Taylor & Francis
Taylor & Francis Group

International Journal of Production Research

ISSN: 0020-7543 (Print) 1366-588X (Online) Journal homepage: www.tandfonline.com/journals/tprs20

Estimating the task content of work: workforce design for AI-driven human-robot collaboration in intralogistics

Pierre Bouquet, Nicolò Piergiovanni Bagnoli & Yossi Sheffi

To cite this article: Pierre Bouquet, Nicolò Piergiovanni Bagnoli & Yossi Sheffi (22 Mar 2026): Estimating the task content of work: workforce design for AI-driven human-robot collaboration in intralogistics, International Journal of Production Research, DOI: 10.1080/00207543.2026.2643477

To link to this article: https://doi.org/10.1080/00207543.2026.2643477

INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH
https://doi.org/10.1080/00207543.2026.2643477

Estimating the task content of work: workforce design for AI-driven human-robot collaboration in intralogistics

Pierre Bouquet, Nicolò Piergiovanni Bagnoli and Yossi Sheffi

Center for Transportation & Logistics, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA

ABSTRACT

This paper addresses the challenge of strategic workforce planning for AI-driven human-robot collaboration (AI-HRC) in intralogistics. We ask two questions: how can task-level full-time equivalent (FTE) estimates be constructed from existing labour statistics, and how can these estimates, combined with AI exposure metrics, inform strategic AI-HRC design and workforce planning? Drawing on U.S. Bureau of Labor Statistics employment data, O∗NET occupational profiles, and task-level AI exposure scores, we develop a stochastic task-time framework that decomposes occupations into tasks and models task frequencies as probability vectors on the simplex. A covariance-completion procedure reconstructs task covariance matrices consistent with survey standard errors, enabling the translation of occupational data into task-level and detailed work activity (DWA)-level FTE estimates with uncertainty bounds. Applying the framework to the U.S. intralogistics workforce, we find that approximately 370,000 FTEs (about 17% of workers) are concentrated in the top 15% most AI-exposed DWAs. These results provide task-specific insight into AI-driven automation and support scenario-based workforce planning by linking alternative AI-HRC adoption paths to task-level FTE impacts, uncertainty bands, and upskilling priorities, thereby offering an analytical foundation for resilient, human-centered AI-HRC systems.

ARTICLE HISTORY
Received 2 December 2025
Accepted 1 March 2026

KEYWORDS
Artificial Intelligence; future of work; workforce design; task automation; human robot collaboration; intralogistics

1. Introduction

Artificial intelligence (AI) is reshaping work organisation, producing heterogeneous effects across tasks, occupations, and industries. Recent audits indicate that approximately 80 percent of the U.S. workforce has at least 10 percent of tasks exposed to LLM capabilities or LLM-powered software, while about 19 percent has 50 percent or more of tasks exposed (Eloundou et al. 2024). Beyond exposure studies, empirical evidence from thousands of real interactions demonstrates high applicability for information-gathering, writing, and advising tasks, particularly where communication and coordination are essential (Tomlinson et al. 2025). Task-based economic research finds that digital technologies often replace routine tasks but augment non-routine problem solving (Autor and Dorn 2013; Autor, Levy, and Murnane 2003). These findings underline the importance of task-level analysis of AI-driven automation. To further understand the impact of AI, it is necessary to embed it within physical task systems, integrating sensing, actuation, and human judgment to enable coordination between robots and frontline workers. This integration is especially relevant in the operational context of logistics.

The logistics sector is undergoing a significant transformation as AI-driven technologies reshape transportation, warehousing, and distribution. Performance in logistics depends on synchronous, high-volume task execution and must adapt to volatility from demand peaks, SKU proliferation, and service-level pressures. Warehousing has shifted from picker-to-parts systems to robotised fulfilment centers, resulting in higher service quality and reduced lead times (Boysen and De Koster 2025). Industry 4.0 technologies and digital twins reduce information lags and enable tighter synchronisation between production and warehousing (De Koster et al. 2025; Li et al. 2023). Digital twins are best understood as cyber-physical systems that maintain real-time, data-driven representations of logistics processes through continuous streams of IoT and operational data (Park, Son,

CONTACT Pierre Bouquet pibou149@mit.edu Center for Transportation & Logistics, Massachusetts Institute of Technology (MIT), 1 Amherst Street, Cambridge, MA, USA

Supplemental data for this article can be accessed online at http://dx.doi.org/10.1080/00207543.2026.2643477.

and Noh 2021). By 2028, approximately 80 percent of warehouses and distribution centers are projected to use some form of robotics or automation, highlighting the rapid pace and scale of change (Ma and Saénz 2025). Autonomous mobile robots (AMRs) and robotic mobile fulfilment systems (RMFS) are seeing growing adoption in intralogistics, as more facilities integrate them into their planning and control processes (Fragapane et al. 2021). However, the sector continues to rely heavily on human workers, particularly in picking, sorting, and exception handling, where variability, damage, and unique items constrain the potential for full automation (U.S. Bureau of Labor Statistics 2024).

These technological shifts have set the stage for AI-driven human-robot collaboration (AI-HRC) within logistics facility operations. This sector – commonly referred to as intralogistics – concerns the management of material and information flows in warehouses, distribution centers, and manufacturing plants. Cobots, AMRs, and perception systems bring speed and precision, while humans provide dexterity, contextual understanding, and exception management (Pasparakis, Vries, and Koster 2023). When well-designed, such collaboration enhances job satisfaction and self-efficacy; when poorly paced or interfaced, it can erode trust and underutilise both humans and robots (Pasparakis, Vries, and Koster 2023). Despite rapid automation, humans remain the backbone of system resilience. In this evolving context, AI-enabled collaboration does not simply replace labour but reshapes it.

To address the evolving nature of work in intralogistics and the challenges posed by AI-HRC and resilience requirements, this study proposes a stochastic task-time framework for strategic workforce planning. The framework generates estimates of the total number of FTEs performing each task, including confidence levels. These estimates are then integrated with AI automation exposure studies to inform AI-HRC workforce planning. Methodologically, this approach extends prior work by Cai et al. (2025) by defining the intralogistics industry as a set of job-task pairs and incorporating uncertainty bounds into functionals of the frequency vector developed by Martin and Monahan (2022).

The results are estimated ranges of the number of FTEs assigned for each task within the U.S. intralogistics workforce. Although the primary focus is on the U.S. intralogistics sector, the methodology is broadly applicable and can be replicated for any workforce. When combined with AI exposure studies, this work can support strategic workforce planning by identifying the human capital required for AI-HRC tasks. To enhance practical relevance, exposures are interpreted in relation to AMR/RMFS, IoT instrumentation, and digital-twin design patterns. We also provide a public repository with all the code and data pipelines, which can be adapted to any workforce representation.

The remainder of the paper is organised as follows. Section 2 reviews the literature and positions this work among related studies. Section 3 describes the data and methodologies developed for the stochastic task-time framework. Section 4 presents the results and the diagnostics of the methodology. Section 5 examines the implications for strategic workforce planning and AI-HRC design, and discusses the assumptions underlying this study. Section 6 summarises key insights and outlines directions for future research.

2. Literature review

We present a comprehensive literature review on AI-driven automation, focussing on intralogistics. The section is organised as follows: first, modelling technological impact with task-based workforce models; next, presenting empirical evidence on AI innovations; finally, reviewing current AI-driven technologies in intralogistics and AI-HRC.

2.1. Modelling the impact of technological innovation through task-based workforce models

The relationship between technological progress and the reorganisation of work has been studied for decades using task-based economics. Foundational work formalised how computers substitute for routine tasks and complement non-routine cognitive and social functions, giving rise to labor-market polarisation analyses (Autor and Dorn 2013; Autor, Levy, and Murnane 2003). Subsequent work extended this framework to automation and AI, showing that new technologies both displace and create task frontiers (Acemoglu and Restrepo 2019; Autor et al. 2024; Autor 2015). This decomposition of occupations into constituent tasks, and the evaluation of their susceptibility to technological change, has become standard in labour economics and management studies (Brynjolfsson and Mitchell 2017; Eloundou et al. 2024; Felten, Raj, and Seamans 2018; Frey and Osborne 2017).

2.2. Empirical evidence on the effects of AI innovations

A parallel body of empirical literature quantifies the effects of AI-driven innovations on productivity and work quality. Randomised and quasi-experimental studies across multiple domains demonstrate substantial but heterogeneous gains that depend on task structure and worker experience. In software engineering, controlled trials of code-assistance tools show measurable reductions in completion time and error rates (Peng et al. 2023). In text-based knowledge work, LLM-based writing assistants boost speed and quality while narrowing performance dispersion among workers (Noy and Zhang 2023). Field studies in customer support and consulting suggest that generative AI improves both throughput and consistency, particularly among less experienced employees (Brynjolfsson, Li, and Raymond 2025; Dell'Acqua et al. 2023). Collectively, these studies indicate that AI-driven technologies are primarily redistributive rather than purely additive. They elevate lower performers, reduce productivity variance, and shift expertise toward supervision and integration with existing workflows, rather than completely redefining them. These behavioural mechanisms influence how AI exposure translates into performance gains and establish a baseline for understanding the augmentation enabled by this technology.

2.3. AI-driven technologies in intralogistics and HRC

While most AI-work studies focus on digital services, intralogistics offers a crucial testing ground where information processing and physical execution converge. Across successive automation waves, warehouse systems have evolved from human-supervised information platforms to nearly autonomous decision-making layers, computerising functions such as navigation, scheduling, routing, and replenishment (Boysen and De Koster 2025). AMRs now coordinate movements once assigned manually, RMFS manage retrieval tasks formerly handled by forklift or pallet operators, and IoT-based predictive maintenance automates monitoring and fault prediction previously performed by technicians (Li et al. 2023). The digital and physical layers have effectively merged – each pick, scan, and movement generates data that feeds learning algorithms for continuous optimisation. Industry 4.0 instrumentation and digital-twin synchronisation reduce information latency and tightly couple production with intralogistics in near real time, supporting scenario analysis, capacity planning, and resilience testing under disruptions (De Koster et al. 2025; Li et al. 2023; Liu, Pan, and Ballot 2024).

Within this technological transition, AI-HRC has become the organising principle of next-generation intralogistics systems. Research based on current technology increasingly supports the superiority of 'humans-in-the-loop' configurations over full automation in high-variability contexts, as anomalies and changeovers remain difficult to codify (De Koster and Roy 2024; De Koster et al. 2025). In this context, robots perform high-speed, repetitive, and physically demanding tasks, while humans contribute dexterity, contextual awareness, and adaptive decision-making (Pasparakis, Vries, and Koster 2023; Perotti et al. 2025). As tasks are divided in AI-HRC, even minor adjustments in automation design can influence workforce composition (Autor and Thompson 2025). Traditional workforce metrics at the occupational level obscure the inherent variability and uncertainty present in mixed human-AI task systems (Nourmohammadi et al. 2025).

3. Methodology

This section introduces the workforce modelling framework, followed by a description of the data sources and preprocessing steps. We outline the statistical approach for estimating task times and associated uncertainties. Finally, these estimates are translated into FTEs and aggregated to inform AI-HRC strategic workforce planning.

3.1. Overview and framework

The framework developed in this study provides a unified set of processes for translating occupational-level data into task-level estimates for AI-HRC workforce planning. The framework consists of five structured stages, as illustrated in Figure 1.

The model integrates three key data sources: (i) intralogistics job employment data; (ii) detailed task-job relationships; and (iii) metrics of task exposure to AI. The model translates occupational data into task-level FTE ranges. Diagnostic tests validate the model's internal consistency. The outputs are interpreted along three dimensions: aggregation of FTEs at the task level, integration of AI exposure scores to identify tasks most suitable for AI-HRC, and strategic workforce planning. The methodological assumptions underlying this framework are stated throughout Section 3 and summarised in Appendix A. A detailed discussion of limitations, scope conditions, and sources of uncertainty is provided in Appendix B.

3.2. Data sources

This section introduces the study data, specifying their sources, structure, and preprocessing steps.

3.2.1. BLS employment and wage data

We obtained employment data for the Transportation and Warehousing sector from the Bureau of Labor Statistics (BLS) Occupational Employment and Wage Statistics (OEWS) program. The OEWS is a semiannual survey

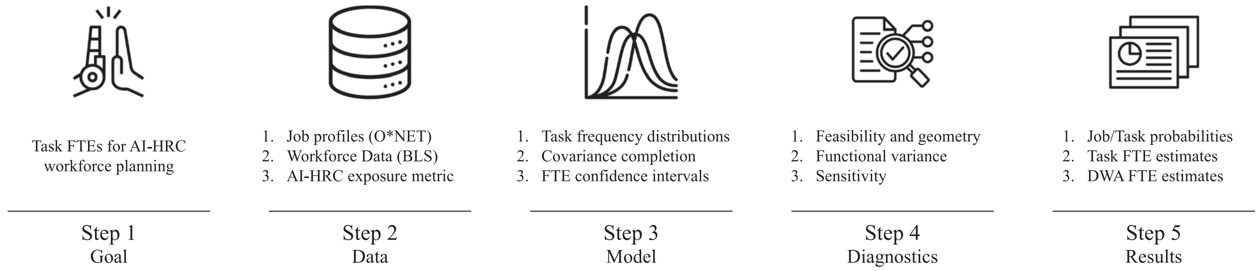

Figure 1. A modelling framework for estimating the number of workers per task within the intralogistics sector.

covering approximately 830 occupations based on a sample of approximately 1.1 million establishments representing 55 percent of total U.S. employment (U.S. Bureau of Labor Statistics 2024). The OEWS reports employment counts by occupation along with standard errors derived from the survey design. This paper employs the May 2024 OEWS release.

3.2.2. O∗NET occupational data

The O∗NET serves as the leading repository of standardised occupational data in the United States and a foundation for workforce planning strategies. Launched in 1998, it is maintained by the U.S. Department of Labor's Employment and Training Administration. The O∗NET Content Model organises occupational information hierarchically (National Center for O∗NET Development 2024). The database includes about 55,000 job titles which are consolidated into approximately 900 detailed occupational profiles. Each profile breaks an occupation into work elements organised hierarchically across several precision levels. At the finest granularity, over 18,000 unique task statements describe specific activities workers perform. They are grouped into about 2000 Detailed Work Activities (DWAs). The structure of the data, survey protocols, and extracts of the data tables used are provided in Appendix C. This study utilises O∗NET version 28.3, released in May 2024.

3.3. AI exposure metrics by task

We adopt the AI exposure metrics proposed by Bouquet, Sheffi, and Kaboli (2026). This methodology embeds DWA descriptions using the SBERT model and compares them against a large corpus of news articles referencing automation, robotics, and artificial intelligence. While cosine similarity determines the semantic relevance of an article to a task, the methodology employs a 'factual relevance' filter derived from the RoBERTa sentiment model to distinguish actual technological capability from media speculation. The metric explicitly prioritises neutral sentiment: in the context of technological news, neutrality serves as a proxy for objective reporting on actual deployments (e.g. 'The warehouse implemented AMRs'), whereas polarised sentiment typically reflects editorialising (e.g. 'Robots will save/destroy the workforce'). By isolating the signal of technical feasibility from the noise of media speculation, the model calculates a cosine-weighted average of these objectivity-filtered scores over a two-year rolling window. This approach ensures comparability across DWAs and robustness to short-term fluctuations in news coverage. The result is a continuous score between 0 and 1 for each DWA, indicating the potential exposure to AI-driven automation.

3.4. Data preparation

The process begins with the O∗NET task ratings file. A sequence of cleaning steps is then applied to prepare the data for modelling and to ensure that all valid and available information is used.

(1) **Preliminary cleaning** involved dropping rows flagged by O∗NET as 'Recommend Suppress = Y' and removing records with missing data values.
(2) **Category completeness** was enforced by requiring complete data for each task in the dataset. Tasks with missing data were excluded from further analysis.

3.5. Statistical modelling

This section introduces a framework for estimating the number of FTEs performing tasks within an organisation, sector, or economy under uncertainty.

3.5.1. Task frequency distributions

Data on task execution frequency is derived from the O∗NET data collection program, which administers structured surveys to job incumbents using a two-stage response protocol. First, respondents indicate whether a specific task is relevant to their occupation using a binary filter. If the task is deemed relevant, respondents then report the frequency of performance using a seven-point ordinal scale, ranging from 'Yearly or less' to 'Hourly or more.' O∗NET aggregates these responses to report the percentage of workers selecting each category, along with design-based standard errors. Further details on the data structure and survey protocols are provided in Appendix C.2.

Let i index occupations and k tasks. Denote by $\widehat{RT}_{i,k}$ the design-based percentage of incumbents who mark task k as relevant in occupation i, with standard error $\mathrm{SE}(\widehat{RT}_{i,k})$. We convert to probabilities via

$$\hat{R}_{i,k} = \frac{\widehat{RT}_{i,k}}{100}, \quad \mathrm{Var}(\hat{R}_{i,k}) = \left(\frac{\mathrm{SE}(\widehat{RT}_{i,k})}{100}\right)^2 \tag{1}$$

Similar to Martin and Monahan (2022), we introduce an explicit unconditional 'never' category (indexed $r = 0$) as the complement of relevance,

$$\hat{p}_{i,k,0} = 1 - \hat{R}_{i,k}, \quad \mathrm{Var}(\hat{p}_{i,k,0}) = \mathrm{Var}(\hat{R}_{i,k}). \tag{2}$$

Conditioned on relevance, O∗NET reports the design-based estimator of the percentage of workers choosing each frequency category $r = 1, \ldots, 7$, denoted $\widehat{FT}_{i,k,r}$ with standard errors $\mathrm{SE}(\widehat{FT}_{i,k,r})$. We first rescale them:

$$\hat{F}_{i,k,r} = \frac{\widehat{FT}_{i,k,r}}{100}, \quad \mathrm{Var}(\hat{F}_{i,k,r}) = \left(\frac{\mathrm{SE}(\widehat{FT}_{i,k,r})}{100}\right)^2, \quad \sum_{r=1}^{7} \hat{F}_{i,k,r} = 1. \tag{3}$$

Combining relevance with conditional frequencies yields the unconditional category probabilities for a randomly selected worker:

$$\hat{p}_{i,k,r} = \begin{cases} 1 - \hat{R}_{i,k}, & r = 0, \\ \hat{R}_{i,k}\,\hat{F}_{i,k,r}, & r = 1, \ldots, 7, \end{cases} \quad \sum_{r=0}^{7} \hat{p}_{i,k,r} = 1. \tag{4}$$

Assuming the design-based estimators $\hat{R}_{i,k}$ and $\hat{F}_{i,k,r}$ are statistically independent (elicited in distinct survey steps), we compute the exact variance of $\hat{p}_{i,k,r}$ using the product variance formula from Goodman (1960). For $r \geq 1$, this yields:

$$\mathrm{Var}(\hat{p}_{i,k,r}) = \hat{F}^2_{i,k,r}\,\mathrm{Var}(\hat{R}_{i,k}) + \hat{R}^2_{i,k}\,\mathrm{Var}(\hat{F}_{i,k,r}) + \mathrm{Var}(\hat{R}_{i,k})\,\mathrm{Var}(\hat{F}_{i,k,r}). \tag{5}$$

If any $\hat{p}_{i,k,r}$ evaluates to zero after construction (or its standard error is numerically zero), we add $\epsilon = 10^{-6}$ for computational stability to the affected entries and renormalise $\hat{\mathbf{p}}_{i,k}$ to maintain the simplex constraint.

To translate the unconditional category probabilities into expected annual occurrence counts, we use an annualisation vector based on the work by Tomlinson et al. (2025).

$$f\mathbf{w} = [0,\ 1,\ 4,\ 24,\ 104,\ 260,\ 780,\ 2080]^{\top} \tag{6}$$

where each entry corresponds to the O∗NET scale [Never, Yearly, More-than-yearly, More-than-monthly,…, Hourly-or-more]. These values implement conservative lower-end midpoints of the 'more than' intervals (e.g. 4 ≈ 2× per half-year; 24 ≈ 2× per month) and standard full-time assumptions (hourly ≈ 2080 work-hours/year). Finally, let $\hat{\mu}_{i,k}$ be the plug-in estimator of the expected annual number of occurrences for task k in occupation i:

$$\hat{\mu}_{i,k} = \mathbf{w}^{\top}\hat{\mathbf{p}}_{i,k} = \sum_{r=0}^{7} w_r\,\hat{p}_{i,k,r}. \tag{7}$$

While the O∗NET survey uses a complex design, we adopt a model-based approach for uncertainty analysis. We treat the empirically derived probability vector $\hat{\mathbf{p}}_{i,k}$ as the latent population distribution and assume each worker's categorical response is an independent and identically distributed draw from $\hat{\mathbf{p}}_{i,k}$.

3.5.2. Covariance completion

Determining the uncertainty of the annualised task frequency, $\hat{\mu}_{i,k} = \mathbf{w}^{\top}\hat{\mathbf{p}}_{i,k}$, requires the full covariance matrix $\widehat{\Sigma}_{i,k} \in \mathbb{S}^8$, the space of symmetric 8 × 8 matrices. However, O∗NET survey data provides only the marginal variances, leaving the cross-category correlations undefined. To recover a valid covariance structure, we formulate a convex completion problem that enforces consistency with both the observed survey design and the geometry of the probability simplex (Grussler, Rantzer, and Giselsson 2018).

Formally, any covariance matrix of a probability vector on the unit simplex must satisfy the tangent space constraint $\Sigma\mathbf{1} = \mathbf{0}$ (Aitchison 1982). We estimate $\widehat{\Sigma}_{i,k}$ by selecting the positive semidefinite (PSD) matrix that satisfies this geometric constraint and matches the reported marginals, while minimising the squared Frobenius norm $\|\Sigma\|_F$. This objective function acts as a parsimonious regularizer (Boyd and Vandenberghe 2004; Horn and Johnson 2012): it recovers the minimum-energy covariance structure required to satisfy the constraints, thereby avoiding the imposition of spurious correlations. This approach is robust to survey noise, as Frobenius norm minimisation prevents the amplification of sampling errors into the off-diagonal terms (Yue et al. 2024).

All semidefinite programs are solved with MOSEK ApS (2025); solver settings are reported in Appendix D.

Let $\mathbf{v}_{i,k} = [v_{i,k,0}, \ldots, v_{i,k,7}]^\top$, where $v_{i,k,r} = \mathrm{Var}(\hat{p}_{i,k,r})$ denotes the reported marginal variance for frequency category r of task k in occupation i. We obtain the covariance matrix $\widehat{\Sigma}_{i,k}$ as:

$$\begin{aligned} \widehat{\Sigma}_{i,k} \in \arg\min_{\Sigma \in \mathbb{S}^8} \|\Sigma\|_F^2, & \\ \text{s.t.} \quad \Sigma \succeq 0, & \\ \mathrm{diag}(\Sigma) = \mathbf{v}_{i,k}, & \\ \Sigma\mathbf{1} = \mathbf{0}. & \end{aligned} \tag{8}$$

Here, $\|\cdot\|_F$ denotes the Frobenius norm, $\mathbb{S}^8$ the space of symmetric 8×8 matrices, and $\Sigma\mathbf{1} = \mathbf{0}$ enforces the unit-sum property of probabilities. The variance of the plug-in estimator $\hat{\mu}_{i,k} = \mathbf{w}^\top \hat{\mathbf{p}}_{i,k}$ follows from the law of propagation of variances (Bevington and Robinson 2003)

$$\mathrm{Var}(\hat{\mu}_{i,k}) = \mathbf{w}^\top \widehat{\Sigma}_{i,k} \mathbf{w}, \tag{9}$$

3.5.3. Model diagnostics

We assess the completed covariance matrices $\widehat{\Sigma}_{i,k}$ along three dimensions: feasibility, geometry on the simplex, and robustness.

3.5.3.1. Feasibility and Simplex Geometry.

We first verify the solver convergence for each solution of problem (8), since there is a potential for ill-conditioning and numerical instability in PSD problems (Wolkowicz, Saigal, and Vandenberghe 2000). Hence we check the feasibility of each converged solution, based on two principal criteria.

The simplex tangent constraint is imposed exactly: each estimated covariance $\widehat{\Sigma}_{i,k}$ must satisfy $\widehat{\Sigma}_{i,k}\mathbf{1} = \mathbf{0}$, ensuring compatibility with compositional geometry (Aitchison 1982). Positive semidefiniteness is verified numerically by requiring $\lambda_{\min}(\widehat{\Sigma}_{i,k}) \geq -10^{-8}$, consistent with standard solver tolerances and reflecting typical floating-point precision (Wolkowicz, Saigal, and Vandenberghe 2000).

Solutions are deemed feasible only if both criteria are met; otherwise, the task is rejected from the model. For additional rigour and reference, we report supplementary diagnostics, including diagonal agreement, total uncertainty, and effective rank, in Appendix E.1.

3.5.3.2. Functional Variance and Feasible Bounds.

To link our diagnostics to inference on annual occurrences, we compute the variance of the plug-in estimator, $\hat{\mu}_{i,k}$, as $\mathbf{w}^\top \widehat{\Sigma}_{i,k} \mathbf{w}$. We benchmark this value against the feasible variance interval $[\sigma^2_{\min}, \sigma^2_{\max}]$, which is obtained by solving two auxiliary problems to minimise and maximise $\mathbf{w}^\top \Sigma \mathbf{w}$, subject to $\Sigma \succeq 0$, $\mathrm{diag}(\Sigma) = \mathbf{v}_{i,k}$, and $\Sigma\mathbf{1} = 0$ (Boyd and Vandenberghe 2004). We report the normalised position

$$\frac{\mathbf{w}^\top \widehat{\Sigma}_{i,k} \mathbf{w} - \sigma^2_{\min}}{\sigma^2_{\max} - \sigma^2_{\min}} \in [0, 1], \tag{10}$$

which quantifies the conservativeness of the completion relative to the range of all feasible solutions under these constraints. To prevent a small number of extreme yet feasible completions from dominating uncertainty budgets, we exclude tasks whose normalised position lies outside the interval $[0.05, 0.95]$, following robust-trimming principles in Huber and Ronchetti (2009); see Appendix E.2 for details.

3.5.3.3. Sensitivity.

Finally, we assess robustness to sampling-error misspecification by perturbing the diagonal targets $\mathbf{v}_{i,k}$ by $\pm 5\%$ (Huber and Ronchetti 2009; Saltelli et al. 2004), recomputing the completion, and recording the absolute and relative change in $\mathbf{w}^\top \widehat{\Sigma}_{i,k} \mathbf{w}$. Small changes indicate that the inferred functional uncertainty is both numerically stable and substantively robust to moderate misspecification of the marginals that could come from O∗NET sampling errors (Yuan and Zhang 2013). For each task, we report the maximum and minimum changes observed, as well as the spread and symmetry of the variance estimates. Tasks exhibiting numerically unstable sensitivity, specifically, those with symmetric relative change exceeding 0.1 after a perturbation of 5% of the marginals are excluded from further analysis; see Appendix E.3 for details.

3.6. From tasks to full-time equivalents

The preceding sections established methods for estimating task occurrence rates $\hat{\mu}_{i,k}$ with full covariance structures. We now translate these annual occurrence estimates into workforce-level metrics. This process involves three stages. First, raw annual task frequencies are converted into time shares to evaluate the proportion of working time that occupation i allocates to task k. Second, time shares are scaled by employment counts to calculate FTE headcounts at the task levels. Third, tasks are aggregated into cross-occupational DWAs to enable standardised comparisons across occupations. Uncertainty is propagated at each stage to provide confidence intervals of workforce estimates.

3.6.1. Occupation-level time shares

Let $\hat{\mu}_{i,k}$ denote the expected annual number of occurrences of task k in occupation i, let $d_{i,k}$ be a fixed task duration (hours per occurrence), and let K_i be the set of task indices within occupation i. Define the occupation-level total number of task occurrences:

$$M_i = \sum_{j \in K_i} \hat{\mu}_{i,j}\, d_{i,j}. \tag{11}$$

The time share $\pi_{i,k}$ that occupation i allocates to task k is

$$\pi_{i,k} = \frac{\hat{\mu}_{i,k}\, d_{i,k}}{M_i}, \quad \text{where} \sum_{k \in K_i} \pi_{i,k} = 1. \tag{12}$$

$\pi_{i,k}$ represents the expected percentage of time someone doing job i will spend on task k. In addition, we know $\mathrm{Var}(\hat{\mu}_{i,k})$ and we assume that $\hat{\mu}_{i,k}$ is asymptotically normal (van der Vaart 2000). Under this assumption, we propagate uncertainty using the delta method, following standard variance approximations for ratio estimators in survey sampling (Särndal, Swensson, and Wretman 1992). The resulting approximation is:

$$\frac{\partial \pi_{i,k}}{\partial \hat{\mu}_{i,k}} = \frac{d_{i,k}}{M_i}\left(1 - \pi_{i,k}\right), \quad \frac{\partial \pi_{i,k}}{\partial \hat{\mu}_{i,j}} = -\frac{\pi_{i,k}\, d_{i,j}}{M_i} \quad (j \neq k), \tag{13}$$

and we get:

$$\mathrm{Var}(\pi_{i,k}) \approx \frac{1}{(M_i)^2}\left[d_{i,k}^2 \left(1 - \pi_{i,k}\right)^2 \mathrm{Var}(\hat{\mu}_{i,k}) + \pi_{i,k}^2 \sum_{\substack{j \in K_i \\ j \neq k}} d_{i,j}^2\, \mathrm{Var}(\hat{\mu}_{i,j}) \right]. \tag{14}$$

While the general expression holds, we follow the standard simplifying assumption that all tasks within an occupation have equal duration per occurrence and set $d_{i,j} = 1$ (Martin and Monahan 2022; Tomlinson et al. 2025). This assumption is required because O∗NET reports task frequencies but not the time spent per execution. Importantly, this does not remove task heterogeneity: differences in total task time remain captured through the annualisation vector w (6), which places greater weight on high-frequency categories (e.g. hourly) than on low-frequency ones (e.g. yearly) (Martin and Monahan 2022). Under this assumption, (11) reduces to $M_i = \sum_{j \in K_i} \hat{\mu}_{i,j}$ and, by independence, $\mathrm{Var}(M_i) = \sum_{j \in K_i} \mathrm{Var}(\hat{\mu}_{i,j})$. Equation (14) then simplifies to:

$$\mathrm{Var}(\pi_{i,k}) \approx \frac{1}{M_i^2}\left[\left(1 - 2\pi_{i,k}\right) \mathrm{Var}(\hat{\mu}_{i,k}) + \pi_{i,k}^2\, \mathrm{Var}(M_i)\right], \tag{15}$$

and we show how to evaluate the proportion of time that occupation i spends on task k, including a measure of uncertainty.

3.6.2. Workforce-level FTE estimation

Having established occupation–task time shares $\pi_{i,k}$ and their associated uncertainty, we scale these proportions by employment levels to obtain task-level FTE counts. Let N_i denote the number of workers in occupation i and $\mathrm{SE}(N_i)$ its sampling standard error. The estimated FTE devoted to task k within occupation i is

$$\widehat{\mathrm{FTE}}_{i,k} = \pi_{i,k}\, N_i, \tag{16}$$

which satisfies $\sum_{k=1}^{K_i} \widehat{\mathrm{FTE}}_{i,k} = N_i$ by construction. To quantify uncertainty, we treat $\pi_{i,k}$ and N_i as independent random variables and apply the exact variance formula for the product of independent random variables (Goodman 1960). Writing $\mathrm{Var}(N_i) = \mathrm{SE}(N_i)^2$, this yields

$$\mathrm{Var}\left(\widehat{\mathrm{FTE}}_{i,k}\right) = N_i^2\, \mathrm{Var}(\pi_{i,k}) + \pi_{i,k}^2\, \mathrm{Var}(N_i) + \mathrm{Var}(\pi_{i,k})\, \mathrm{Var}(N_i). \tag{17}$$

When employment counts are treated as fixed, $\mathrm{SE}(N_i) \approx 0$ and the terms involving $\mathrm{Var}(N_i)$ vanish. The resulting estimate of task-level FTE is reported with a $(1-\alpha)$ confidence interval:

$$\widehat{\mathrm{FTE}}_{i,k} \pm z_{1-\alpha/2} \sqrt{\mathrm{Var}\left(\widehat{\mathrm{FTE}}_{i,k}\right)}, \tag{18}$$

truncated to the interval $[0, \sum_i N_i]$ to enforce natural bounds.

3.6.3. Aggregating task FTEs

A key element of O∗NET is the tasks to DWA assignments, which link occupation-specific tasks to shared DWAs through a many-to-many relationship. This structure enables an inter-occupation workforce-level analysis of FTEs since DWAs are shared between occupations. However, aggregating task-level FTEs to DWA FTEs introduces potential double-counting. To address this, we assume a uniform distribution of FTEs across linked DWAs and implement double-counting safeguards.

Let $\mathcal{M}_{i,k}$ denote the set of DWAs m linked to task k in occupation i. For each $m \in \mathcal{M}_{i,k}$, the FTE contribution from task k in occupation i to DWA m is allocated as:

$$\widehat{\mathrm{FTE}}_{i,k,m} = \frac{\widehat{\mathrm{FTE}}_{i,k}}{|\mathcal{M}_{i,k}|}, \tag{19}$$

where $|\mathcal{M}_{i,k}| \geq 1$ (in the special case when $|\mathcal{M}_{i,k}| = 1$, the full $\widehat{\mathrm{FTE}}_{i,k}$ is assigned to the single DWA). The total FTE for a given DWA m is obtained by summing over all contributing task-occupation pairs:

$$\widehat{\mathrm{FTE}}_m = \sum_i \sum_{\substack{k \\ m \in \mathcal{M}_{i,k}}} \widehat{\mathrm{FTE}}_{i,k,m}. \tag{20}$$

The mappings $|\mathcal{M}_{i,k}|$ are fixed from the O∗NET dataset and thus introduce no additional uncertainty. Under the assumptions of independence across tasks within an occupation and across occupations (as in prior variance propagations), the variance of each contribution simplifies to:

$$\mathrm{Var}(\widehat{\mathrm{FTE}}_{i,k,m}) = \left(\frac{1}{|\mathcal{M}_{i,k}|}\right)^2 \mathrm{Var}(\widehat{\mathrm{FTE}}_{i,k}), \tag{21}$$

using the variance derived for $\widehat{\mathrm{FTE}}_{i,k}$ in previous sections. The total variance for $\widehat{\mathrm{FTE}}_m$ is then:

$$\mathrm{Var}(\widehat{\mathrm{FTE}}_m) = \sum_i \sum_{\substack{k \\ m \in \mathcal{M}_{i,k}}} \mathrm{Var}(\widehat{\mathrm{FTE}}_{i,k,m}), \tag{22}$$

with a corresponding $(1-\alpha)$ confidence interval given by:

$$\widehat{\mathrm{FTE}}_m \pm z_{1-\alpha/2}\sqrt{\mathrm{Var}(\widehat{\mathrm{FTE}}_m)}, \tag{23}$$

truncated to $[0, \sum_i N_i]$ to respect natural bounds. This aggregation transforms task-level estimates into DWA-level workforce FTEs. Since many DWAs are shared among workers in the same industry, this enables direct comparison of activities across occupations in the intralogistics sector. Provided with uncertainty measures, this enables workforce planning and AI-HRC design by reporting the estimated FTE count performing each activity.

4. Results

This section provides a structured overview of our findings. The first subsection presents results on the data cleaning procedure and descriptive statistics of the intralogistics sector. The subsequent subsection evaluates the performance and diagnostics of the covariance completion methodology. The final subsection combines DWA-level FTE estimates with AI exposure scores to characterise employment concentration across DWAs and quantify the share of intralogistics workers in highly exposed activities.

4.1. Results – data cleaning and descriptive statistics

The data cleaning methodology described in Section 3.4 is applied to the intralogistics sector. Occupations are initially filtered to keep only the 'Transportation and Warehousing' sector ones. Because this classification is broader than intralogistics, the analysis focuses on occupations within transportation and warehousing that involve internal material handling, coordination, and operational management within logistics facilities, excluding external transport and last-mile delivery. The selected occupations include both manual and managerial roles, such as *stockers*, *machine operators*, *logisticians*, and *supply-chain managers* (Appendix F).

Table 1. Intralogistics dataset attrition by step: occupations and tasks.

	Occupations		Tasks	
Step	Count	Cum. loss (%)	Count	Cum. loss (%)
Initial intralogistics filter	20	–	408	–
After Task Ratings cleaning	14	30.00	219	46.32
After BLS 2024 match	11	45.00	167	59.07
Final vs. initial	11	45.00	167	59.07

The initial filtering of the intralogistics subset yields 20 Standard Occupational Classification (SOC) and 408 unique task statements. The Task Ratings cleaning procedures produce 14 occupations and 219 tasks. Alignment with the 2024 BLS naming convention further reduces the subset to 11 occupations and 167 tasks mapped to 98 DWAs. Table 1 summarises the retention statistics.

Figure 2 illustrates the evolution of total employment in the sector from 2015 to 2024, based on data from the U.S. Bureau of Labor Statistics (2024). Employment increased steadily until 2021 and then plateaued. The spike in 2019 is attributable to an O∗NET occupational reclassification (see Appendix G). Since 2022, the intralogistics sector has consistently employed approximately 2.2 million workers in the United States, representing about 1.4 percent of the total American workforce.

One interesting insight is the breakdown of the U.S. intralogistics workforce by education level. Roles requiring a higher education degree grew rapidly, increasing from approximately 40,000 in 2015 to over 70,000 in 2024. In contrast, employment among workers with only a high school diploma grew from 1.34 million to about 2.01 million, reflecting slower relative growth rates of 5.8 percent and 4.8 percent compound annual growth rate, respectively. This percentage-point difference is primarily attributable to an accelerated growth in recent years. These patterns support a gradual but persistent shift toward higher educational attainment within the intralogistics sector (Figure 3).

4.2. Results – covariance completion diagnostics and robustness

This section presents diagnostics for the covariance matrices, focussing on feasibility, geometric conformity, and sensitivity. Results are provided for both the complete cleaned dataset ($N = 11{,}834$) and the subset of intralogistics DWAs ($N = 167$).

(1) The assessment of covariance feasibility involves verifying the numerical and geometric validity of the

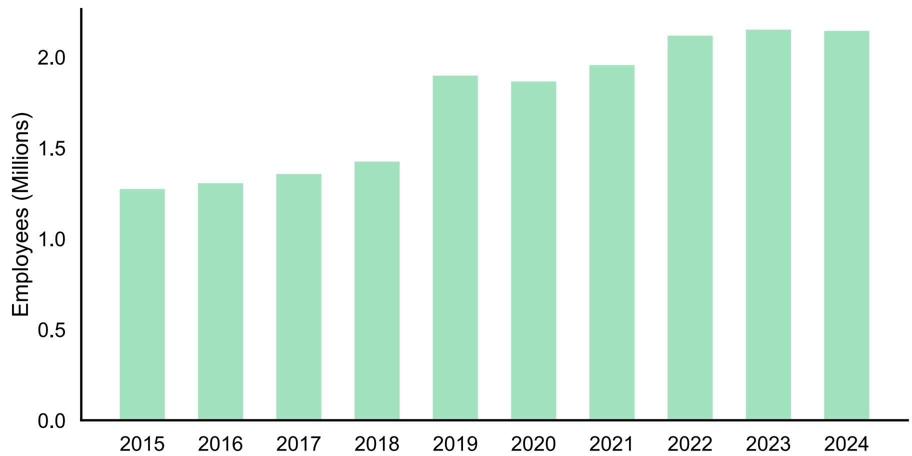

Figure 2. Total employment in the filtered intralogistics sector, USA (2015–2024). Source: U.S. Bureau of Labor Statistics (2024).

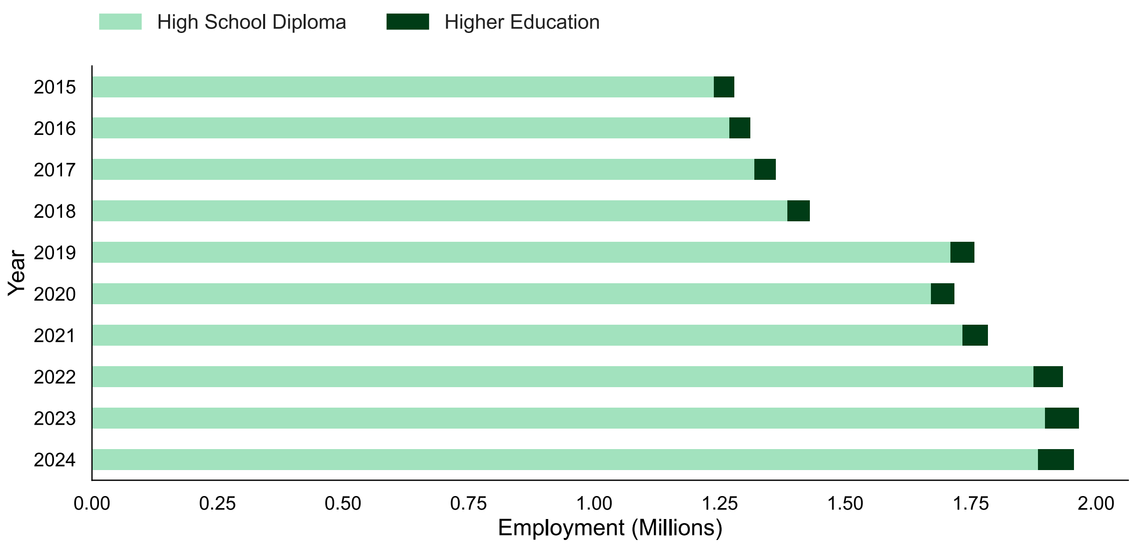

Figure 3. Filtered intralogistics employment by education level (2015-2024). Source: U.S. Bureau of Labor Statistics (2024) & National Center for O∗NET Development (2024).

completed covariance matrices. All task covariance matrices in the datasets are positive semi-definite. The maximum tangent residual is less than 10^{-10} indicating that the covariance matrices reside on the feasible simplex. Detailed results are provided in Appendix E.1.

(2) For each task, the estimated variance of the annual occurrence, derived from the completed covariance matrix, is compared to its feasible range. This range is determined through convex optimisation with identical marginal constraints, resulting in task-specific lower and upper variance bounds. Each task's realised variance is normalised to a value between zero and one within these bounds. The mean normalised position for the intralogistics sector is 0.39 with a standard deviation of 0.12. No mean normalised position outlier was present for the intralogistics sector's tasks. In contrast, 177 outliers (1.5 percent) were identified in the full dataset of 11,834 tasks. Detailed results are available in Appendix E.2.

(3) The sensitivity of the covariance matrix to minor changes in estimates is evaluated by perturbing the target marginal variances by plus or minus 5 percent and recalculating the full covariance completion for each task. The symmetrised relative change in variance is then measured. All intralogistics tasks exhibit stable and well-conditioned behaviour, as modest survey errors do not result in statistically significant changes in task variance. In comparison, the full dataset has 1 (0.01 percent) tasks failing directional checks. Detailed results are provided in Appendix E.3.

Diagnostics for the entire dataset identified 177 unique tasks (1.5 percent), where the covariance matrix

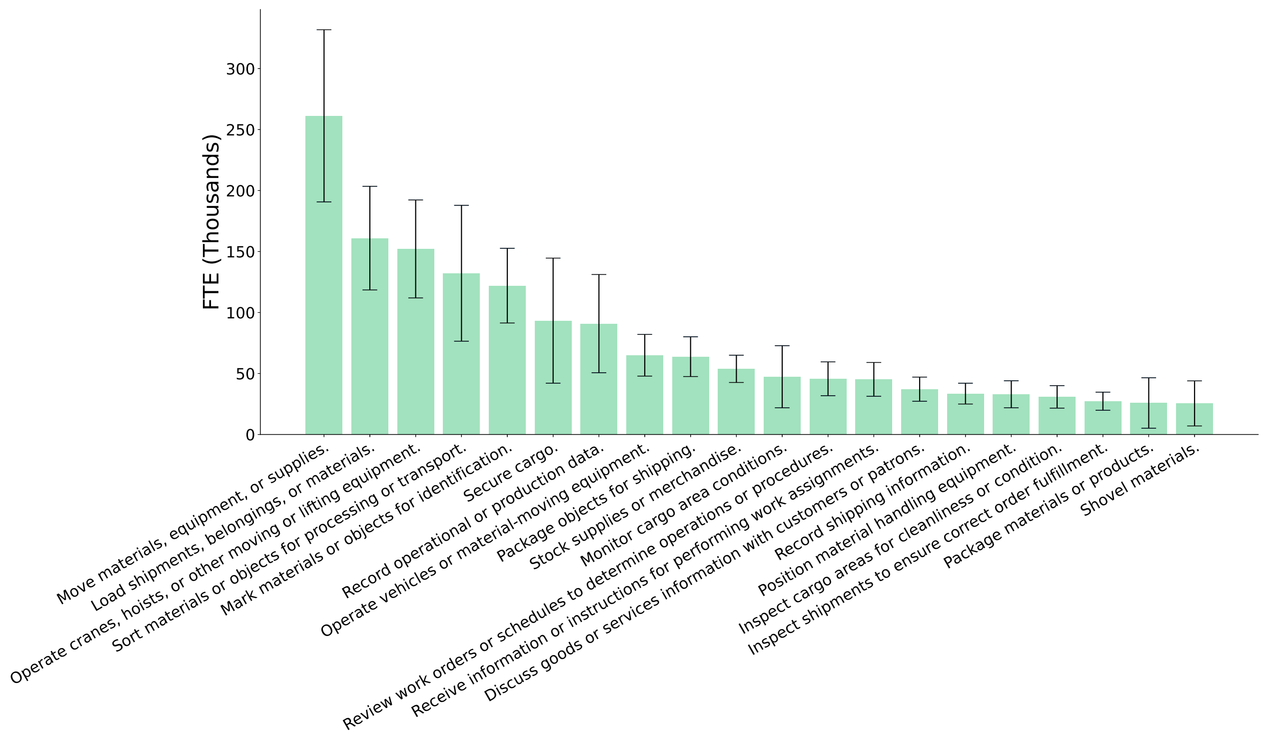

Figure 4. Estimated FTEs by DWA Title with 90 percent Confidence Intervals.

did not satisfy at least one criterion. The small proportion supports the use of covariance completion to reconstruct matrices from marginal standard errors for task frequency estimates. No covariance matrix was removed for intralogistics tasks, yielding 167 appropriately behaved task matrices. These diagnostics support the reliability of our method used to calculate the FTEs of the task for the intralogistics workforce.

4.3. Results – exposure to AI from intralogistics workers

The main findings are now discussed in detail. Figure 4 illustrates estimated employment concentration across the ten largest DWAs and their associated uncertainty in the intralogistics sector. The columns indicate mean FTE estimates in thousands, and the error bars represent 90 percent confidence intervals. For example, the DWA *'move materials, equipment, or supplies'* is estimated to employ between 200,000 and 340,000 FTEs. Notably, employment is concentrated in a limited number of DWAs, which correspond to the most common operational roles. For instance, the ten largest DWAs account for 54 percent of the total intralogistics workforce (in terms of FTEs). This concentration reflects the sector's DWA structure, where recurrent operations such as moving materials or supplies, monitoring cargo area conditions, and recording production data account for the largest FTE counts. The top ten DWAs by FTE count are listed in Appendix H.

Figure 5 associates each DWA to its AI exposure score. The x-axis represents exposure to AI-driven automation, and the y-axis indicates FTE for each DWA. Each point represents a DWA's exposure and its mean FTE estimate with a 90 percent confidence interval. The shaded area depicts the cumulative FTE distribution. The 85th percentile of the automation score identifies the most exposed DWAs which are highlighted as red crosses (Bouquet, Sheffi, and Kaboli 2026). This categorisation reveals that the equivalent of 16.4 percent, or approximately 366,000 FTEs of the intralogistics workforce may be affected by the implementation of AI-HRC. The list of DWAs, their FTEs and confidence intervals are depicted in Table 2.

Within the high-exposure DWAs in Table 2, three tasks stand out by FTE magnitude: (i) *Mark materials or objects for identification*, (ii) *Monitor cargo area conditions*, and (iii) *Review work orders or schedules to determine operations or procedures*. Although these activities differ in function, they share factors that make them particularly susceptible to partial automation; namely, repetitive visual or procedural inspection and the integration of digital work-order systems.

These results collectively indicate that while employment remains concentrated in a few dominant DWAs, the subset of tasks positioned at the higher end of the AI exposure distribution tends to include both operational and supervisory functions. This heterogeneity suggests that the next stage of technological transformation in

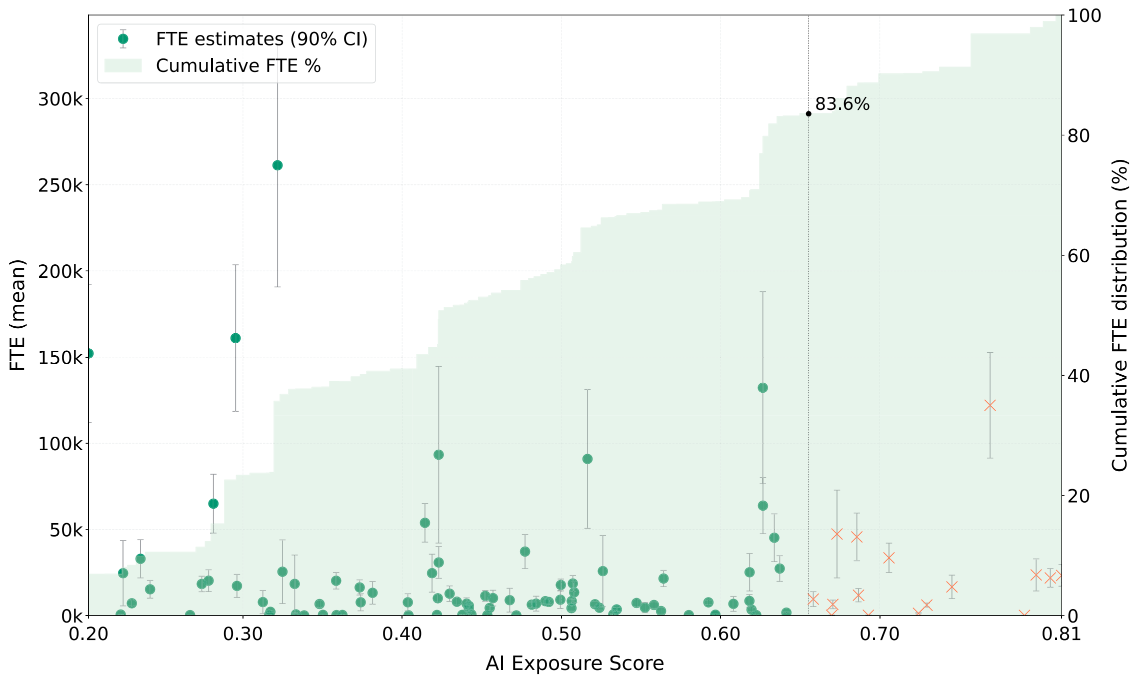

Figure 5. FTE Mean vs. AI Exposure with 90 percent Confidence Intervals and Cumulative Distribution.

Table 2. DWAs above 85th percentile AI-HRC exposure, ordered by FTE.

DWA Title	FTE Estimate	SE
Mark materials or objects for identification	122,047	18,623
Monitor cargo area conditions	47,346	15,487
Review work orders or schedules to determine operations or procedures	45,661	8424
Record shipping information	33,522	5191
Attach identification information to products, items, or containers	23,610	5659
Receive shipments	23,314	3798
Read work orders to determine material or setup requirements	21,862	3283
Prepare documentation for contracts, transactions, or regulatory compliance	16,669	4149
Verify information or specifications	11,842	2349
Analyze shipping information to make routing decisions	9579	2611
Store records or related materials	6502	1531
Measure product or material dimensions	6140	723
Record production information	1352	202
Prepare products for testing	198	71
Collect samples of materials or products for testing	198	71
Prepare informational or reference materials	128	51
Total	**366,958**	**29,087**

intralogistics may not involve the replacement of full occupations, but rather the augmentation of specific tasks. Consequently, precise strategic workforce planning is necessitated at the DWA level to effectively manage this transition.

5. Discussion

The uncertainty-aware approach developed in this study provides an understanding of the range of labour distribution across tasks. The following discussion connects this new methodology to three operational domains: strategic workforce planning, AI-HRC design, and resilience under uncertainty.

5.1. Functional areas and task-Level reallocation for AI-HRC

Using our results, DWAs can be grouped into three domains ripe for AI-HRC: administrative and coordination tasks, material handling, and monitoring.

In administrative and coordination functions, enterprise resource planning systems already streamline documentation, data transfer, and scheduling. Generative AI assistants now extend these systems by automatically summarizing orders and drafting reports, enhancing efficiency in what were once manual clerical and supervisory workflows (Pennathur et al. 2024).

In material handling and storage, automation technologies such as AMRs, RMFS, and robotic process automation embody the physical side of AI integration. Computer vision, path planning, and digital twin models enable robots to navigate, transport goods, and coordinate warehouse operations dynamically (Boysen and De Koster 2025; Fragapane et al. 2021; Li et al. 2023). In these contexts, automation can replace human labour. Tasks once done manually by workers can be handled by machines, while oversight, exception management, and safety remain with human operators. These supervisory roles entail distinct responsibilities and may be filled by personnel other than those who performed manual labour.

Finally, in monitoring and inspection, perception-driven AI and sensor networks are automating quality control and environmental supervision. Vision-based defect detection, predictive maintenance, and anomaly detection systems continuously collect and interpret data, shifting the human role from routine checking toward diagnostic and decision-making functions (Li et al. 2023; Liu et al. 2022).

Overall, these three domains illustrate how AI-HRC transforms intralogistics into a hybrid ecosystem where automation enhances precision and productivity while humans are responsible for oversight and exception handling. In this context, targeted, task-specific interventions allow for more precise allocation of training resources and can improve workforce placement outcomes as AI-HRC displaces human labour. A detailed section with the selected DWAs and ready-to-deploy AI-HRC technologies is presented in Appendix I.

5.2. Strategic workforce planning and real-time decision making

Traditional workforce planning typically employs a top-down approach, with managers estimating headcount based on projected revenue and demand (Cantrell and Klosk 2025). Instead of viewing workforce planning as a periodic activity, the proposed framework supports continuous, real-time monitoring of task allocations between human workers and robots. The task-level model can be updated as employees are hired, reassigned, or their employment ends. When integrated with data on AI-driven robotic innovations, this approach transforms AI-HRC and workforce planning from a static process into a dynamic organisational intelligence system (Teece, Pisano, and Shuen 1992; Xu, Liu, and Montreuil 2024).

This dynamic approach facilitates scenario-based staffing strategies, allowing organisations to simulate the effects of AI-HRC and directly connect automation to workforce impact. These simulations provide insights into operational risk and enable the evaluation of workforce strategies for AI-HRC design. For instance, the largest intralogistics task, *'move materials, equipment, or supplies'*, has a 90 percent confidence interval of approximately 200,000 and 340,000 FTEs, supporting multiple staffing scenarios grounded in empirical data. A conservative AI-HRC scenario with a 90/10 human-to-robot ratio would affect between 20,000 and 34,000 FTEs, while an aggressive 40/60 split would reduce the workforce by at least 120,000 FTEs. Even limited automation can significantly affect workforce size and wage bill costs.

As human labour may remain less expensive than AI-HRC, this research can help identify the inflection point at which automation becomes economically viable for the organisation and the expected amortization period from AI-HRC investment. Scenario analysis provides a proactive, data-driven basis for workforce planning and clarifies accountability for workforce and investment decisions.

5.3. Employee retention and supply-chain resilience

The intralogistics sector faces substantial workforce volatility, with an annual turnover rate of 49 percent, nearly three times the U.S. average and among the highest across all sectors (U.S. Bureau of Labor Statistics 2024). This labour instability is a major supply chain vulnerability, as disruptions to human capital can cause operational delays and performance losses (Nagurney and Ermagun 2022). AI-HRC helps mitigate these risks by improving working conditions in hazardous environments, reallocating personnel from repetitive tasks to higher-value activities, and increasing employee tenure and job satisfaction (Baltrusch et al. 2021).

Identifying workers in high-exposure DWAs who are suitable for upskilling and internal mobility prior to installing automation preserves firm-specific knowledge and stabilises the workforce. The proposed framework supports AI-HRC planning by enabling organisations to quantify workforce impacts before deployment of automation and to invest in targeted capability development. From this perspective, AI-HRC represents an opportunity to enhance operations and invest in structured upskilling and career advancement. Such investments generate both financial and strategic benefits. Internal workforce development reduces recruiting costs and knowledge attrition, whereas external hiring is generally more expensive and associated with lower initial performance and higher turnover compared to internal movers (Bidwell 2011). Furthermore, organisational career growth is a strong predictor of employee retention, thereby reducing workforce volatility (Rani, Sood, and Chaurasia 2024).

In summary, proactive workforce planning strengthens supply chain resilience by promoting workforce continuity and delivering financial advantages compared to external hiring.

6. Conclusion, managerial implications, and future research

We present a stochastic, task-based workforce framework that converts O∗NET task frequencies into annualised occurrences, completes covariances under simplex constraints via a semidefinite program, and propagates uncertainty to task- and DWA-level full-time-equivalent (FTE) estimates. The method yields interpretable confidence intervals and detailed employment maps that show where intralogistics labour time is concentrated. Results indicate that the most AI-exposed DWAs span both operational and supervisory activities.

Three findings stand out. First, intralogistics employment is concentrated in a relatively small set of DWAs, so a focussed subset of activities accounts for a large share of the sector's labour time. Second, the top 15 percent of DWAs suitable for AI-HRC encompasses several hundred thousand FTEs, indicating that even conservative adoption paths will affect workforce allocations at scale. Third, the most exposed DWAs span operational as well as coordination and administrative functions, so that near-term change may be multifaceted and will require careful, task-level workforce planning.

RQ1: can we construct task-level full-time equivalent (FTE) estimates and characterise their statistical uncertainty?

Our findings show that a stochastic workforce framework can systematically estimate FTE requirements per task by modelling task frequencies as probability vectors on the simplex, completing their covariance structure, and propagating this uncertainty to task-level FTEs. The approach makes its assumptions explicit and evaluates mathematical consistency through feasibility, geometry, and sensitivity diagnostics. The resulting confidence intervals translate task frequency data from surveys into a range of workforce estimates, supporting more granular planning by identifying where labour time concentrates and how robust these estimates are to sampling and modelling uncertainty.

RQ2: how can task-level FTE estimates combined with AI exposure metrics inform strategic AI-HRC design and workforce planning?

Applied to the intralogistics sector, the stochastic workforce framework provides FTE estimates per task and per DWA, which can be combined with task-level automation exposure metrics to inform AI-HRC design. By linking FTE estimates to exposure indices, the framework identifies DWAs with both substantial labour demand and high suitability for AI-HRC, thereby quantifying the number of workers whose tasks are most exposed to reallocation or elimination under different automation scenarios. This supports scenario-based workforce planning (e.g. conservative versus aggressive AI-HRC adoption), allows managers to anticipate where headcount reductions or role redesign are most likely to be required, and highlights priority areas for upskilling and redeployment. In turn, such targeted planning can improve employee retention and, ultimately, enhance intralogistics and supply chain resilience.

6.1. Managerial implications

- **AI-HRC design.** Combining FTE estimates with automation exposure metrics enables targeted prioritisation of AI-HRC initiatives. High-FTE, high-exposure DWAs can be treated as priority candidates for intervention, whether through deploying collaborative robots, implementing advanced decision-support tools, or redesigning workflows. Conversely, DWAs with lower exposure or high uncertainty in FTE estimates may warrant more cautious experimentation. Aligning AI-HRC investments with this task-level analytics can help ensure that resources are focussed on activities where they are most likely to generate productivity gains.
- **Strategic workforce planning.** Task-level FTE estimates and their confidence intervals support scenario-based staffing decisions. By simulating conservative and aggressive AI-HRC adoption paths (for example, shifting a fraction of task time to automated systems), managers can quantify how many FTEs are affected under each scenario and assess the risk associated with different automation strategies. This moves workforce planning beyond deterministic headcount projections toward uncertainty-aware, task-driven analyses.
- **Resilience under uncertainty.** The framework supports proactive talent management. By identifying the DWAs where AI-HRC is most likely to change work content, firms can design upskilling and reskilling pathways that anticipate task reallocation rather than reacting to it ex post. In this way, AI-HRC can be positioned as an instrument for workforce stabilisation and long-term operational robustness.

6.2. Future research

Several avenues for future research emerge from this study. First, while the present analysis is deliberately scoped to intralogistics and AI-HRC, the proposed task-level FTE framework is inherently generalisable as it builds on the O∗NET occupational taxonomy which spans the full U.S. economy. Extending the empirical implementation to additional industries therefore requires no changes to the core methodology, only a redefinition of the occupational subset and associated task mappings. Such extensions would enable cross-industry workforce impact assessments and allow the framework to inform economy-wide workforce transition strategies. Second, relaxing the assumptions of equal task durations within occupations and uniform split of task time across linked DWAs using empirical data would allow richer heterogeneity as well as more precise task-DWA mappings. Third, the current analysis treats task frequencies and employment levels as cross-sectional. Extending the framework to panel data would enable analysis of dynamic adjustments in task composition and FTE allocations around AI-HRC adoption, including quasi-experimental designs. Fourth, combining the stochastic workforce framework with wage data would permit evaluation of how AI-HRC affects the distribution of labour income across DWAs, not just quantities of labour. Finally, the automation exposure metrics are derived from external text corpora; validating them with direct measures of technology deployment would strengthen the framework and better support the design of operationally effective and socially sustainable AI-HRC systems.

Acknowledgements

The authors are grateful to Iona Clark for her valuable feedback and support. They also thank colleagues at EPFL and the MIT Center for Transportation & Logistics for their insightful discussions and helpful comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Pierre Bouquet is a Ph.D. candidate at the Massachusetts Institute of Technology (MIT), working under the supervision of Prof. Yossi Sheffi. He received both his B.Sc. and M.Sc. degrees in Mechanical Engineering from the Swiss Federal Institute of Technology in Lausanne (EPFL), where he specialised in automation systems and pursued a minor in data science. His research interests include optimisation, machine learning, and data-driven methods for analyzing and designing complex systems. For inquiries or collaborations, he can be reached at pierre.bouquet.ch@gmail.com or via LinkedIn.

Nicolò Piergiovanni Bagnoli received a Bachelor's degree in Mechanical Engineering from the Università di Bologna and a Master's degree in Mechanical Engineering from EPFL. His academic research focuses on the impact of artificial intelligence on the workforce, with an emphasis on task composition, productivity, and organisational structure. This work is informed by prior operations experience in the aerospace, chemicals, and medical sectors, where he worked on manufacturing, operational processes, and strategy across Italy, Switzerland, and the United States.

Professor Yossi Sheffi is Director of the MIT Center for Transportation & Logistics (MIT CTL), and Director and Founder of the Master of Engineering in Logistics Program. He is a faculty member of the MIT Civil and Environmental Engineering Department, as well as the Institute for Data, Systems, and Society. He is a global expert in systems optimisation, risk analysis and supply chain management.

Data availability statement

The authors confirm that the data supporting the findings of this study are available. The experiments were conducted in a reproducible manner, and the source code, as well as data are available on a GitHub repository: https://github.com/Pierre-Bouquet/Estimating_the_Task_Content_of_Work_Workforce_Design_for_AI_HRC.

References

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives* 33 (2): 3–30. https://doi.org/10.1257/jep.33.2.3.

Aitchison, John. 1982. "The Statistical Analysis of Compositional Data." *Journal of the Royal Statistical Society: Series B (Methodological)* 44 (2): 139–160. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x.

Aleotti, Jacopo, Alberto Baldassarri, Marcello Bonfè, Marco Carricato, Davide Chiaravalli, Roberto Di Leva, Cesare Fantuzzi, et al. 2021. "Toward Future Automatic Warehouses: An Autonomous Depalletizing System Based on Mobile Manipulation and 3d Perception." *Applied Sciences* 11 (13): 5959. https://doi.org/10.3390/app11135959.

Autor, David, Caroline Chin, Anna Salomons, and Bryan Seegmiller. 2024. "New Frontiers: The Origins and Content of New Work, 1940–2018." *Quarterly Journal of Economics* 139 (3): 1399–1465. https://doi.org/10.1093/qje/qjae008.

Autor, David, and Neil Thompson. 2025. "Expertise." *Journal of the European Economic Association* 23 (4): 1203–1271. https://doi.org/10.1093/jeea/jvaf02.

Autor, David H. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29 (3): 3–30. https://doi.org/10.1257/jep.29.3.3.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review* 103 (5): 1553–1597. https://doi.org/10.1257/aer.103.5.1553.

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

Baltrusch, Saskia, Frank Krause, Aijse de Vries, and Michiel P. de Looze. 2021. "What about the Human in Human–Robot Collaboration? A Literature Review on HRC's Effects on Aspects of Job Quality." *Ergonomics* 65 (9): 1–35.

Bevington, Philip R., and D. Keith Robinson. 2003. *Data Reduction and Error Analysis for the Physical Sciences*. 3rd ed. Boston: McGraw-Hill.

Bidwell, Matthew. 2011. "Paying More to Get less: The Effects of External Hiring versus Internal Mobility." *Administrative Science Quarterly* 56 (3): 369–407. https://doi.org/10.1177/0001839211433562.

Bouquet, Pierre, Yossi Sheffi, and Amin Kaboli. 2026. "News Sentiment as a Dynamic Predictor of Job Automation Risk." *SSRN Electronic Journal* MIT Center for Transportation & Logistics Research Paper No. 2026/002. https://papers.ssrn.com/sol3/papers.cfm?abstract_id = 6168446.

Boyd, Stephen, and Lieven Vandenberghe. 2004. *Convex Optimization*. Cambridge, UK: Cambridge University Press.

Boysen, Nils, and René De Koster. 2025. "50 Years of Warehousing Research–an Operations Research Perspective." *European Journal of Operational Research* 320 (2): 449–464. https://doi.org/10.1016/j.ejor.2024.03.026.

Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond. 2025. "Generative AI at Work." *The Quarterly Journal of Economics* 140 (2): 889–942. https://doi.org/10.1093/qje/qjae044.

Brynjolfsson, Erik, and Tom Mitchell. 2017. "What Can Machine Learning Do? Workforce Implications." *Science* 358 (6370): 1530–1534. https://doi.org/10.1126/science.aap8062.

Cai, Min, Gen Wang, Xinggang Luo, and Xueqi Xu. 2025. "Task Allocation of Human–Robot Collaborative Assembly Line considering Assembly Complexity and Workload Balance." *International Journal of Production Research* 63 (13): 4749–4775. https://doi.org/10.1080/00207543.2024.2442546.

Cantrell, Sue, and Russell Klosk. 2025. "Is It Time to Break Workforce Planning out of Its Silo?" *Deloitte Insights*. Accessed November 01, 2025.

De Koster, René, and Debjit Roy. 2024. *Research: Warehouse and Logistics Automation Works Better with Human Partners*. Harvard Business Review. https://think.taylorandfrancis.com/special_issues/hrc-intralogistics/.

De Koster, René, Debjit Roy, Yun Fong Lim, and Subodha Kumar. 2025. "Internet of Things in Intralogistics: Applications and Emerging Research." *Production and Operations Management*. https://think.taylorandfrancis.com/special_issues/hrc-intralogistics/.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan R. Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and R. Karim. 2023. *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality*. Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013, The Wharton School Research Paper. https://ssrn.com/abstract = 4573321.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2024. "GPTs are GPTs: Labor Market Impact Potential of LLMs." *Science* 384 (6702): 1306–1308. https://doi.org/10.1126/science.adj0998.

Felten, Edward W., Manav Raj, and Robert Seamans. 2018. "A Method to Link Advances in Artificial Intelligence to Occupational Abilities." *AEA Papers and Proceedings* 108:54–57. https://doi.org/10.1257/pandp.20181021.

Fragapane, Giuseppe, René de Koster, Fabio Sgarbossa, and Jan Ola Strandhagen. 2021. "Planning and Control of Autonomous Mobile Robots for Intralogistics: Literature Review and Research Agenda." *European Journal of Operational Research* 294 (2): 405–426. https://doi.org/10.1016/j.ejor.2021.01.019.

Frey, Carl Benedikt, and Michael A. Osborne. 2017. "The Future of Employment: How Susceptible Are Jobs to Computerisation?" *Technological Forecasting and Social Change* 114:254–280. https://doi.org/10.1016/j.techfore.2016.08.019.

Goodman, Leo A. 1960. "On the Exact Variance of Products." *Journal of the American Statistical Association* 55 (292): 708–713. https://doi.org/10.1080/01621459.1960.10483369.

Grussler, Christian, Anders Rantzer, and Pontus Giselsson. 2018. "Low-Rank Optimization with Convex Constraints." *IEEE Transactions on Automatic Control* 63 (11): 4000–4007. https://doi.org/10.1109/TAC.9.

Horn, Roger A., and Charles R. Johnson. 2012. *Matrix Analysis*. 2nd ed.. Cambridge University Press. https://www.amazon.com/Matrix-Analysis-Second-Roger-Horn/dp/0521548233.

Huber, Peter J., and Elvezio M. Ronchetti. 2009. *Robust Statistics*. 2nd ed. Hoboken, New Jersey: Wiley. https://onlinelibrary.wiley.com/doi/book/10.1002/9780470434697.

Li, Mingxing, Daqiang Guo, Ming Li, Ting Qu, and George Q. Huang. 2023. "Operation Twins: Production–Intralogistics Synchronisation in Industry 4.0." *International Journal of Production Research* 61 (15): 5193–5211. https://doi.org/10.1080/00207543.2022.2098874.

Liu, Huwei, Li Zhou, Junhui Zhao, Fan Wang, Jianglong Yang, Kaibo Liang, and Zhaochan Li. 2022. "Deep-Learning-based Accurate Identification of Warehouse Goods for Robot Picking Operations." *Sustainability* 14 (13): 7781. https://doi.org/10.3390/su14137781.

Liu, Yu, Shenle Pan, and Eric Ballot. 2024. "Unveiling the Potential of Digital Twins in Logistics and Supply Chain Management: Services, Capabilities, and Research Opportunities." *Digital Engineering* 3:100025. https://doi.org/10.1016/j.dte.2024.100025.

Ma, Benedict Jun, and María Jesús Saénz. 2025. "AI Can Improve How Humans and Robots Work: Warehouse Automation Is Booming as Companies Ramp up Their Use of Robotics." *MIT Sloan Management Review* 67 (1): 80–85. https://sloanreview.mit.edu/article/ai-can-improve-how-humans-and-robots-work/.

Martin, Josh, and Ellys Monahan. 2022. *Developing a Method for Measuring Time Spent on Green Tasks*. Technical Report. Office for National Statistics.

MOSEK ApS. 2025. "The MOSEK Python Fusion API manual. Version 11.0." https://docs.mosek.com/latest/pythonfusion/index.html.

Nagurney, Anna, and Alireza Ermagun. 2022. "Resilience of Supply Chain Networks to Labor Disruptions." *Findings*.

National Center for O∗NET Development. 2024. "O∗NET 28.3 Database." May. Version 28.3. https://www.onetcenter.org/db_releases.html.

Nourmohammadi, A., T. Arbaoui, M. Fathi, and A. Dolgui. 2025. "Balancing Human–Robot Collaborative Assembly Lines: A Constraint Programming Approach." *Computers & Industrial Engineering* 205:111154. https://doi.org/10.1016/j.cie.2025.111154.

Noy, Shakked, and Whitney Zhang. 2023. "Experimental Evidence on the Productivity Effects of Generative Artificial

Intelligence." *Science* 381 (6654): 187–192. https://doi.org/10.1126/science.adh2586.

Park, Kyu Tae, Yoo Ho Son, and Sang Do Noh. 2021. "The Architectural Framework of a Cyber Physical Logistics System for Digital-Twin-based Supply Chain Control." *International Journal of Production Research* 59 (19): 5721–5742. https://doi.org/10.1080/00207543.2020.1788738.

Pasparakis, Alexandros, Jelle de Vries, and René de Koster. 2023. "Assessing the Impact of Human–Robot Collaborative Order Picking Systems on Warehouse Workers." *International Journal of Production Research* 61 (22): 7776–7790. https://doi.org/10.1080/00207543.2023.2183343.

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. 2023. "The Impact of Ai on Developer Productivity: Evidence from Github Copilot." *arXiv preprint arXiv:2302.06590*.

Pennathur, Priyadarshini R., Valerie Boksa, Arunkumar Pennathur, Andrew Kusiak, and Beth A. Livingston. 2024. "The Future of Office and Administrative Support Occupations in the Era of Artificial Intelligence: A State of the Art Review and Future Research Directions." *International Journal of Industrial Ergonomics* 104:103665. https://doi.org/10.1016/j.ergon.2024.103665.

Perotti, Sara, Luca Cannava, Jörg M. Ries, and Eric H. Grosse. 2025. "Reviewing and Conceptualising the Role of 4.0 Technologies for Sustainable Warehousing." *International Journal of Production Research* 63 (6): 2305–2337. https://doi.org/10.1080/00207543.2024.2396015.

Rani, Nidhi, Gaytri Sood, and Shiv S. Chaurasia. 2024. "Exploring Organizational Career Growth: A Systematic Literature Review and Future Research Directions." *The International Journal of Human Resource Management* 35 (17): 4081–4117.

Saltelli, Andrea, Stefano Tarantola, Francesca Campolongo, and Marco Ratto. 2004. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. Wiley.

Särndal, Carl-Erik, Bo Swensson, and Jan H. Wretman. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag. https://www.google.com/books/edition/Model_Assisted_Survey_Sampling/ufdONK3E1TcC?hl=en.

Teece, David J., Gary P. Pisano, and Amy Shuen. 1992. *Dynamic Capabilities and Strategic Management*. Berkeley: Center for Research in Management, University of California.

Tomlinson, Kiran, Sonia Jaffe, Will Wang, Scott Counts, and Siddharth Suri. 2025. "Working with AI: Measuring the Applicability of Generative AI to Occupations." *arXiv preprint arXiv:2507.07935*.

Tsacoumis, Suzanne, and Chad H. Van Iddekinge. 2006. *A Comparison of Incumbent and Analyst Ratings of O∗NET Skills*. Technical Report. O∗NET Resource Center. https://www.onetcenter.org/dl_files/SkillsComp.pdf.

U.S. Bureau of Labor Statistics. 2024. "Occupational Employment and Wages – May 2024." Apr. News Release Summarizing May 2024 OEWS. https://www.bls.gov/news.release/ocwage.htm.

van der Vaart, Aad W. 2000. *Asymptotic Statistics*. Cambridge University Press. https://www.google.com/books/edition/Asymptotic_Statistics/UEuQEM5RjWgC?hl=en.

Wolkowicz, Henry, Romesh Saigal, and Lieven Vandenberghe. 2000. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*. Springer.

Xu, Y., Y. Liu, and B. Montreuil. 2024. "Dynamic Workforce Scheduling and Relocation in Hyperconnected Parcel Logistic Hubs." https://www.semanticscholar.org/paper/Dynamic-Workforce-Scheduling-and-Relocation-in-Hubs-Xu-Liu/e4ad040ce164e66069dd0506cdc1344cb651ef54.

Yuan, Ming, and Yin Zhang. 2013. "Robust Covariance Matrix Completion." *Biometrika* 100 (4): 905–920.

Yue, Man-Chung, Yves Rychener, Daniel Kuhn, and Viet Anh Nguyen. 2024. "A Geometric Unification of Distributionally Robust Covariance Estimators: Shrinking the Spectrum by Inflating the Ambiguity Set." *arXiv preprint arXiv:2405.20124*.