we refer to V exposure – meaning all tasks directly exposed via tools like ChatGPT or the OpenAI Playground are considered twice as exposed as tasks requiring some complementary innovation.

3.4 Limitations of our methodology

3.4.1 Subjective human judgments

A fundamental limitation of our approach lies in the subjectivity of the labeling. In our study, we employ annotators who are familiar with LLM capabilities. However, this group is not occupationally diverse, potentially leading to biased judgments regarding LLMs' reliability and effectiveness in performing tasks within unfamiliar occupations. We acknowledge that obtaining high-quality labels for each task in an

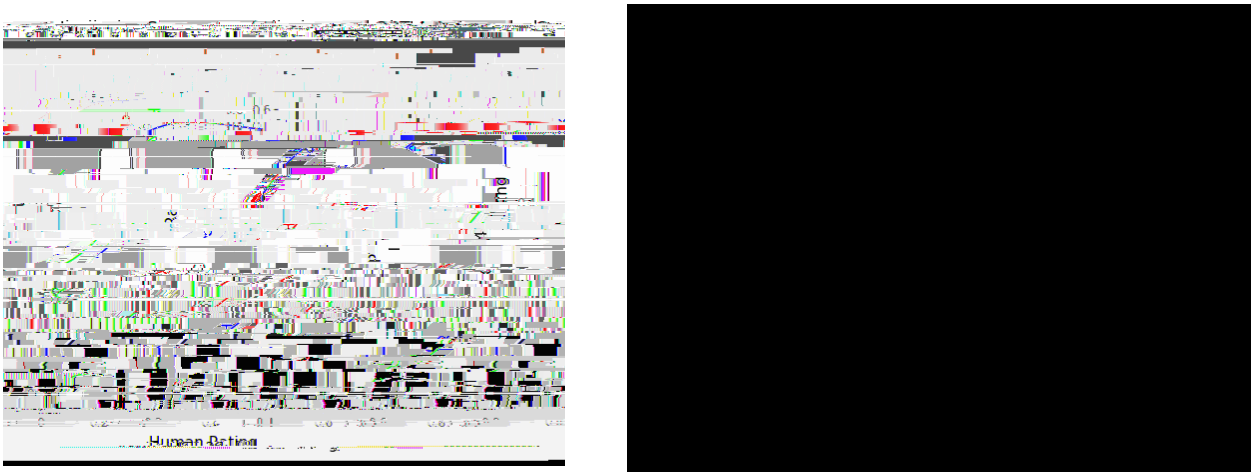

Figure 2: Human raters (x-axis) and GPT-4 ratings (y-axis) show a high degree of agreement about LLM

LLMs' potential impact on the labor market is limited since it does not consider total factor productivity or capital input potential. In addition to their influence on labor, LLMs may also influence these dimensions.

Occupation Level Exposure

Figure 4: The binscatter plots depict the exposure to language models (LLMs) in various occupations, as assessed by both human evaluators and GPT-4. These plots compare the exposure to LLM and partial LLM-powered software (V) at the occupation level against the log of total employment within an occupation and log of the median annual wage for occupations. While some discrepancies exist, both human and GPT-4 assessments indicate that higher wage occupations tend to be more exposed to LLMs. Additionally, numerous lower wage occupations demonstrate high exposure based on our rubric. Core tasks receive twice the weight of supplemental tasks within occupations when calculating average exposure scores. Employment and wage data are sourced from the BLS-OES survey conducted in May 2021. In aggregating tasks to the occupation-level, we assign half the weight to O*NET supplemental tasks as we do for core tasks. All weights within an occupation sum to one.

Figure 5: *V*

On The Job Training Required	Median Income	Tot Emp (000s)	H *U*	M *U*	H *V*	M *V*	H *Z*	M *Z*
None	$77,440	90,776	0.20	0.16	0.42	0.46		

	Min	25th Perc.	Median	75th Perc	Max	Mean	Std. Dev.	Count

	GPT-4 Exposure Rating 1		GPT-4 Exposure Rating 2		Human Exposure Rating	
	(1)	(2)	(3)	(4)	(5)	(6)
Software (Webb)	0•00113	0•00123	0•00111	0•00119	0•00096	0•00101
	[1]0•00031 s / F3	/ F26[illegible])	0003 1			

6 Discussion

6.1 GPTs as a General-Purpose Technology

Earlier in this paper we discuss the possibility that LLMs could be classified as a general-purpose technology. This classification requires LLMs to meet three core criteria: improvement over time, pervasiveness throughout the economy, and the ability to spawn complementary innovations (Lipsey et al., 2005). Evidence from the AI and machine learning literature thoroughly demonstrates that LLMs meet the first criteria – they are improving in capabilities over time with the ability to complete or be helpful for an increasingly complex set of tasks and

6.2 Implications for US Public Policy

The introduction of automation technologies, including LLMs, has previously been linked to heightened economic disparity and labor disruption, which may give rise to adverse downstream effects (Acemoglu and of LLMs on the workforce, policymakers and stakeholders can make more informed decisions to navigate the complex landscape of AI and its role in shaping the future of work.

7.1 LLM Conclusion (GPT-4's Version)

Generative Pre-trained Transformers (GPTs) generate profound transformations, garnering potential technological growth, permeating tasks, greatly impacting professions. This study probes GPTs' potential trajectories,

mean an exacerbation of Baumol's cost disease. In other words, if LLMs are likely to increase productivity differentially across industries, one concern is that the most productive would become even more productive. With inelastic demand for the production of those industries, the most productive sectors would shrink as a

Acemoglu, D. (2002). Technical change, inequality, and the labor market. Journal of Economic Literature, 40.

Baumol, W. J. (2012). The cost disease: Why computers get

Cheng, Z., Lee, D., and Tambe, P. (2022). Innovae: Generative ai for understanding patents and innovation.