Occupational tasks and AI exposure

What can a task score tell us about a job?

Sergio Ocampo
ECON 3399 · AI and Robots at Work

Economics and business career paths · O*NET · Evidence on task exposure

Can my job be taken by AI?

AI (or robots) do not take over jobs ⟶ they take over tasks.

> Ask "How exposed are my tasks to AI?" rather than "Will my job be taken?"

Consider an **accountant** who prepares a financial report.

- Retrieve & organize records
- Check calculations; investigate unusual entries
- Explain findings to the client; take responsibility for the report

> AI and robots can change some of these activities while the occupation keeps its name.

O*NET describes occupations in the U.S. economy

O*NET describes occupations through the tasks people perform.

- Workers: task descriptions + ratings of their work
- Experts + trained analysts: additional judgments
- Other sources: employer postings, research & government records

Economists (19-3011.00)

2025 median wage: **$124,720/year**
2025 employment: **18,600**
2025–35 projection: **+5–6%**

Open profile O*NET data

Credit Analysts (13-2041.00)

2025 median wage: **$83,510/year**
2025 employment: **64,700**
2025–35 projection: **-1% or lower**

Open profile O*NET data

- Identify the tasks performed in each occupation
- Connect each task to the skills needed

> Occupations are connected to education programs through skills.

O*NET data collection · Economists: BLS figures · Credit Analysts: BLS figures

Economics and business programs connect to occupations

Program	Selected linked occupations
Economics programs	
Economics, general	Economists; survey researchers; social science research assistants
Applied economics	Market research analysts; economists; social science research assistants
Econometrics & quantitative economics	Data scientists; business intelligence analysts; statisticians
Business/managerial economics	Management analysts; financial risk specialists; economists
Business programs	
Business administration	Management analysts; general & operations managers; sales managers
Finance	Credit analysts; financial & investment analysts; financial risk specialists
Marketing	Market research analysts; marketing managers; sales managers

O*NET links programs to occupations they can prepare students for; degree level is unspecified.

Economic and business knowledge is used across these occupations

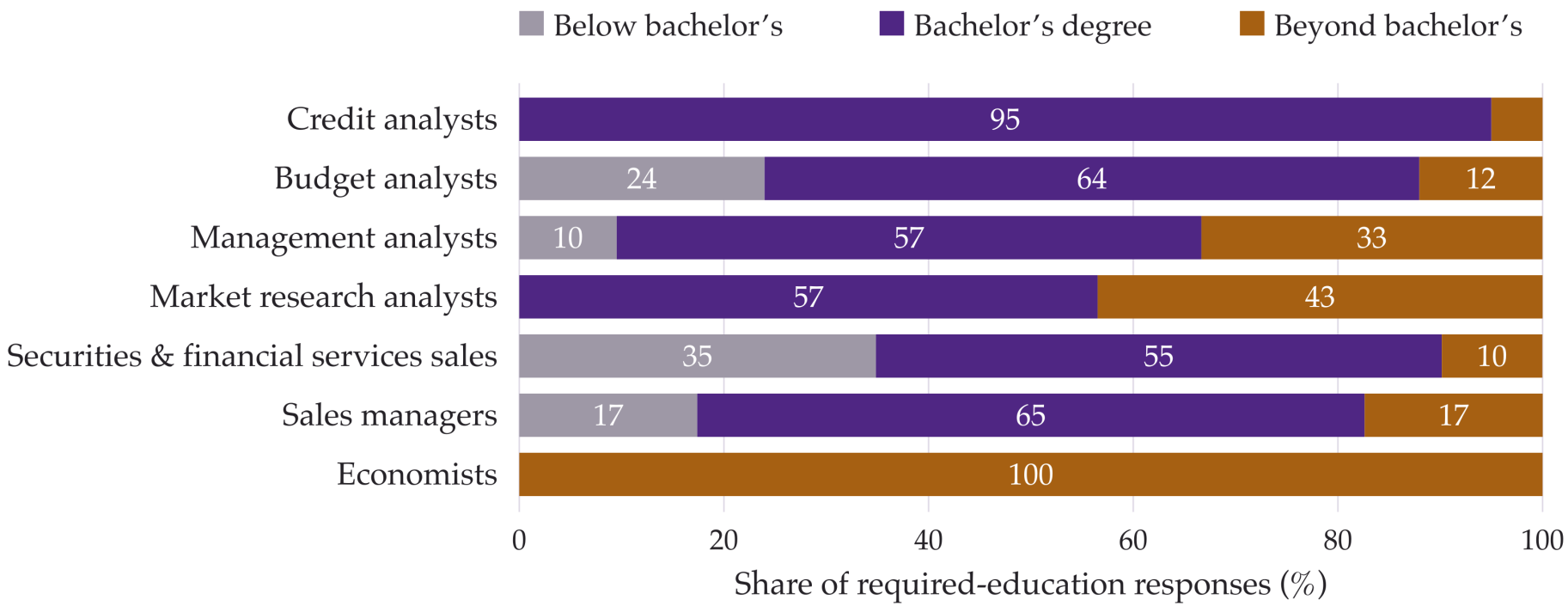
	Economics & accounting	Administration & management	Sales & marketing	Mathematics
Economists	4.50	2.57	1.87	4.57
Management analysts	3.48	4.38	3.30	3.62
Credit analysts	4.50	2.80	2.05	3.90
Budget analysts	4.03	3.71	1.15	3.90
Market research analysts	3.09	3.74	4.09	4.04
General & operations managers	3.47	4.35	3.47	3.56
Sales managers	2.91	4.00	4.70	3.09
Securities & financial services sales	3.93	3.42	3.71	3.73

Importance scale: 1 = not important · 3 = important · 5 = extremely important

Original 1–5 importance ratings, O*NET 27.2. "Economics and accounting" is one combined knowledge category.

> Credit analysts rate economics and accounting as highly as economists do (4.50). Management and sales work combine different kinds of knowledge.

Which of these jobs call for undergraduate preparation?



Required education, not respondents' own qualifications. O*NET 27.2; observations: 2016–2022. Rounded percentages.

> A bachelor's degree is the most common response for these analyst and sales roles. The "economist" occupation is much more oriented toward advanced study.

A reproducible route to the data

O*NET database archive

Open in a browser tab

1. Use release 27.2 for the exposure-paper exercise
2. Text database: Occupation Data + Task Statements
3. Identifiers: occupation codes + task IDs
4. Record release, retrieval date & filters

From O*NET descriptions to AI exposure

Question	**Object we would like to measure**
Could AI help perform the task?	Technical exposure
Do workers use AI for that task?	Adoption and intensity of use
What changes because workers gain access?	Productivity, employment, wages, or quality

$$c(i) = \min\left\{\frac{w}{a_L(i)}, \frac{r}{a_K(i)}\right\}$$

Today we start with exposure. The next two lectures examine adoption and effects.

GPTs are GPTs: Labor Market Impact Potential of LLMs



INSIGHTS

Check for updates

POLICY FORUM

ARTIFICIAL INTELLIGENCE

GPTs are GPTs: Labor market impact potential of LLMs

Research is needed to estimate how jobs may be affected

By **Tyna Eloundou[1], Sam Manning[2], Pamela Mishkin[1], Daniel Rock[3]**

We propose a framework for evaluating the potential impacts of large-language models (LLMs) and associated technologies on work by considering their relevance to the tasks workers perform in their jobs. By applying this framework (with both humans and using an LLM), we estimate that roughly 1.8% of jobs could have over half their tasks affected by LLMs with simple interfaces and general training. When accounting for current and likely future software developments that complement LLM capabilities, this share jumps to just over 46% of jobs. The collective attributes of LLMs such as generative pretrained transformers (GPTs) strongly suggest that they possess key characteristics of other "GPTs," general-purpose technologies (*1*, *2*). Our research highlights the need for robust societal evaluations and policy measures to address potential effects of LLMs and complementary technologies on labor markets.

We consider the progress of LLMs' capabilities and the potential breadth of the complementary technologies that they spawn, underscoring that maximizing the impact of LLMs requires their integration within broader systems (*3*–*5*). The rubric that we develop (see supplementary materials (SM) section 3.1.2) defines task exposure to LLMs, in the spirit of prior work on quantifying exposure to machine learning (*6*–*8*). Following prior literature, we use the concept of exposure as a proxy for potential economic impact, irrespective of labor-augmentation or displacement effects (see section 2 of the SM for further discussion of the literature). Our approach differs from the broader scope of complementary innovations in general-purpose technology discussions, focusing more narrowly on advanced software capabilities than on the potential for business process reengineering, new intangible assets creation, or workforce retraining. General-purpose technologies such as electricity or computing historically have had far-reaching effects that took decades to fully materialize. With evidence of the general-purpose technology potential of LLMs, we urge caution in making long-term predictions while offering an outline of where work might change.

[1]OpenAI, San Francisco, CA, USA. [2]Centre for the Governance of AI, Oxford, UK. [3]Wharton School, University of Pennsylvania, Philadelphia, PA, USA. Email: pamela@openai.com

RATING TASKS AND OCCUPATIONS

We use the O*NET 27.2 database (*9*), which covers 1016 occupations and their Detailed Work Activities (DWAs) and tasks, focusing on the 923 occupations with available task information.

DWAs describe broader activities that represent the work performed within an occupation, e.g., a DWA of "Execute sales or other financial transactions" might be associated with a range of occupations. Tasks are more specific work units linked to one or more DWAs, such as the task "Deliver email confirmation of completed transactions and shipment." O*NET categorizes tasks as "Core" or "Supplemental," with the latter weighted half as much as the former in our analysis (or noted otherwise). Results are robust to alternative weightings (SM section 9.3). We sourced employment and wage data from the 2020–2022 Occupational Employment series by the US Bureau of Labor Statistics (BLS).

Although the term "LLM" can refer to a broad array of data modalities, we confine our definition of LLM-powered software to principally apply to text, code, and images. We do not tie this analysis to a particular model or model provider but rather to a paradigm of release (through chatbot and application programming interfaces) that was dominant in mid-2023. Progress has been made integrating LLMs with robotics, but the uncertainty around their reliability and adoption, as well as the difficulty conveying the potential of these integrations in our rubric format, led us to exclude assumptions about future robotics integrations from our analysis.

We collect results by applying our exposure rubric (SM section 3.1), where we define exposure as the capacity of an LLM or LLM-powered system to reduce the time required for a human to complete a task by at least 50% while preserving or improving quality. The chosen 50% time reduction threshold, though arbitrary, was used for practical purposes to facilitate interpretation by human annotators and GPT-4. Our estimates of task-level potential impacts are likely to be greater than real-world productivity gains, which will be influenced by factors other than the capabilities of LLMs and complementary technologies (*10*).

Our analysis uses four key measures to gauge exposure levels at the occupation level: (i) E1, the share of an occupation's tasks where access to an LLM alone or with a simple interface would lead to 50% time savings; (ii) E2, the share of an occupation's tasks where additional software is needed on top of an LLM to realize 50% time savings; (iii) E1 + E2, the share of tasks exposed assuming full LLM and software integration; and (iv) E1 + 0.5 × E2, an intermediate measure of exposure while complementary technologies are not fully implemented.

The authors, along with a group of contractors experienced in evaluating LLM outputs, labeled the exposure of DWAs and tasks in the O*NET dataset according to the rubric. We ask human annotators to use their best judgment in labeling DWAs and tasks as E0, E1, or E2 accord-

Human and GPT-4 ratings

Mean and standard deviation of ratings from human and GPT-4 labels are highly similar. Occupation level represents the average share of tasks within an occupation exposed. Task level represents the share of tasks across all occupations exposed.

Occupation-level exposure

	Human		GPT-4	
	Mean	SD	Mean	SD
E1	0.14	0.14	0.14	0.16
E1 + 0.5 × E2	0.30	0.21	0.34	0.22
E1 + E2	0.46	0.30	0.55	0.34

Task-level exposure

	Human		GPT-4	
	Mean	SD	Mean	SD
E1	0.15	0.36	0.14	0.35
E1 + 0.5 × E2	0.31	0.37	0.35	0.35
E1 + E2	0.47	0.50	0.56	0.50

Downloaded from https://www.science.org at University of Western Ontario on April 13, 2026

Published article · First page

Tyna Eloundou, Sam Manning,
Pamela Mishkin & Daniel Rock

Science · 2024

Open the paper

- Describe occupational tasks
- Apply rubric: human annotators + GPT-4
- Task labels ⟶ occupation exposure measures

The exposure rubric

LLM Exposure: Using LLM reduces time by at least 50% (preserving quality)

- Distinguish core tasks (weight $\omega = 1$) from supplemental tasks (weight $\omega = 1/2$)

Label	Interpretation
E0	The task does not meet the criterion.
E1	An LLM alone or a simple interface meets the criterion.
E2	Additional software built around the LLM is needed.

Three exposure measures: How AI is adopted and integrated in production

$$\alpha = E1$$

$$\beta = E1 + \tfrac{1}{2}E2$$

$$\gamma = E1 + E2$$

An illustrative calculation

Task	Type	Label	Weight
Prepare a standard draft	Core	E1	1
Check an unusual case	Core	E0	1
Organize source material	Core	E2	1
Prepare a presentation	Supplemental	E2	$\frac{1}{2}$

$$\alpha = \frac{1}{3.5} \approx 0.29$$

$$\beta = \frac{1 + 0.5(1.5)}{3.5} = 0.50$$

$$\gamma = \frac{2.5}{3.5} \approx 0.71$$

Consulting: the same job combines different exposure labels

Management analyst task (abbreviated)	Human label
Plan a study of work problems and procedures	E1 · LLM alone
Analyze collected data and develop alternative solutions	E2 · software needed
Interview personnel and observe work on site	E0 · criterion not met

> The potential gain spans study design and analysis. Learning how an organization actually works can remain a separate, time-consuming part of the assignment.

The task bundle helps us see where a consultant's workflow could change, and where cheaper analysis may increase the value of gathering good information.

Finance and sales: communicating with clients can also be exposed

Occupation and task (abbreviated)	Label
Credit analyst: complete loan applications and credit summaries	E1
Credit analyst: calculate ratios to evaluate financial status	E2
Credit analyst: contact customers to collect overdue payments	E0
Financial services sales: explain stock-market terms to clients	E1
Financial services sales: analyze factors moving prices	E2
Financial services sales: buy or sell securities for dealers	E0

> The distinction is task-specific. Explaining information, analyzing data, executing transactions, and collecting payments receive different labels.

Accessing the paper's files

GPTs-are-GPTs public repository

Open in a browser tab

1. Occupation scores: data/occ_level.csv
2. Task statements + exposure labels: data/full_labelset.tsv
3. Compare human_rating_beta with dv_rating_beta
4. Exposure labels ≠ the repository's automation rubric

Explore the occupations relevant to this class

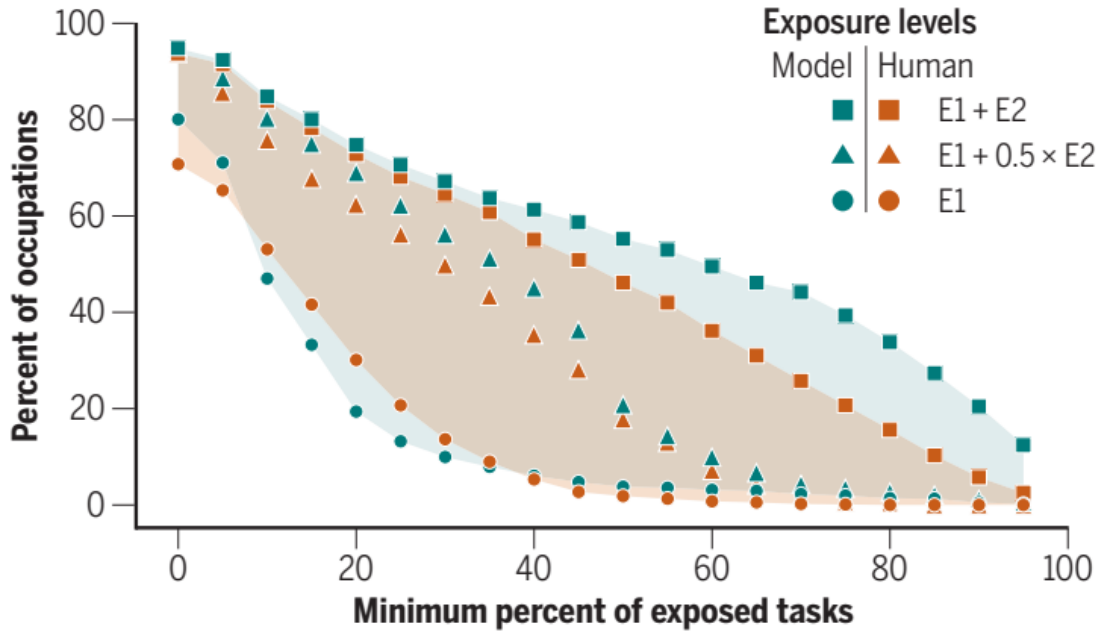
Management Analysts · 13-1111.00 · human labels

α: LLM alone	β: partial software	γ: full software
19.0%	50.0%	81.0%

Selected task statement	Type	Label
Gather and organize information on problems or procedures.	Core	E2
Confer with personnel concerned to ensure successful functioning of newly implemented systems or procedures.	Core	E0
Analyze data gathered and develop solutions or alternative methods of proceeding.	Core	E2
Document findings of study and prepare recommendations for implementation of new systems, procedures, or organizational changes.	Core	E2
Plan study of work problems and procedures, such as organizational change, communications, information flow, integrated production methods, inventory control, or cost analysis.	Core	E1

Saved example. Exposure depends on the paper's rubric and technology assumptions. Use the HTML presentation to change occupations or labels.

Exposure is widespread, but its intensity varies



Exposure curves: x = task-share threshold; y = share of occupations above the threshold.

> At the 50% task threshold, the human-rated share of occupations rises from 1.8% with an LLM alone to 46.2% with complementary software.

Complementary software changes the size of exposure

Mean exposed task share per occupation	Human ratings	GPT-4 ratings
LLM alone: α	14%	14%
Partial software: β	30%	34%
Full software: γ	46%	55%

Published occupation means: equal weight across occupations; weighted tasks within each occupation.

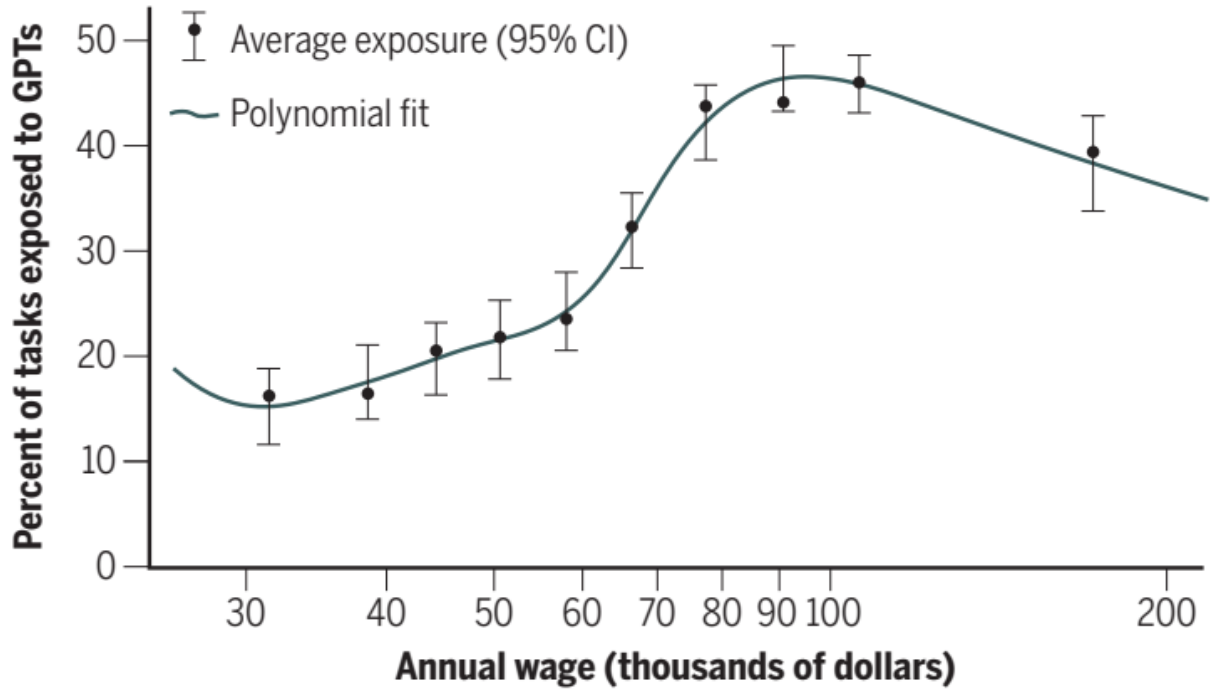
> The largest change comes from what is built around the model. An LLM-alone score misses many tasks that require software integration.

Exposure reaches many workers, with different intensities

Measure and denominator	Reported share
Occupations with ≥50% tasks exposed, α	1.8%
Occupations with ≥50% tasks exposed, β	17.8%
Occupations with ≥50% tasks exposed, γ	46.2%
Workers in occupations with ≥10% tasks exposed, β	About 80%
Workers in occupations with >50% tasks exposed, β	18.5%

> Compare α with γ: 1.8% versus 46.2% of occupations have at least half their tasks exposed. Complementary software expands potential time savings.

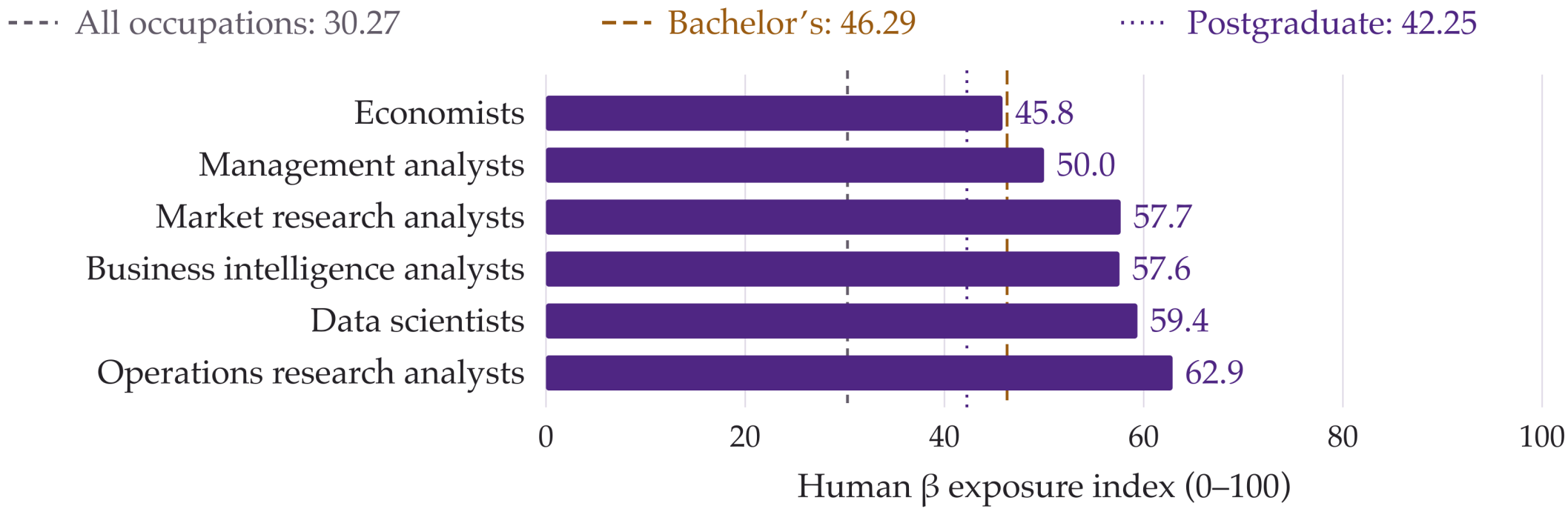
Higher-paying occupations are often more exposed



Human β exposure by occupational median annual wage. Wage bins; bars show 95% confidence intervals.

> Exposure rises with occupational pay. Text, code, and information processing put many well-paid jobs within the reach of LLM assistance.

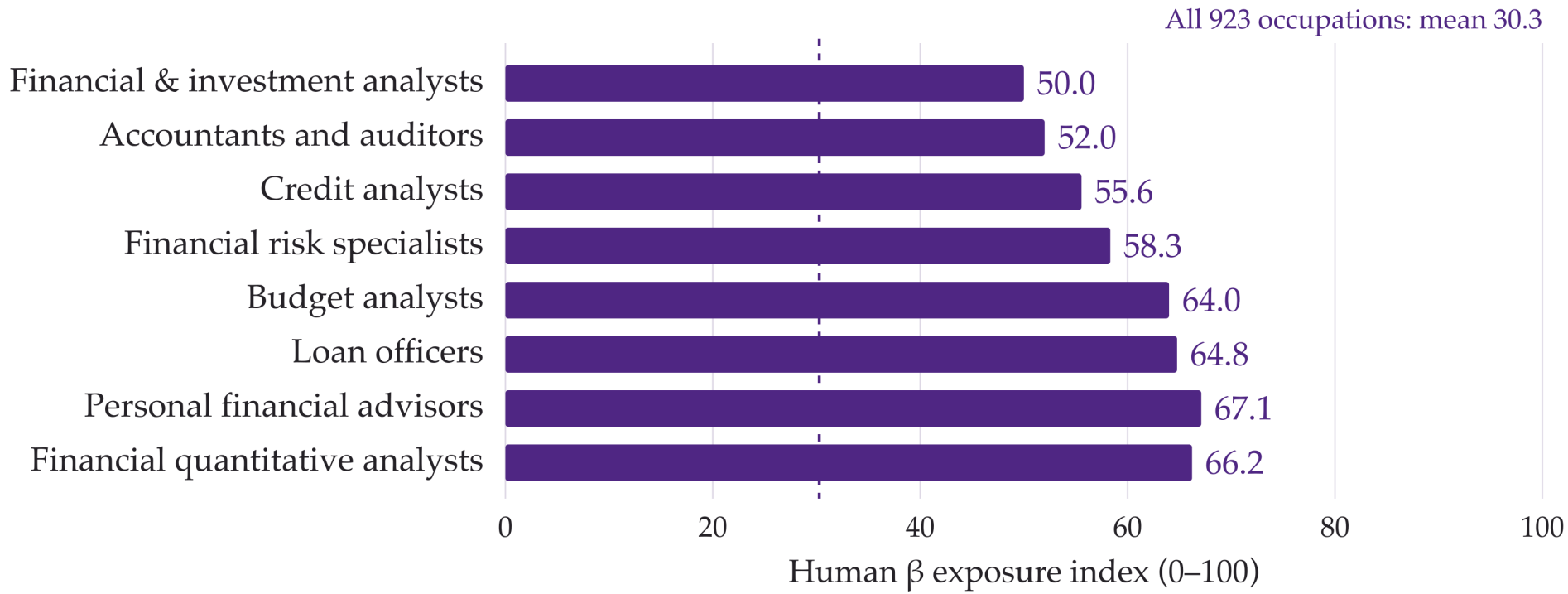
Consulting and analysis: exposure extends beyond the economist title



Unweighted human β. Modal O*NET education: bachelor's (211 occupations); postgraduate (133).

> Both degree groups exceed 30.27 overall. Bachelor's jobs average more exposure than postgraduate jobs, reflecting different task mixes.

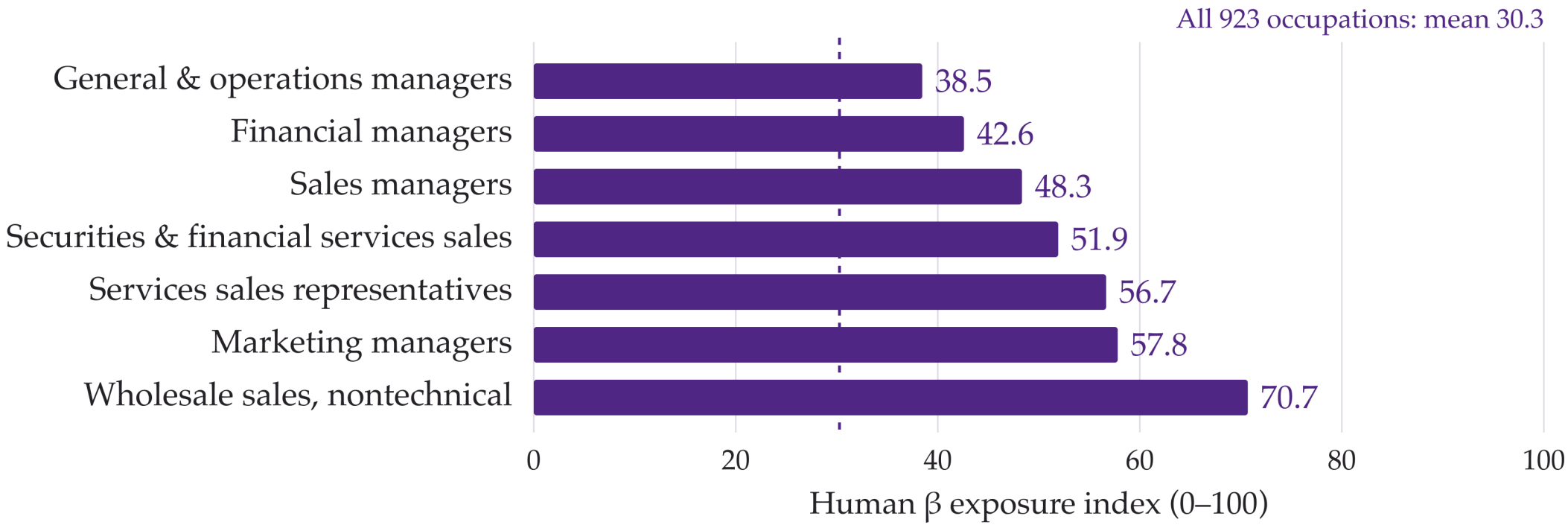
Banks and finance: many different jobs have substantial exposure



Human β index; selected occupations, not a finance-sector employment average. Dashed line: all-occupation mean.

> Credit, risk, and investment analysis are all exposed. Personal financial advice also scores highly (67.1), even though it involves clients.

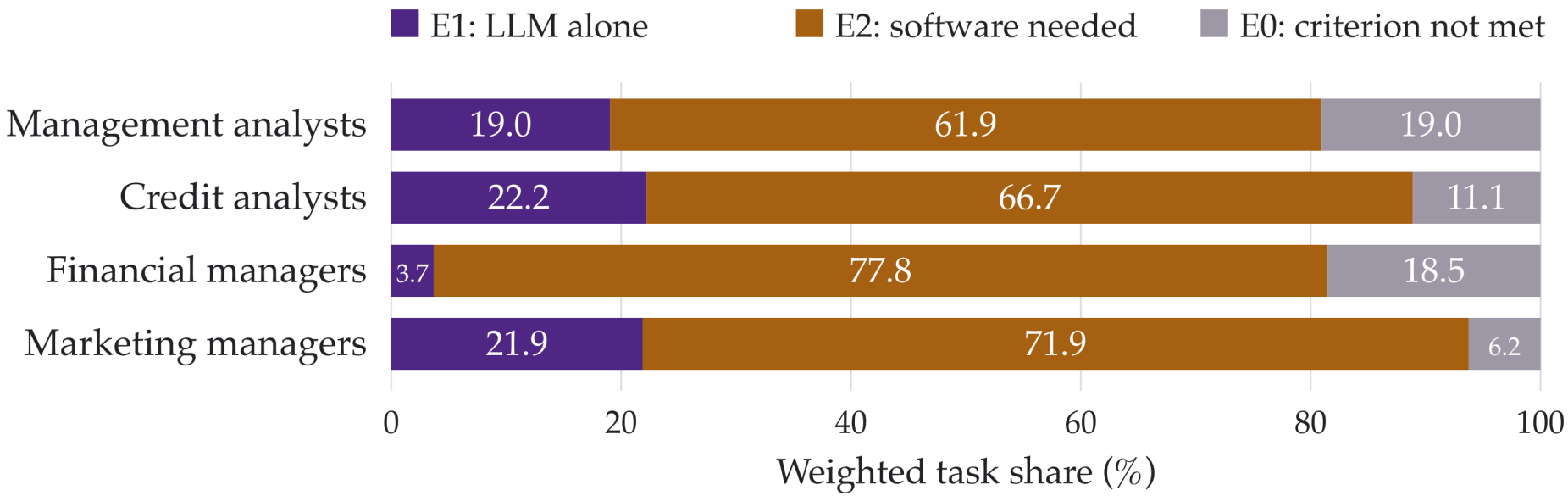
Management and sales are exposed in different ways



Shortened labels; "wholesale sales, nontechnical" excludes technical and scientific products. Human β index.

> General managers score 38.5, but nontechnical wholesale sales representatives score 70.7. "Working with people" is not enough to predict low exposure.

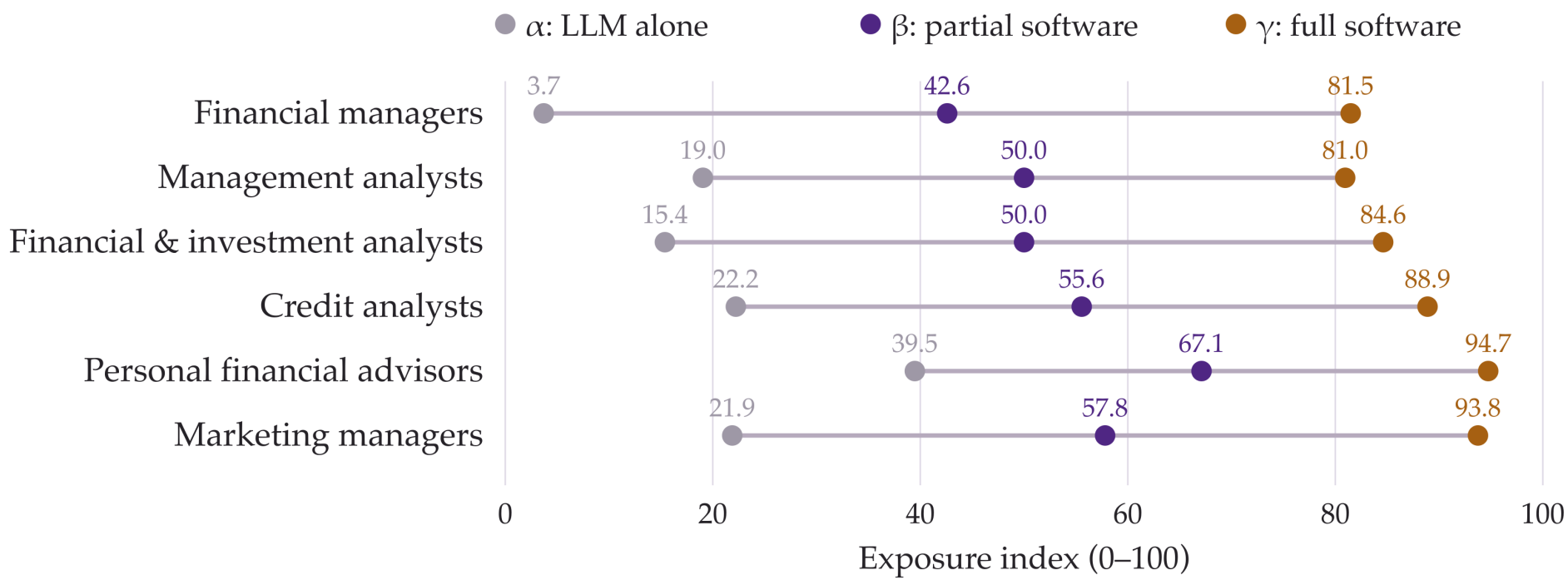
What is behind the occupation average?



Human task labels. Core tasks receive weight 1; supplemental tasks $\frac{1}{2}$. These four occupations have no missing task types.

For management analysts, $19.0 + \frac{1}{2} \times 61.9 = 50.0$. Most exposure comes from tasks needing complementary software, while 19.0% do not meet the criterion.

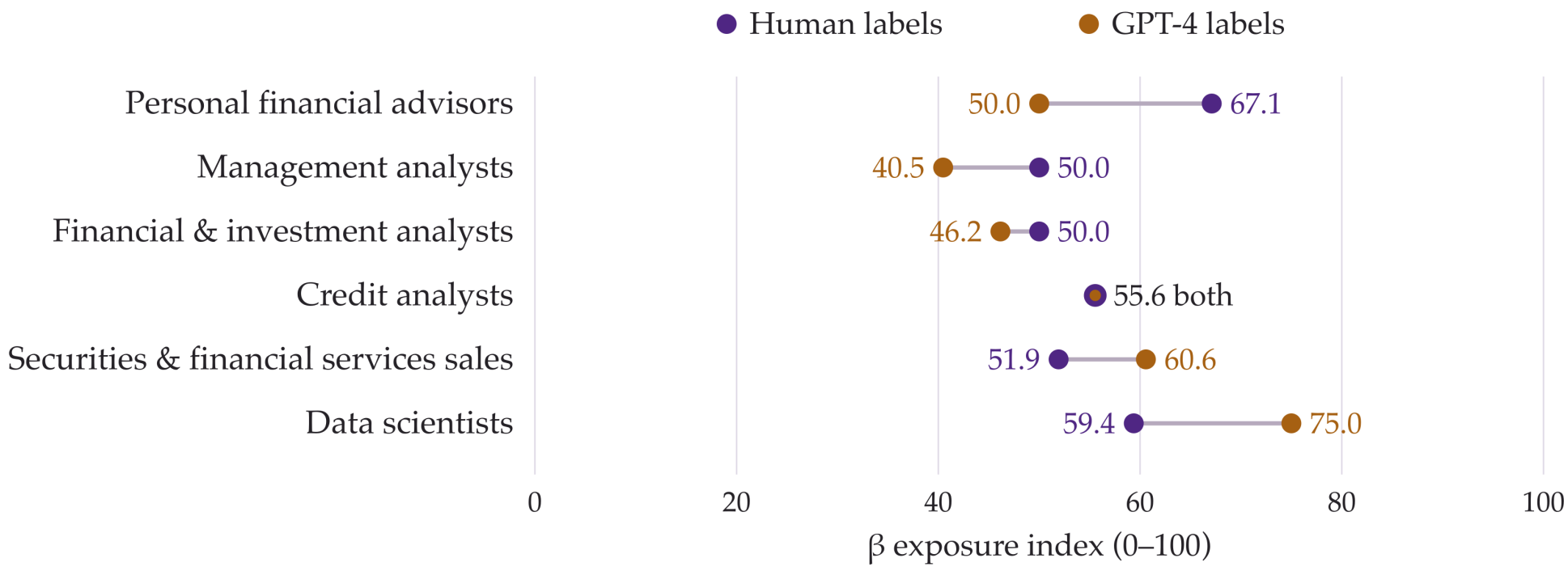
Software integration changes the comparison



Three capability assumptions applied to the same occupational task descriptions. Connecting lines are not confidence intervals.

> Financial managers move from 3.7 with an LLM alone to 81.5 with full software support. Their low α score hides a large set of E2 tasks.

The ranking also depends on who assigns the labels



Same data + β definition; different raters. Lines show rater disagreement, not sampling uncertainty.

Occupational exposure to artificial intelligence



Publisher's article page

Strategic Management Journal · 2021 · Open the paper

Felten: AI applications ⟶ abilities ⟶ occupations

- **Only 10 AI applications**: language, vision, speech, games, music
- **Online mTurk survey**: people link applications to **52 O*NET abilities**
- Ability importance × level ⟶ occupation exposure

	Eloundou et al.	**Felten, Raj & Seamans**
Building block	Task descriptions	Occupational abilities
AI assessment	≥50% time saving	Application–ability links
Output	Weighted task share	Standardized AIOE score

Felten et al. (2021), §2

Felten, Section 3: why occupations differ

Occupation pair	AIOE percentile	Abilities behind the gap
Surgeons	52nd	Similar dexterity;
Meat slaughterers	2nd	more reasoning in surgery
Accountants & auditors	99th	Same cognitive share;
Mathematical technicians	68th	different sensory demands

Accounting + mathematical technicians: **80% cognitive abilities**.

> The mix of abilities matters, even among highly cognitive occupations.

Felten et al. (2021), §3.2

Anthropic: observed AI exposure


ANTHROP\C

Labor market impacts of AI:
A new measure
and early evidence

Published
March 5, 2026

Authors
Maxim Massenkoff and Peter McCrory

Acknowledgements
Ruth Appel, Tim Belonax, Keir Bradwell, Andy Braden, Dexter Callender III, Miriam Chaum, Madison Clark, Jake Eaton, Deep Ganguli, Kunal Handa, Ryan Heller, Lara Karadogan, Jennifer Martinez, Jared Mueller, Sarah Pollack, David Saunders, Carl De Torres, Jack Clark.

We additionally thank Martha Gimbel, Anders Humlum, Evan Rose, and Nathan Wilmers for feedback on earlier versions of this note.

Research report · March 5, 2026

Labor market impacts of AI:
A new measure and early evidence

Maxim Massenkoff
Peter McCrory

Potential capability + observed use

Visit the research website
Explore the public data

Three sources, matched task by task

Source	What it contributes
O*NET	Task descriptions + occupation links
Eloundou et al.	Task feasibility: LLM alone or LLM + software
Anthropic Economic Index	Claude usage + work/automation labels

- Claude.ai: retain work-related conversations
- API: software-integrated use, treated as automated work

> Task descriptions join capability assessments to evidence of actual use.

From tasks to an occupation exposure score

$$E = \sum_t w_t f_t u_t a_t$$

Factor	Meaning
w_t: time weight	Share of working time, estimated by Claude
f_t: feasible	1 if Eloundou's task rating is 0.5 or 1; otherwise 0
u_t: observed at work	1 if work-related usage count $\geq$100; otherwise 0
a_t: automation weight	0.5 for augmentation; 1 for automation; mixtures between

> More work time + more automated use $\longrightarrow$ greater observed exposure.

One task's contribution: classifying patient records

Health information technologists · compile + classify patient data

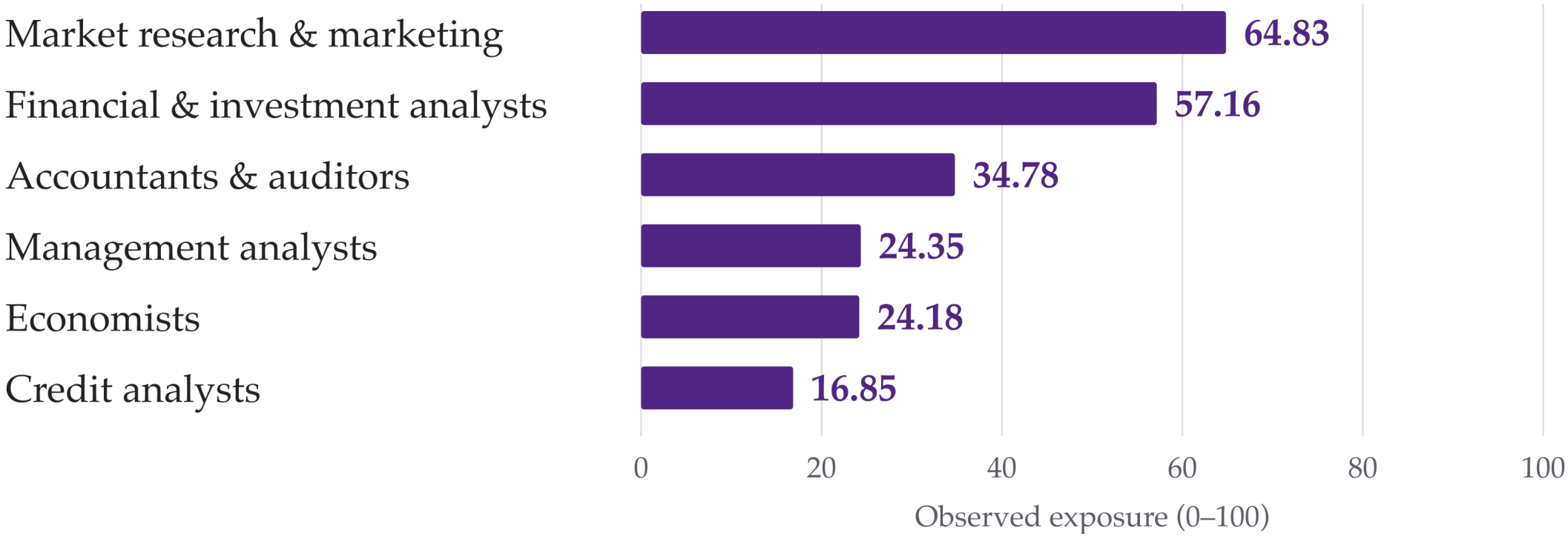
Time share	Feasible?	Work usage	Automation weight
$w_t = 0.13$	$f_t = 1$	>2,000 uses: $u_t = 1$	$a_t = 0.96$

$$0.13 \times 1 \times 1 \times 0.96 = 0.1248$$

> This task adds ≈12.48 points to the occupation's 0–100 score.

Published appendix example; calculation uses its rounded inputs.

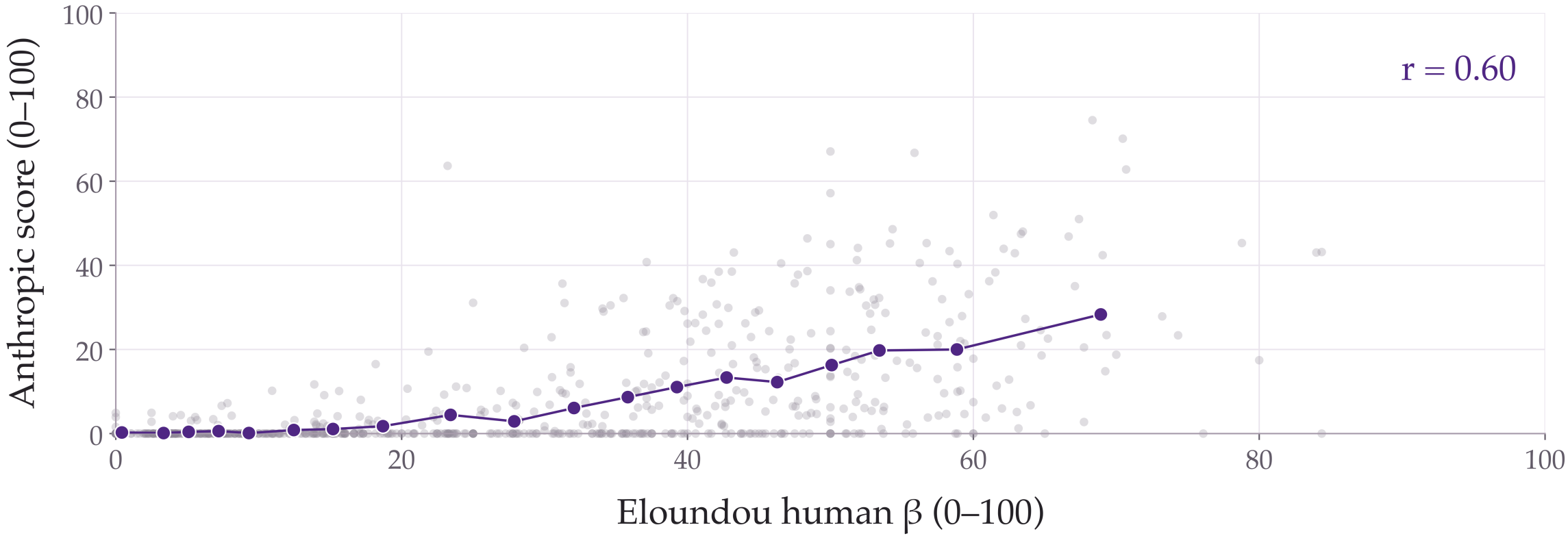
Observed exposure in familiar careers



Released occupation scores ×100; selected occupations, shortened labels.

> Finance roles span 16.85–57.16: task mix + observed use both matter.

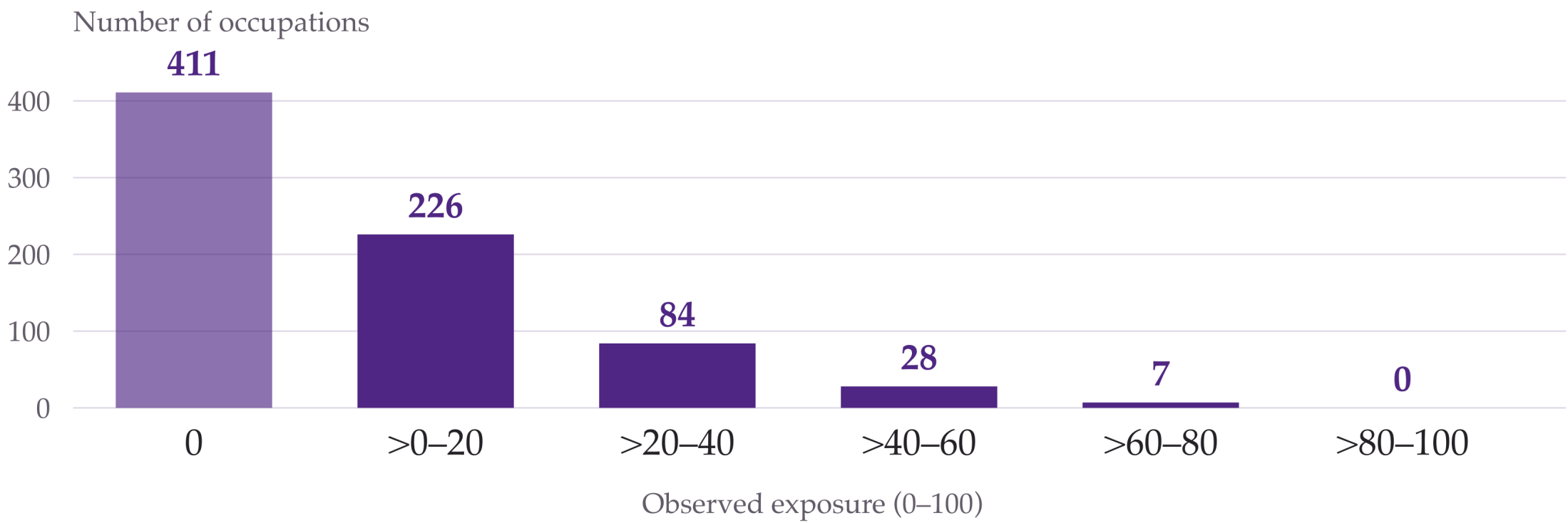
Potential exposure and observed AI use



684 matched occupations · gray: individual scores · purple: averages within 19 bins

> Higher potential exposure goes with more observed use, but occupations vary widely.

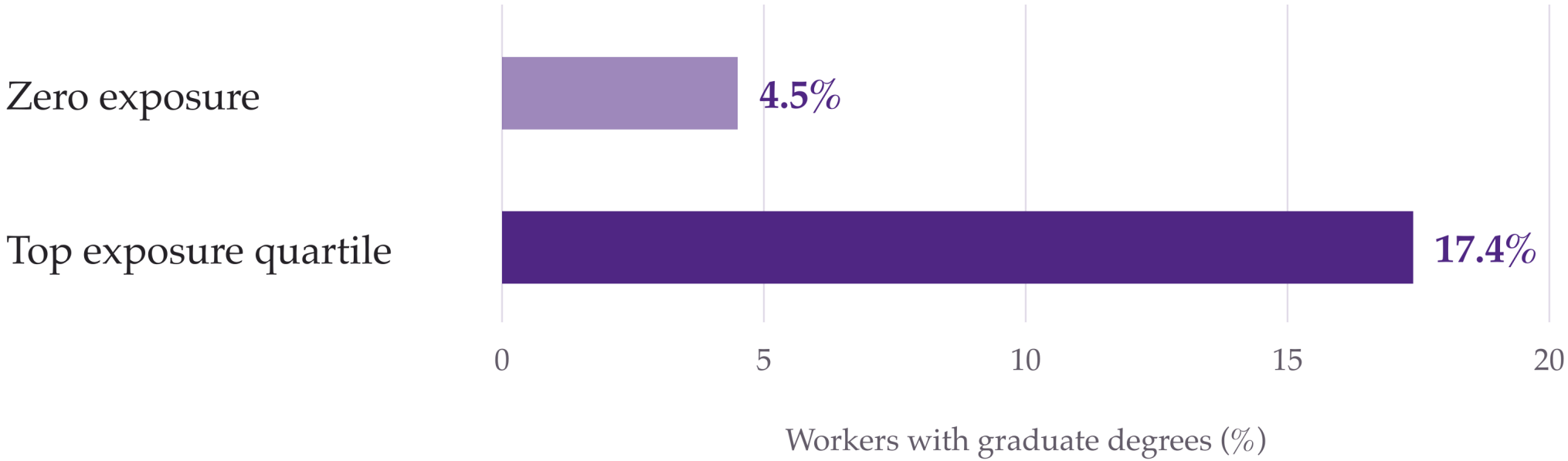
Observed exposure is concentrated



All 756 released occupations; each occupation receives equal weight.

> Zero exposure: 54.37% of occupations; 30% of workers in the paper.

Higher exposure, more graduate education



Worker groups in the CPS, August–October 2022; paper Figure 5.

> About **17 out of 100 highly exposed workers** have graduate degrees, compared with **4–5 out of 100 workers with zero exposure**.

Appendix: Importance, relevance, and frequency measure different things

Rating	Question it answers
Importance (1–5)	How important is this task to the occupation?
Relevance (%)	What share of respondents say the task belongs to their job?
Frequency categories	How often is the task performed?

Core: relevance ≥67% + mean importance ≥3. Otherwise: supplemental.

> Ratings do not measure time shares. “Core” ≠ “automatable.”

Appendix: From a profile to a usable research file

File	Keep these identifiers / fields
Occupation Data	Occupation code, title, description
Task Statements	Occupation code, Task ID, text, core/supplemental
Task Ratings	Task ID, scale, value, sample and precision information
Tasks to Detailed Work Activities (DWAs)	Task ID and DWA identifier

Use the **27.2 archive** to reproduce the exposure exercise. Record the release before joining to wages, employment, or AI labels.

Appendix: What the file contains

Field	Meaning
O*NET-SOC Code	Occupation identifier, retained as text
human_rating_beta	Intermediate exposure score from human labels
dv_rating_beta	Intermediate exposure score from GPT-4 labels
Task ID	Identifier for a task statement
Task Type	Core or supplemental

A row in the occupation file and a row in the task file have different units.

Appendix: Checking a merge before interpreting it

- Keep full occupation codes, including punctuation & suffixes
- Check the unit: detailed occupation or aggregate?
- Count matched, unmatched & duplicate identifiers
- Document crosswalks + score aggregation

> Match occupation codes + units. Similar names alone do not establish a match.

TaskNet: a worker-grounded map of tasks

The **Stanford Task Network (TaskNet)** records what workers do.

Task statements

AI-guided interviews: everyday + emerging duties.
Workers confirm the task lists.

Worker ratings

Relevance + importance.
Frequency, duration & training.
Can quality be specified + checked?

First release: **six occupations**. New task lists can change measured AI exposure.

How TaskNet builds an occupation's task list

1 · Discover

132 interviews

AI-guided task discovery.
Workers confirm or revise.

2 · Aggregate

2,726 reports

Similar reports ⟶
202 standardized tasks

3 · Validate

318 raters

O*NET analyst checks.
7,923 worker ratings

Six occupations: **79 tasks new to O*NET + 58 refined statements**.

Download TaskNet CSV · Browse tasks by occupation

Example: AI creates a new task for lawyers

Worker-reported AI-related duty

Check + revise AI legal drafts before filing:
accuracy, compliance & supporting authority.

Task inventory: Interviews reveal emerging verification work.

Exposure measure: Higher measured exposure for lawyers using TaskNet tasks.

Task lists shape exposure scores. Employment & wage effects need separate evidence.