premium firms) with an increase in the high-skill share of 1.3 percentage points (versus 0.6 for Low-low). Compensation-premium firms show a slightly larger increase in the high-skill share (1.4 percentage points)

but the slowest net growth of R&D staff, whereas Skill-underutilization firms show the fastest net growth of R&D staff (0.121); these cross-group differences are descriptive and are not interpreted as causal. In fixed-effects regressions with the full control set (Panel B), High-high is associated with higher net growth of R&D staff (0.042, t = 3.26), a larger increase in the high-skill share (0.7 percentage points, t = 2.71), and a larger increase in the R&D-staff ratio (2.00 percentage points, t = 11.41). In the four-group specification, in which Compensation-premium and Skill-underutilization firms form the reference group, High-high remains positive and significant (0.036, t = 2.71) while Low-low is associated with lower net growth of R&D staff (−0.033, t = −3.24). The ratio-based variant of the R&D-staff measure (the growth ratio, a monotone transformation of net growth that we report for transparency) is not significantly related to configuration (0.025, t = 0.58), and the same holds for technical staff (−0.034, t = −1.46); because such ratios are disproportionately influenced by firms with a small base headcount, we base our reading on the log-difference measure. Using the executive-pay-adjusted configuration, the coefficient on High-high for net growth of R&D staff is 0.024 (t = 1.85), i.e., attenuated but still positive. We read this as evidence consistent with the skill-accumulation side of the channel; the compensation-to-retention link remains an interpretation that our data cannot test directly, and the corresponding statements in the theory section are now phrased as associations rather than as a tested mechanism. The evidence is therefore corroborative rather than a fully identified mediation result.

Table 16. Results of Examination by Sub-dimensions.

Variable	(1) A_I Patents	(2) A_I Patent Inventory	(3) $\ln P_{AT}$	(4) $E_{si,it}$
$W_{H,it} > 0$	−0.017 (−0.812)	−0.015 (−0.826)	−0.060 ** (−2.356)	0.091 *** (9.299)
$R_{H,it} > 0$	0.094 *** (4.239)	0.079 *** (3.929)	0.060 ** (2.374)	−0.036 *** (−3.468)
S_{ize}	0.395 *** (13.523)	0.385 *** (14.230)	0.603 *** (15.759)	0.603 *** (33.117)
L_{ev}	−0.243 *** (−2.591)	−0.319 *** (−3.601)	−0.031 (−0.269)	0.348 *** (5.814)
R_{OA}	−0.092 (−0.674)	−0.794 *** (−6.524)	−0.069 (−0.390)	2.131 *** (24.870)
T_{op1}	−0.087 (−0.426)	−0.006 (−0.033)	0.515 ** (2.161)	0.020 (0.210)
F_{IXED}	0.092 (0.699)	0.256 * (1.960)	0.040 (0.243)	−0.759 *** (−9.096)
A_{ge}	−0.312 (−1.179)	0.320 (1.205)	−0.058 (−0.185)	0.317 *** (2.603)
constant	−6.292 *** (−6.830)	−7.182 *** (−7.689)	−9.750 *** (−8.391)	−4.141 *** (−8.607)
year	yes	yes	yes	yes
Firm	yes	yes	yes	yes
N	20,978	20,978	20,978	20,978
Within R^2	0.029	0.047	0.047	0.438

Note: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics clustered at the firm level are reported in parentheses. Columns (1) and (2) adopt AI patent measures as dependent variables. Columns (3) and (4) use firm innovation and export technological sophistication as dependent variables, respectively, to facilitate comparison of the differentiated associations across the two dimensions. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

Cross-proxy assessment of the AI-related measures. Our four measures represent related but conceptually different observable AI-related outcomes. Supplementary Material Table S7 examines whether the keyword measures may partly capture disclosure behavior. Panel A reports Pearson correlations. The two disclosure-based measures are closely related (r = 0.94), but this relationship is partly mechanical because the MDA section is contained

in the annual report from which the broader keyword measure is constructed. We therefore do not treat this correlation as independent evidence of convergent validity. AI patents and the number of AI subsidiaries are also strongly correlated (r = 0.92), whereas the four correlations across the disclosure- and action-based measures range from 0.45 to 0.51. Panel B examines whether the disclosure-based measures covary with an organizational-action-based proxy. When entered separately, both keyword measures are positively associated with ln(1 + number of AI subsidiaries) (annual-report keywords: 0.029, (t = 4.658); MDA keywords: 0.022, (t = 3.617). When the two keyword measures and AI patents are included jointly, however, the keyword coefficients are close to zero (annual-report keywords: −0.005, (t = −0.807); MDA keywords: 0.002, (t = 0.282). This pattern is consistent with the possibility that keyword frequencies partly capture reporting behavior, but it is not sufficient to establish that interpretation. Given the high collinearity between the keyword measures, the specification cannot determine whether the change reflects multicollinearity, overlap with AI patenting, or a genuine absence of incremental association. Panel C re-estimates the configuration association using each measure as the dependent variable. The coefficient is positive for annual-report keywords (0.062, t = 3.129), MDA keywords (0.045, t = 2.258), and AI patents (0.087, t = 3.474). It is also positive but only marginally significant for the number of AI subsidiaries (0.022, t = 1.883, p = 0.060). The estimates are therefore directionally similar across the two measurement families, although the evidence is weakest for AI-subsidiary establishment.

These comparisons are subject to two important limitations. First, none of the four measures directly captures the operational deployment of AI within firms. Keyword frequencies measure corporate disclosure, patents measure innovation output that may not be deployed, and AI subsidiaries may be established for reasons including tax, financing, or market access. Second, cross-measure agreement does not constitute formal convergent validity because the measures are partly nested and capture different outcomes. We therefore interpret the findings as evidence concerning AI-related disclosure, patenting, and AI-subsidiary establishment, rather than direct evidence of in-house AI adoption. Establishing operational deployment would require firm-level records of AI systems in use, which are not systematically available in our data for the firms and period examined.

5.3. Heterogeneity Analysis

To identify the heterogeneous associations of High-high with firms' export technological sophistication, this study avoids inferring group differences through simple split-sample regressions. Instead, it adopts a full-sample interaction model combined with the Wald test for intergroup coefficient differences to ensure rigorous identification of heterogeneity. The results are reported in Table 17.

With the western region as the reference group, the coefficient of the interaction term between High-high and the eastern region is 0.068 (t = 1.532), and the coefficient of the interaction term for the central region is 0.075 (t = 1.355). Neither reaches conventional statistical significance. The Wald test further indicates that the coefficient differences between eastern and western regions, central and western regions, and eastern and central regions are all insignificant, with p-values of 0.126, 0.176, and 0.172, respectively. The estimated coefficients are slightly larger in the eastern and central regions relative to the western region, yet the between-group differences are statistically insignificant. These results only reveal numerical disparities and do not provide sufficient evidence of significant regional heterogeneity in the association of High-high.

Table 17. Regression results of regional heterogeneity.

Variable	$E_{si,it}$
$H_{H,it}$	−0.023 (−0.531)
$H_{H,it}\times$ east	0.068 (1.532)
$H_{H,it}\times$ central	0.075 (1.355)
S_{ize}	0.603 *** (33.100)
L_{ev}	0.354 *** (5.885)
R_{OA}	2.146 *** (24.800)
T_{op1}	0.009 (0.089)
F_{IXED}	−0.768 *** (−9.159)
A_{ge}	0.336 *** (2.759)
constant	−4.190 *** (−8.728)
year	yes
Firm	yes
N	20,978
Within R^2	0.433

Note: *** denote significance at the 1% levels. Robust t-statistics clustered at the firm level are reported in parentheses. The western region is set as the regional reference group. The *p*-values from the Wald tests for between-group coefficient differences are 0.126, 0.176 and 0.172, respectively. Source: Author's calculations based on panel data of Chinese listed companies from 2015 to 2023, using Stata/MP 18.

6. Discussion

6.1. Discussion on Baseline Regression Results

The baseline regression results show that $S_{mis,it}$, which captures the overall degree of deviation, is not significantly associated with firms' export technological sophistication. In contrast, firms with High-high exhibit higher export technological sophistication than those with the reverse configuration. This finding suggests that the relevant margin is not the magnitude of deviation but its direction and the combination of the two dimensions.

The positive association of High-high is consistent with a potential complementary relationship between above-benchmark compensation and skilled labor. In principle, compensation above the industry benchmark may be associated with employee screening, effort incentives, and retention, which could in turn be related to access to high-ability workers, lower turnover risk, and greater production efficiency. These mechanisms, however, cannot be directly distinguished in our data. Chinese listed firms disclose period-end headcounts but not hires, separations, or employee-level retention. Supplementary Material Table S6 therefore examines changes in workforce size and composition rather than retention itself. High-high is associated with faster net growth in R&D staff and a larger increase in the high-skill share, whereas the broader headcount-growth measures are unrelated to configuration. These results are consistent with skill accumulation, although the observed changes may reflect hiring, attrition, or employee reclassification. A larger skilled workforce may broaden firms' knowledge base and strengthen their capacity for technology identification, absorption, and re-innovation. Thus, the High-high results are consistent with a potentially synergistic combination of compensation and skilled labor, rather than providing direct evidence of a retention, engagement, or efficiency mechanism.

Further dimensional decomposition reveals that the compensation and skill dimensions yield differentiated effects with opposite signs. Compensation-premium is associated with higher contemporaneous ETS but with lower innovation output. Compensation-premium attracts workers with higher productivity and stronger adaptability to complex production tasks. Meanwhile, the efficiency-wage interpretation would imply lower employee turnover, though our headcount-ratio measures are too crude to test this directly. These effects are rapidly reflected in firms' contemporaneous productivity and export

technological sophistication. Nevertheless, rising compensation increases firms' recurring expenditures. Under internal financing constraints, it may be associated with lower innovation output through the crowding out of resources available for R&D. The costs of skilled personnel are recognized immediately, whereas R&D outputs generally require knowledge accumulation, experimental development and commercialization before translating into tangible outcomes. This creates a temporal mismatch between upfront current costs and future benefits. The High-high combines both dimensions: this would imply a potential association between market-oriented compensation and talent retention, while the pool of high-skilled personnel is associated with innovation capacity and with a rough balance between contemporaneous upgrading and long-run accumulation.

It should also be noted that the positive association of High-high does not imply "the higher the configuration, the better". This study only identifies the association between relative allocation above the industry mean and higher ETS. It does not establish that such allocations are optimal or rule out potential efficiency losses arising from excessive labor costs and resource misallocation beyond reasonable thresholds. The findings do not necessarily contradict static matching theory. Rather, they are consistent with the possibility that contemporaneous matching efficiency and longer-term upgrading involve different considerations. Under the conceptual framework developed in Section 3.1, above-benchmark compensation and a highly skilled workforce may constitute input-side conditions conducive to sensing, seizing, and reconfiguring. However, these capabilities, as well as the firm specificity and imitability of the two resources, are not directly measured in this study. They should therefore be understood as theoretical interpretations rather than empirically established mechanisms. The observed asymmetry between the two dimensions is also consistent with the possibility that the returns to capability building emerge later than its costs, although the present evidence cannot establish this temporal mechanism or exclude alternative explanations. Overall, the findings highlight the potential importance of distinguishing short-run efficiency objectives from longer-term value-chain upgrading objectives in the context examined.

6.2. Discussion on Association Channel Test

The results of channel association tests reveal that High-high is positively associated with firms' innovation output and AI orientation. Meanwhile, the innovation channel and AI-orientation channel are significantly correlated, indicating that these channels overlap rather than being fully independent.

Firms with a High-high configuration are characterized by a relatively large pool of highly skilled employees. Such a workforce may provide a human-capital foundation conducive to internal knowledge exchange, knowledge diffusion between production and R&D functions, and patenting activity, although these processes are not directly observed in our data. Empirically, the Skill-underutilization dimension is positively associated with innovation output, whereas the accompanying Compensation-premium dimension is associated with fewer R&D resources and lower innovation output. These opposing associations may partially offset one another. High-high is also associated with faster net growth in R&D staff and greater R&D investment, but the headcount evidence does not measure employee retention. Taken together, the results are consistent with the coexistence of additional labor costs and capacity accumulation and with higher patenting and ETS. They do not, however, establish optimal compensation–skill matching or a causal capacity-building mechanism.

Turning to the AI-related outcomes, the association is concentrated primarily among firms with a larger stock of highly skilled personnel. One possible interpretation is that skilled employees provide capabilities relevant to AI-related innovation and organizational

adjustment, including algorithmic problem-solving, process redesign, and human–machine coordination. These activities, however, are not directly observed in our data. Consistent with this interpretation, firms offering relatively high compensation without a corresponding stock of technical personnel do not exhibit a comparable association with AI-related disclosure and patenting. The AI-related measures are also positively associated with ETS in the cross-section. Taken together, these patterns are compatible with a potential AI-related pathway, but they neither establish a causal mechanism nor provide direct evidence of operational AI deployment within firms.

These two associated channels are not entirely separate transmission paths. AI orientation itself constitutes an important component of firms' innovation activities, and AI patents count toward firms' innovation output. Improvements in firms' overall innovation capacity, in turn, are consistent with a greater potential to assimilate and adopt AI technologies. This finding suggests that the High-high human capital configuration provides a talent base for both activities, which are jointly associated with firms' export upgrading. At the same time, it must be clarified that the available evidence only demonstrates statistical associations among variables. We cannot fully rule out unobserved firm strategic and governance factors that simultaneously drive human-capital allocation, innovation, AI orientation and export upgrading. Therefore, the results should not be overinterpreted as a definitive causal mediation chain, which constitutes a limitation of this study. Because the IV estimates are imprecise, the OLS magnitudes reported here may also be affected by reverse causality or omitted variables and should be read as conditional associations.

6.3. Discussion on Heterogeneity Analysis Results

The regional heterogeneity tests do not identify statistically significant differences: the interaction terms of High-high with the eastern and central regions fail to reach statistical significance, and the Wald tests likewise detect no significant between-group differences (p = 0.126, 0.176 and 0.172). Because the western region serves as the reference category, the implied High-high association is −0.023 in the western region, 0.045 in the eastern region and 0.052 in the central region; these linear combinations are not individually distinguishable from zero. One tentative reading—which our design cannot test directly, because we do not observe regional cost or talent-supply conditions—is that region-specific cost factors work in opposite directions and partly offset one another. In the eastern region, labor-market competition is intense and housing and living costs are high, so that a larger share of the pay premium would be absorbed by the cost of competing for high-skilled talent; in the central and western regions, high-skilled talent is relatively scarce but land and labor are cheaper, so that firms reaching a high-compensation–high-skill configuration face milder cost pressure. If these two forces roughly cancel, cross-regional differences would be small—an interpretation that is not inconsistent with the insignificant interaction terms, but that remains a conjecture rather than a tested mechanism. This does not mean that regional talent endowments are irrelevant. For firms in the central and western regions, the binding constraint is more likely to lie in whether a High-high configuration can be attained at all—raising an above-industry pay premium while simultaneously building a high-skill workforce is harder when skilled workers are scarce and outward migration is substantial—rather than in whether such a configuration would be accompanied by upgrading once it is in place. Neither the insignificant coefficients nor our design, however, establish that the configuration operates in the western region, where the reference-group point estimate is negative and imprecisely estimated.

7. Conclusions and Policy Recommendations

7.1. Research Conclusions

1. The coefficient of $S_{mis,it}$ is -0.043 and fails the significance test. In contrast, the baseline two-way fixed-effects estimate for the High-high configuration is 0.037 and significant at the 1% level. When the skill dimension is measured by the share of R&D personnel, the share of employees with college education or above, and the share of employees with postgraduate education or above, the coefficients of High-high remain significantly positive and range from 0.036 to 0.042. This suggests that our estimates are not sensitive to a specific skill measure. The Low-low configuration, however, exhibits no stable positive association. Given the imprecision of the instrumental-variable estimates, the baseline conditional association estimated within the two-way fixed-effects framework may partly reflect reverse causality or time-varying omitted variables and is therefore interpreted as associational rather than causal evidence.

2. The compensation and skill configurations exhibit associations of opposite signs across the two outcome dimensions, a pattern consistent with what we term the "value paradox". The Compensation-premium is significantly positively associated with export technological sophistication, with a coefficient of 0.091, yet it bears a significant negative association with firm innovation output, with a coefficient of -0.060. Skill-underutilization is significantly positively associated with firm innovation output, with a coefficient of 0.060, but shows a significant negative association with export technological sophistication, with a coefficient of -0.036. Comparisons across the four configuration types reveal that relative to Low-low configuration firms, Compensation-premium is linked to higher export technological sophistication, while Skill-underutilization exhibits a negative association. Accordingly, Compensation-premium is more consistent with advantages in the contemporaneous export product mix, whereas Skill-underutilization is more closely related to innovation accumulation. Since this study does not directly identify the realization process of dynamic benefits, the above findings should be interpreted as structural differences across outcome dimensions rather than a strict intertemporal causal mechanism.

3. Channel association tests indicate that High-high is positively associated with firms' innovation output and AI orientation. These two channels exhibit substantial overlap. Sub-dimensional results further reveal that AI patenting activities are mainly associated with R&D staffing configuration rather than with Compensation-premium alone. Subject to inherent limitations of the measurement approach, the above evidence constitutes channel consistency checks and cannot be interpreted as rigorous mediation effects or causal transmission mechanisms.

4. We find no significant regional heterogeneity in the associations of High-high. Taking the western region as the reference group, the coefficient of the interaction term between High-high and the eastern region is 0.068, while the coefficient for the interaction term with the central region is 0.075; neither is statistically significant. The Wald test further confirms that there are no significant coefficient differences between eastern and western regions, central and western regions, and eastern and central regions.

7.2. Policy Recommendations

Because the coefficients of interest are interpreted as conditional associations estimated within a two-way fixed-effects framework, and because the instrumental-variable estimates are too imprecise to corroborate the magnitudes of the baseline associations, these suggestions should be understood as highlighting configuration patterns that firms and policymakers may wish to consider, rather than as prescriptive recommendations based on causally identified effects. They are not prescriptions that are guaranteed to

produce upgrading; in particular, none of them should be read as implying that raising compensation by itself is associated with higher export technological sophistication.

1. Support a high-compensation–high-skill configuration: a pattern associated with firm upgrading.

Our estimates show that the High-high configuration—above-industry compensation combined with a high-skill share—is positively associated with ETS. Compensation-premium firms also exhibit higher contemporaneous ETS, but they are negatively associated with innovation output. Among the configurations examined, High-high is the only one whose point estimates are positive for both ETS and innovation output. However, this pattern does not by itself establish that High-high performs significantly better than every alternative configuration; such a conclusion requires direct tests of the relevant coefficient differences. The findings are consistent with a capability-building interpretation, but they do not rule out alternative explanations such as firm selection, reverse causality, unobserved productivity, or inefficient compensation expenditure. Accordingly, the results should not be interpreted as showing that firms can achieve upgrading simply by moving toward a High-high configuration. They instead suggest that the joint alignment of compensation and workforce skills warrants further causal investigation. Any policy support for such alignment—for example, carefully targeted and evaluated hiring or training incentives—should therefore be regarded as a testable policy option rather than a direct implication of the estimates.

2. Associate skill allocation with firm upgrading through technology adoption and its associated enabling patterns.

Skill allocation is associated with firm upgrading, and the observed patterns are consistent with a role for AI orientation and related complementary capabilities. Policymakers may consider supporting the joint development of AI-related capabilities and high-skilled talent, subject to careful policy evaluation. Two qualifications are important. First, the evidence concerning AI orientation is sensitive to how AI activity is measured. Second, our channel analyses provide consistency checks rather than formal mediation estimates and therefore do not establish that AI orientation transmits the association between skill allocation and upgrading. Policy evaluations should consequently triangulate multiple verifiable, activity-based indicators—such as AI patenting and the establishment and substantive operation of AI-focused subsidiaries—rather than relying primarily on firms' narrative disclosures, which may not correspond to actual implementation.

3. Build an open, collaborative innovation ecosystem along the industrial chain.

Such an ecosystem would lower the cost of reallocating high-skilled labor at the level of the industrial chain. Policies could reduce barriers to knowledge and technology flows—between firms, universities, and research institutes—so that high-skilled talent can be reallocated on a broader scale. Regional innovation alliances and shared R&D platforms may support the mobility of underutilized high-skilled labor within industrial clusters and are consistent with more efficient regional human-capital allocation. Regionally coordinated ecosystems may also reduce duplicative R&D investment, an efficiency gain that is itself consistent with more sustainable industrial development. In this sense, knowledge spillovers from high-skilled workers are expected to cross firm boundaries and diffuse along the industrial chain. An open, coordinated ecosystem of this kind may be associated with a higher collective technological level of industrial clusters and with stronger bargaining positions within GVCs—an association rather than a guaranteed upgrading effect.

4. Address SMEs' skill-upgrading dilemmas amid the polycrisis.

Compared with large firms, small and medium-sized enterprises (SMEs) generally face tighter liquidity constraints, greater difficulty in attracting skilled workers, and less organizational slack, particularly when multiple economic shocks overlap. Because the

joint development of competitive compensation and workforce skills may require substantial fixed investment and organizational capacity, these constraints could make such configurations more difficult for SMEs to attain. Overlapping crises may further limit their ability to sustain human-capital investment [38]. Prior research suggests that growing SMEs often encounter constraints in both managerial capacity and skill supply during expansion and that organizational flexibility, external talent networks, and gradual approaches to skill development may help alleviate these constraints [39]. Possible policy responses include shared R&D and training platforms, carefully targeted and evaluated tax incentives or subsidies for skill investment, and cluster-based joint-recruitment or talent-sharing arrangements. Such measures could reduce the constraints SMEs face in investing in skills and organizational capabilities, although whether they promote the adoption of upgrading-related configurations remains an empirical question. In this limited sense, they are broadly aligned with the joint pursuit of decent work and economic growth (SDG 8) and industry, innovation, and infrastructure (SDG 9).

7.3. Limitations and Future Research

This study has several noteworthy limitations that point toward avenues for future research. First, the skill dimension adopts the share of R&D personnel as a proxy for high-skilled human capital. Although validated by construct checks using educational attainment measures, this indicator cannot directly quantify the gap between workers' skills and the technical requirements of their jobs. For the compensation dimension, it is also difficult to fully disentangle three alternative explanations: strategic talent investment, efficiency wages, and rent-sharing. For the retention component of the compensation mechanism, data availability imposes a further constraint: Chinese listed firms disclose only period-end headcounts by staff category, not hires or separations, so we observe net headcount growth rather than turnover (Supplementary Material Table S6). The compensation-to-retention link is therefore presented as an interpretation throughout, and the corresponding statements in the theory section are phrased as associations rather than as a tested mechanism. Future research may incorporate more refined measurement indicators to address this limitation. Second, although this study employs firm fixed effects, firm-level clustering, lagged variables, leave-one-out mean instrumental variables, and multiple alternative measures, causal identification is not fully established. Skill configuration exhibits high persistence, making it difficult to fully satisfy the exclusion restriction for lagged instrumental variables. Moreover, the instrumental variable coefficients are insignificant within the firm fixed-effects framework. Accordingly, this study does not claim to resolve endogeneity concerns. Consistent with this, all coefficients reported from the two-way fixed-effects regressions—including the baseline estimate of 0.037—are interpreted as conditional associations rather than as causal effect sizes. Third, the baseline dependent variable is industry-level export technological sophistication. This indicator is adjusted by firm-level relative total factor productivity. Since both the dependent variable and core explanatory variable are constructed from raw labor-related data, there exists a risk of mechanical correlation. Although robustness tests using the TFP-excluded dependent variable confirm stable regression results, this study still interprets the findings as associations and refrains from strong causal claims. Future research should attempt to identify more appropriate dependent variables. Fourth, the AI-orientation channel is measured using four proxies that capture related but distinct facets of AI-related activity. The two disclosure-based measures partly reflect reporting rhetoric rather than operational deployment, and neither patents nor subsidiary establishment directly measures in-house AI use. The channel should therefore be read as a bundle of AI-related disclosure, patenting and strategic positioning, not as evidence that operational AI deployment causally improves ETS. Future research may

use direct measures of in-house AI adoption, such as AI staffing surveys or firm-level IT expenditure data.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/su182010540/s1. Table S1: VIF. Table S2: Correlation matrix of main variables. Table S3: Robustness to executive-pay contamination in the compensation measure. Table S4: Supplementary robustness tests. Table S5: Lead-lag tests of anticipatory recruitment (DV: $E_{si,it}$). Table S6: Configuration types and talent-related considerations. Table S7: Validation of the AI proxies.

Author Contributions: Conceptualization, Z.D. and X.C.; methodology, Z.D.; software, Z.D.; validation, X.C.; formal analysis, Z.D.; investigation, Z.D.; resources, Z.D. and X.C.; data curation, Z.D.; writing—original draft preparation, Z.D. and X.C.; writing—review and editing, Z.D. and X.C.; visualization, Z.D.; supervision, X.C.; project administration, X.C.; funding acquisition, Z.D. and X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Major Project of Humanities and Social Sciences Research of the Department of Education of Hebei Province, grant number [ZD202417]; Major Project of the Hebei Provincial Social Science Fund; grant number [HB25ZD003]; Hebei Provincial Central Government-Guided Local Science and Technology Development Fund Project, grant number [246Z7606G].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data reported in this study are available upon request from the corresponding author. The data used to support the findings of this study are included in the Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Bank. *Medium and High-Tech Exports (% of Manufactured Exports)*; Indicator TX.MNF.TECH.ZS.UN; World Bank: Washington, DC, USA, 2026. Available online: https://data.worldbank.org/indicator/TX.MNF.TECH.ZS.UN (accessed on 8 October 2026).
2. Humphrey, J.; Schmitz, H. How does insertion in global value chains affect upgrading in industrial clusters? *Reg. Stud.* **2002**, *36*, 1017–1027. [CrossRef]
3. Pietrobelli, C.; Rabellotti, R. Global value chains meet innovation systems: Are there learning opportunities for developing countries? *World. Dev.* **2011**, *39*, 1261–1269. [CrossRef]
4. Liu, Z.; Zhang, J. From integrating into global value chains to constructing national value chains: Strategic reflections on China's industrial upgrading. *Acad. Mon.* **2009**, *41*, 59–68. (In Chinese)
5. Nunn, N. Relationship-specificity, incomplete contracts, and the pattern of trade. *Q. J. Econ.* **2007**, *122*, 569–600. [CrossRef]
6. Huang, Q.; Yu, Y.; Zhang, S. Internet development and manufacturing productivity improvement: Internal mechanism and China's experience. *China Ind. Econ.* **2019**, *8*, 5–23. (In Chinese)
7. Baldwin, R. *The Great Convergence: Information Technology and the New Globalization*; Harvard University Press: Cambridge, MA, USA, 2016.
8. Du, Y.; Zhou, Y. Domestic market scale, siphon effect, and global value chain upgrading: Evidence from cross-country panel data. *J. Guangxi Norm. Univ. Philos. Soc. Sci. Ed.* **2025**, *61*, 103–120. (In Chinese)
9. Coe, N.M.; Yeung, H.W.-C. *Global Production Networks: Theorizing Economic Development in an Interconnected World*; Oxford University Press: Oxford, UK, 2015.
10. Kong, X. Human capital drives manufacturing industries toward the mid-to-high end of value chains: Mechanisms, influencing factors, and countermeasures. *Theory J.* **2025**, *5*, 132–141. (In Chinese)
11. McGuinness, S.; Pouliakas, K.; Redmond, P. Skills mismatch: Concepts, measurement and policy approaches. *J. Econ. Surv.* **2018**, *32*, 985–1015. [CrossRef]
12. Sicherman, N. "Overeducation" in the labor market. *J. Labor Econ.* **1991**, *9*, 101–122. [CrossRef] [PubMed]
13. Mahy, B.; Rycx, F.; Vermeylen, G. Educational mismatch and firm productivity: Do skills, technology and uncertainty matter? *Economist* **2015**, *163*, 233–262. [CrossRef]
14. Ji, W.; Lai, D. Human capital allocation and China's innovation performance. *Econ. Perspect.* **2018**, *11*, 19–31. (In Chinese)

15. Zheng, Y.; Dai, Y. The impact of skill mismatch on global value chain division position. *J. Zhongnan Univ. Econ. Law.* **2021**, *5*, 101–112. (In Chinese)
16. Dai, X.; Liu, M. How talent becomes dividend: Evidence from global value chain upgrading. *China Ind. Econ.* **2018**, *4*, 98–116. (In Chinese)
17. Atiq-ur-Rehman; Rauf, A.; Khan, G.Y. Human capital-growth nexus: The role of skill mismatch. *Intellect. Econ.* **2021**, *15*, 15–30. [CrossRef]
18. Verhaest, D.; Omey, E. Objective over-education and worker well-being: A shadow price approach. *J. Econ. Psychol.* **2009**, *30*, 469–481. [CrossRef]
19. Hsieh, C.T.; Klenow, P.J. Misallocation and manufacturing TFP in China and India. *Q. J. Econ.* **2009**, *124*, 1403–1448. [CrossRef]
20. Shao, M.; Bao, Q. The impact of export enterprise transformation on employment and wages in China: An empirical analysis based on propensity score matching estimation. *J. World Econ.* **2011**, *34*, 48–70. (In Chinese)
21. Ning, G.; Lin, Z. Information technology application, corporate organizational change, and shifts in labor skill demand. *Econ. Res. J.* **2014**, *49*, 79–92. (In Chinese)
22. McGowan, M.A.; Andrews, D. *Skills Mismatch, Productivity and Policies: Evidence from the Second Wave of PIAAC*; OECD Economics Department Working Papers No. 1403; OECD Publishing: Paris, France, 2017. [CrossRef]
23. Acemoglu, D.; Restrepo, P. Artificial intelligence, automation, and work. In *The Economics of Artificial Intelligence: An Agenda*; Agrawal, A., Gans, J., Goldfarb, A., Eds.; University of Chicago Press: Chicago, IL, USA, 2019; pp. 197–236. [CrossRef]
24. Teece, D.J.; Pisano, G.; Shuen, A. Dynamic capabilities and strategic management. *Strateg. Manag. J.* **1997**, *18*, 509–533. [CrossRef]
25. Teece, D.J. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strateg. Manag. J.* **2007**, *28*, 1319–1350. [CrossRef]
26. Eisenhardt, K.M.; Martin, J.A. Dynamic capabilities: What are they? *Strateg. Manag. J.* **2000**, *21*, 1105–1121. [CrossRef]
27. Helfat, C.E.; Peteraf, M.A. Managerial cognitive capabilities and the microfoundations of dynamic capabilities. *Strateg. Manag. J.* **2015**, *36*, 831–850. [CrossRef]
28. Barney, J. Firm resources and sustained competitive advantage. *J. Manag.* **1991**, *17*, 99–120. [CrossRef]
29. Peteraf, M.A. The cornerstones of competitive advantage: A resource-based view. *Strateg. Manag. J.* **1993**, *14*, 179–191. [CrossRef]
30. Nelson, R.R.; Winter, S.G. *An Evolutionary Theory of Economic Change*; Harvard University Press: Cambridge, MA, USA, 1982. [CrossRef]
31. Cohen, W.M.; Levinthal, D.A. Absorptive capacity: A new perspective on learning and innovation. *Admin. Sci. Q.* **1990**, *35*, 128–152. [CrossRef]
32. March, J.G. Exploration and exploitation in organizational learning. *Organ. Sci.* **1991**, *2*, 71–87. [CrossRef]
33. Hausmann, R.; Hwang, J.; Rodrik, D. What you export matters. *J. Econ. Growth* **2007**, *12*, 1–25. [CrossRef]
34. Sheng, B.; Mao, Q. Does import trade liberalization affect Chinese manufacturing export technological sophistication? *J. World Econ.* **2017**, *40*, 52–75. [CrossRef]
35. Wang, J.; Huang, Q. The Impact of Digital Transformation on the Export Technology Complexity of Manufacturing Enterprises. *Sustainability* **2025**, *17*, 2596. [CrossRef]
36. Gao, X.; Yuan, K. Cleaner production environmental regulation and corporate export technological sophistication: Micro evidence and influence mechanism. *J. Int. Trade.* **2020**, *2*, 93–109. (In Chinese)
37. Jiang, T. Mediating effect and moderating effect in empirical causal inference research. *China Ind. Econ.* **2022**, *5*, 100–120. (In Chinese)
38. Aranitou, V. *The Economic Impact of Small and Medium-Sized Enterprises: Analytical Approaches to Growth and Innovation Challenges Amid Crises in Europe*; Palgrave Macmillan: Cham, Switzerland, 2024.
39. Kindström, D.; Carlborg, P.; Nord, T. Challenges for Growing SMEs: A Managerial Perspective. *J. Small Bus. Manag.* **2024**, *62*, 700–723. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.