



ТЕХНОЛОГІЧНІ ЗМІНИ, ІННОВАЦІЙНИЙ РОЗВИТОК І ЕКОНОМІКА ЗНАНЬ

TECHNOLOGICAL CHANGE, INNOVATION-DRIVEN DEVELOPMENT AND KNOWLEDGE ECONOMY

<https://doi.org/10.15407/economyukr.2026.07.085>

UDC 331.2:331.5:330.4

JEL: J31, J23, J24, O33, D14, C34

SEMKO Roman, PhD (Econ.),

Associate Professor of Department of Finance

National University of “Kyiv-Mohyla Academy”

2, Hryhoriia Skovorody St., 04070, Kyiv, Ukraine

e-mail: rsemko@ukma.edu.ua

ORCID: <https://orcid.org/0000-0002-3777-4064>

LARGE LANGUAGE MODEL EXPOSURE AND INDIVIDUAL INCOME IN WARTIME UKRAINE: EVIDENCE FROM A 2023—2024 HOUSEHOLD SURVEY

This study examines the relationship between household members' income and occupational exposure to Large Language Models (LLMs). Using a Heckman sample selection model on Ukrainian household survey data for 2023—2024, we find that higher occupational LLM exposure is positively correlated with higher earnings. The results suggest that LLM technology may boost productivity and may also affect wages as LLMs become more integrated into production, implying a need for government reskilling and upskilling programs.

Keywords: household members' income; employment; Large Language Models; task automation; Heckman selection model.

Russia's full-scale invasion in February 2022 has had a tremendous impact on the Ukrainian economy and its labor market. Millions of people were internally displaced or migrated as refugees abroad, which substantially affected labor market dynamics. For example, the number of open vacancies posted by the State Employment Service of Ukraine dropped by almost 50% in 2022 and returned to pre-full-

C i t a t i o n: Semko, R. (2026). Large Language Model exposure and individual income in wartime Ukraine: evidence from a 2023—2024 household survey. *Economy of Ukraine*. 69. 07(776). 85—100. <https://doi.org/10.15407/economyukr.2026.07.085>

© Publisher PH «Akadempriodyka» of the NAS of Ukraine, 2026. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

scale invasion level only in 2024¹. At the same time, Ukraine faces severe labor shortages due to the large migration and high demand for country rebuilding and defense. Wages increased gradually over the period², increasing production costs.

There are several strategies to address the shortage of workers needed for reconstruction and development. One of the approaches is automation. Specifically, the introduction of cognitive AI tools like Large Language Models (LLMs). In November 2022, OpenAI company released ChatGPT interface (based on Generative Pretrained Transformers model GPT-3.5) which broke into the market demonstrating superior reasoning capabilities and huge potential to fulfil complex cognitive tasks. One of the first papers to assess the impact of LLMs was T. Eloundou et al. (2024). The authors analyzed thousands of O*NET occupations³ and underlying tasks, and calculated their exposure to LLMs based on: (i) human expert classification and (ii) GPT-4 model classification. Three exposure metrics were calculated: α -exposure that measures the share of LLM-impacted tasks, β -exposure which includes LLM impact as well as half of the impact of building additional software above LLM, and γ -exposure that incorporates fully LLM and LLM-powered software. For example, γ -exposure provided by human experts for cashier occupation of 0.49 would mean that 49% of cashiers' tasks can be largely automated with the use of LLM and LLM-based software.

In order to assess whether such occupational exposure is associated with income in Ukraine, appropriate microdata is required. State Statistics Service of Ukraine conducted regular Household Budget Surveys (HBS) with the goal to assess based on the representative sample of households and its members their socio-economic and demographic status, reporting income and expenditures, available durable goods and services. However, with the Russia's full-scale aggression in February 2022, this and other major data collection procedures were paused until 2026. Notwithstanding this fact, Institute of Demography and Quality of Life Problems of the National Academy of Sciences of Ukraine and the Ukrainian Center for Social Reforms conducted a nationwide (except occupied territories) household socio-economic status survey⁴ during three-month period from December 2023 to February 2024. Similarly to HBS, it covers all key aspects of the households and their members, and this data will be used for the analysis.

The problem statement of this study is that despite the growing Ukrainian and international literature on labor market shortage and occupational exposure to LLMs, there is still limited empirical evidence on how such exposure is associated

¹ Number of vacancies and registered unemployed individuals. *State Employment Service of Ukraine*. URL: <https://www.dcz.gov.ua/stat/statfile> (accessed: 10.01.2026).

² Average monthly wages of regular employees by types of economic activity. *State Statistics Service of Ukraine*. URL: https://ukrstat.gov.ua/operativ/operativ2021/gdn/szpshp/arh_szp_ed_u.html (accessed: 10.01.2026).

³ O*NET 27.2 database (Version 27.2). *O*NET Resource Center*. URL: https://www.onetcenter.org/db_releases.html (accessed: 14.12.2025).

⁴ Socio-Economic Status of Households in Ukraine. *UNICEF Ukraine*. URL: <https://www.unicef.org/ukraine/en/documents/socio-economic-status-of-households> (accessed: 02.12.2025).

with individual income in Ukraine. Large Language Models offer substantial potential for the automation and augmentation of many cognitive tasks. They may have a large impact on employment levels and individual income and can at the same time pose challenges and provide new opportunities for the Ukrainian economy, whose labor market is under high pressure due to Russia's aggression, mass migration, and potential labor shortage during reconstruction.

In line with this problem statement, the objective of the research is to assess the association between LLM exposure and household member income in Ukraine. Namely, these main and alternative hypotheses will be tested:

H_0 : LLM exposure is not statistically significantly correlated with household member income,

H_1 : LLM exposure and individual income are correlated,

H_{1a} : the relationship is negative (automation hypothesis),

H_{1b} : the relationship is positive (augmentation hypothesis) or LLMs are not integrated yet.

The dynamics of the labor market of Ukraine has been actively studied during the wartime period by the Ukrainian scientists from different universities and institutes of the National Academy of Sciences of Ukraine. Usually, comprehensive analysis covers such problematic aspects as low income and inequality, labor shortage, rigidity and inefficiencies on the market, regional imbalances, and human capital depletion. One of the first papers which mentioned the detrimental impact of Russia's full-scale invasion on household income and quality of life was V. Heyets et al. (2022). The authors defined the number one priority as supporting employment and reasonable income levels, especially for the most vulnerable groups and internally displaced people. Comparing the results from 2021 HBS with 2023—2024 survey of socio-economic status of households, V. Shyshkin (2024) calculated that real household income was relatively stagnant, absolute poverty and inequality levels increased. V. Blyzniuk et al. (2025) highlight the problem of human capital deterioration and states that the key driver of recovery must be qualitative rather than quantitative development through mutually beneficial symbiosis of educational sector and labor market needs. We support above propositions, namely, by considering LLMs as a powerful tool that can potentially improve productivity, lower the inequality, and increase the quality of life of the households.

When speaking about LLMs specifically, there are two main groups of papers on the impact of LLMs (or AI broadly) on economic systems. The first focuses on assessments of LLM exposure of different occupations and industries, while the second analyzes its effect on salaries and employment dynamics. The methods used vary from classical impact analysis approaches such as difference-in-difference (DiD) to classical regression models.

K. Handa et al. (2025) mapped four million conversations (prompts) from Anthropic's Claude.ai model to O*NET's occupational framework using tree-based structure where tasks are grouped in broad categories and then classified into specific occupations. The next occupations show the highest association with AI (in decreasing order): Mathematics and Computer Science, Arts, Sports, and Media, Education,

Business and Finance. On the other hand, physical labor-intensive jobs are the least related to LLMs, for example Farming, Fishing, and Forestry. In terms of labor-augmentation vs automation, they were approximately equally spread across the sample.

D. Chen et al. (2025) used earnings and unemployment data over the 2010–2025 period to estimate the effect of LLM introduction on these indicators. Authors built on K. Handa et al. (2025) automation/augmentation findings. The key method utilized is a (synthetic) DiD approach that allows to isolate LLM impact from other trends in panel data. It appears that the changes in employment rates are not statistically significant at least during short-term period, while weekly wages are increasing by 89 USD measured in 2010 prices for the occupations exposed to LLM. It means that LLMs are rather increasing productivity for occupations sensitive to them than replacing the tasks or jobs.

Completely different results were reported by B. Klein Teeselink (2025) for the UK labor market. They used unique dataset collected by the company “Revelio Labs” over 2021–2025 that includes millions of work profiles and job postings from professional networking firms, company websites, and online job posting platforms. In particular, LinkedIn employee profiles were used to count the number of unique people per company before and after ChatGPT introduction in November 2022. Based on DiD method, it was found that companies exposed to LLM technology reduced employment, especially junior positions were affected. Similarly, the number of vacancies dropped for technical and analytical jobs, while it did not change for the ones requiring human interaction and relationship building skills. High-salary jobs were also the most affected in terms of new role posting: both the number of vacancies and compensation packages dropped if the tasks were highly exposed to LLM automation. This effect was not observed for low-income jobs. It took time to adjust to ChatGPT breakthrough — results became statistically significant after 10–18 months.

Different methodological approach was utilized by E. Engberg et al. (2025) who analyzed three waves of surveys (2006, 2012, and 2018) in Germany covering 20,000 employees working at least 10 hours per week. They study the impact of AI and robotic technology on the labor market using AI exposure dynamic measure from E. Engberg et al. (2024) and robotic index from M. Webb (2020). Based on Mincer wage regression, AI exposure is positively associated with wage increases. Both AI and robots shift the content of the tasks performed by workers suggesting vulnerable occupations may require reskilling and upskilling.

Less evidence is available for developing economies. Q. Chen et al. (2025) started with the data collection from the key Chinese online vacancy portals. Using similar methodology as T. Eloundou et al. (2024), they split occupations on underlying typical tasks and classified them based on their potential for automation with the help of OpenAI GPT-4 and local LLM models. Authors rejected routinization hypothesis which states that technological progress is automating routine tasks and does not affect more cognitive jobs. High LLM exposure is associated with higher salaries and poses significant risk of job displacement in the technological and related sectors. Nevertheless, based on the proposed theoretical model, there might

be a positive effect for low-skilled workers. Equipped with LLMs, they can perform more tasks that are now done by lawyers, programmers, or even college professors which potentially will reduce wage gap in the future.

This article presents novel results in several important aspects. **First**, instead of using announced salaries from vacancy postings that are not necessarily realized ones, we rely on actually reported wages (or income broadly) from the household members. **Second**, from methodological point of view, DiD and wage-only regression models may suffer from sample selection bias when unemployed individuals' salaries are not observed. This drawback is expected to be addressed with the application of Heckman sample selection model (Heckit).

Therefore, **the purpose of this article** is to test whether household members whose occupations have high exposure to LLMs have higher income.

Research methods. The study combines occupational mapping, descriptive statistical analysis, and regression modelling. Occupations reported in the survey are matched with O*NET data. Mapping is conducted with the help of LLM models as well as with human manual review. Finally, Heckit sample selection model is utilized to estimate the association between income and LLM exposure.

The sample of individual household members was restricted to people aged 18–65 which resulted in 11,640 records or 62% of total population. A small number of people who were working abroad were also removed from the sample since foreign salaries may significantly differ from the local ones biasing the analysis. Then, only those who belong to the labor force were identified: (i) 6,737 members employed, or running the business, or involved in any other income-generating activity plus, (ii) those 600 who are unemployed or working on their own family farm (unpaid) but are actively searching for the job. Next, if the person reported less than 10 hours per week and monthly income was less than 6,000 UAH or hourly salary was lower than state-defined minimum wage, it was also set to missing because it is likely inaccurate (given minimum wage was 6,700 and 7,100 UAH in 2023 and early 2024, respectively). There are 196 records where respondents refused to provide their occupations or gave general answers such as self-employed without description of the activity. Also, 213 observations represent people serving for the military. Since O*NET list is based only on civilian jobs and the vast majority of cases provided no specific specialization like military driver or infantry was provided — these observations were excluded from the sample. As a result, there are 4,995 individuals with reported income and 600 unemployed people that will be used in the regression model.

Taking into account survey representative weights, 70% of individuals from the final sample were living in urban settlements and the remaining 30% in rural areas. 54% were male and 46% female with average age of 42 years old which signals men more often belong to the labor force. The largest number of respondents were from Kyiv City, Dnipro, Kyiv, Lviv, Odesa, and Kharkiv regions. The smallest shares were from Kherson, Donetsk, Mykolaiv, and Zaporizhzhia regions which are largely occupied by aggressor areas and/or are close to the frontline. 71% of respondents were married (officially registered or cohabiting), 11% were divorced, 15% were never

married, and 3% were widowed. Individuals' household size varied from 1 to 13 with the average number of members almost 3. There were only 0.7 children on average with more than half of individuals living in households without children.

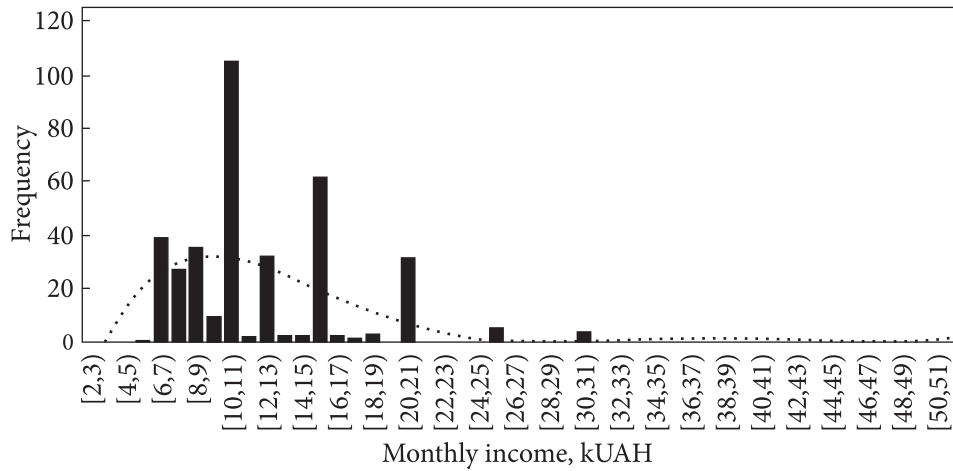
Higher or incomplete higher education was reported by 53% of respondents: 22% achieved Master's degree, 24% Bachelor's degree, 7% had short-cycle (incomplete or associate) higher education. There were 0.6% of respondents with academic degrees (PhD, candidate or doctor of sciences). The largest share was vocational education or training programs — 33%. Upper (complete general) or lower secondary education level was reported by 14% of individuals. Slightly more than 1% of people had primary or no education. 53% of individuals provided fields of study (usually those with higher education) including Engineering/Technical sciences (14%), Economics (13%), Education/Pedagogy (6%), Humanities (5%), Medicine (4%), and Law (4%).

In terms of employment, paid job with formal documentation (labor contract) accounted for 69.75% of household members, without formal documents (verbal agreement) — 11.86%, business owners or private entrepreneur (including FOP) were 5.73%, self-employed (unregistered, no employees) were 3.22%. Duration of the last employment varied from 0.1 to 44.4 years with an average of 8.4 years (median equals to 6.0).

The largest share of respondents — 14% — were employed in wholesale and retail trade as well as repair of motor vehicles and motorcycles sector. Industrial/manufacturing sector share is 10%. Then comes transportation and storage, education, construction, health, and agriculture each with a 6—9% share.

Technicians and associate professionals had almost 14% share. Professionals such as researchers, mathematicians, physicians, engineers, and so on constituted also 14%. 13% of people held elementary occupations. Service and sales workers (shop assistants, cashiers) share was 11%, craft and related trades workers — 9%, managers (enterprise, institution, organization, and unit managers) — 7%, clerical support workers (clerks, secretaries) — 6%, plant and machine operators and assemblers — 5%, skilled agricultural, forestry and fishery workers constituted only 3% of the sample.

Interviewed household head provided occupations of its members. This field is not standardized and contains 2,440 distinct responses. There are numerous typos, omissions, or errors. We have mapped these responses to 923 occupations from the U.S. economy — 2018 Standard Occupations Codes O*NET-SOC. In order to minimize human errors during mapping, OpenAI GPT-5.2 Thinking LLM was used to translate and map provided occupations from Ukrainian to English O*NET ones. It was done in multiple chunks (several hundreds of records per iteration) in order to stay in the model's context window using prompts requesting the translation and matching of Ukrainian occupations to O*NET occupations. Then, each pair was manually reviewed and corrected where necessary. During correction, another model was employed to get the best consensus suggestion — Google Gemini 3 Pro which was asked to highlight any potential inconsistencies. As a result, more than 10% of mappings were adjusted or corrected. Then, the final mapping



Household members' reported monthly income, kUAH. Dotted curve shows a six-degree polynomial trend line fitted to the monthly income distribution

Source: calculated and compiled by the author.

Table 1. Household member occupations mapped to O*NET (top frequencies)

O*NET Occupations	Share, %
Retail Salespersons	8.38
Light Truck Drivers	5.51
General and Operations Managers	5.15
Secondary School Teachers, Except Special and Career / Technical Education	4.06
Accountants and Auditors	2.71
Construction Laborers	2.67
Automotive Service Technicians and Mechanics	2.54
Security Guards	2.14
Registered Nurses	1.92
Industrial Engineers	1.92
Laborers and Freight, Stock, and Material Movers, Hand	1.85
Helpers-Production Workers	1.72
Maintenance and Repair Workers, General	1.55
Software Developers	1.53

Source: calculated by the author.

was provided for the analysis by above two models to highlight any remaining mistakes and they have been addressed by the author through manual review and updates. The results of such mapping are 387 distinct occupations with the top frequencies provided in Table 1.

Also, T. Eloundou et al. (2024) calculated LLM exposure metrics for university professors with high granularity, however, in the sample, people usually did not

provide specialization. Consequently, average across all disciplines value was used for university teaching staff.

The average reported household members' income equals 15,079 UAH with minimum of 2,000 for those who worked 10 hours a week and maximum of 150,000. As can be seen from Figure, the numbers are highly skewed and follow Log-Normal distribution which is typical for this type of data. We can also see clear behavioral pattern: heaping — when respondents are using cognitive shortcut and report the income near round numbers like 10,000 or 20,000. Almost 1% of respondents have reported income above 51,000.

Correlation between individual members' income (salary and entrepreneurial) and the income of other household members is low and constitutes 12%. Expectedly, average household income is larger than individuals' and equals to 26,256 with maximum value of 325,000 UAH.

Correlation between household members' reported income and exposure to α -, β -, and γ -based automation is 0%, 15%, and 21% for human expertise classification and 16%, 25%, and 23% for GPT-4 classification, respectively.

METHODOLOGY

The standard Heckman sample selection model will be used to estimate the correlation between LLM exposure and individual income. The model itself can be derived starting from the reservation wage — the minimum amount at which the person agrees to accept the job opening (see M. Verbeek (2004) for more details):

$$w_i^r = z_i' \gamma + \mu_i, \quad (1)$$

where z_i' is the transposed vector of personal utility preferences, γ is the vector of parameters, and μ_i represents the error term capturing the effect of other variables not included in the equation. Next, the person is offered a wage defined as follows:

$$w_i^* = x_i^{1'} \beta^1 + e_i^1, \quad (2)$$

where $x_i^{1'}$ is the transposed vector of individual and job characteristics, β^1 is the vector of parameter, and e_i^1 is the error term. The offered wage w_i^* will be accepted if it is larger or equal than respective reservation wage w_i^r . We can define wage differential as a selection equation

$$h_i^* = w_i^* - w_i^r = x_i^{1'} \beta^1 - z_i' \gamma + e_i^1 - \mu_i = x_i^{2'} \beta^2 + e_i^2 \quad (3)$$

with $x_i^{2'}$ stacking $x_i^{1'}$ and z_i' variables, β^2 consisting of β^1 and γ parameters, and $e_i^2 = e_i^1 - \mu_i$.

Now, we can say that the wage is observed $w_i = w_i^*$ and the person accepts job offer (is employed: $h_i = 1$) if $h_i^* \geq 0$. Otherwise, actual wage w_i is not observed and the person is unemployed ($h_i = 0$) when $h_i^* < 0$.

Probability of being employed would be equal to

$$P(h_i = 1) = P(h_i^* \geq 0) = P(x_i^{2'} \beta^2 + e_i^2 \geq 0) = P(-e_i^2 \leq x_i^{2'} \beta^2) = \Phi(x_i^{2'} \beta^2), \quad (4)$$

where $\Phi(x)$ is standard Normal cumulative distribution function (CDF) assuming $e_i^2 \sim N(0,1)$.

Based on above derivation and adding Heckman's (1979) lambda (inverse Mill's ratio) to the wage equation, the following model will be estimated for expected individual income:

$$E(w_i | h_i = 1) = x_i' \beta^1 + \sigma_{1,2} \frac{\phi(x_i^{2'} \beta^2)}{\Phi(x_i^{2'} \beta^2)}, \quad (5)$$

$$P(h_i = 1) = \Phi(x_i^{2'} \beta^2), \quad (6)$$

where $\lambda(x_i^{2'} \beta^2) = \frac{\phi(x_i^{2'} \beta^2)}{\Phi(x_i^{2'} \beta^2)}$ is Heckman's lambda, $\phi(x)$ and $\Phi(x)$ are standard Normal density and cumulative distribution functions, respectively. Error terms follow

bivariate Normal distribution $\begin{pmatrix} e_i^1 \\ e_i^2 \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & 1 \end{pmatrix} \right)$. Wage equation is specified in the log-level form. The model will be estimated using maximum likelihood method with survey weights and based on the *R* package sample Selection developed by O. Toomet & A. Henningsen (2008).

RESULTS AND DISCUSSION

We estimated six regression specifications for α -, β -, and γ -exposure measured by human experts and GPT-4 model. Since equations are estimated jointly, coefficients vary both in the wage (5) and selection (6) equations even though the former does not include LLM exposure variable.

It is assumed that age, household head status, marriage, number of children, non-labor income, type of settlement, sex, macro region, education level, house ownership affect employment outcome. Identification of the Heckman sample selection model is based on exclusion restrictions. In our case they are age, age squared, respondent status, marital status, number of children, other income, and housing status. These factors are expected to influence labor market participation through reservation wage but do not impact the income directly. Similar factors are usually present in other studies as highlighted by F. Blau et al. (2024) although we admit that some of them may be weak in selection.

As can be seen from Table 2, age, marriage, and the sum of non-labor income do not have statistically significant effect on employment. The number of children is significant at least at 90% confidence level with more children present reducing the probability of employment. On the contrary, being a household head increases employment and is statistically significant for β - and γ -exposure specifications. Living in rural settlement (in comparison to the urban), or in any macro region except Kyiv City, or being a female reduces the probability of employment. Comparing the level of education with the reference category (primary or lower level), we can see that both higher and vocational education are positively impacting employment, while secondary education does not have statistically significant effect. Interestingly, renting a house or apartment decreases the probability of employment in comparison to the reference situation of owning it.

Analyzing the wage equation in Table 3, we can see that when duration of the individual employment (experience) on the current job increases, the wage also in-

Table 2. Sample selection equation: Employment

Coefficients	Human expertise			GPT-4 classifications		
	α	β	γ	α	β	γ
Constant	1.514 ***	1.460 ***	1.460 ***	1.497 ***	1.474 ***	1.483 ***
Age	0.011	0.011	0.011	0.012	0.011	0.011
Age ² /100	-0.004	-0.006	-0.006	-0.006	-0.006	-0.005
Household head	0.070	0.078 *	0.078 *	0.070	0.080 *	0.083 *
Married	0.012	0.018	0.025	0.012	0.031	0.036
Number of children	-0.042 *	-0.047 *	-0.048 *	-0.041 *	-0.049 *	-0.051 **
Sum of non-labor income	0.000	0.000	0.000	0.000	0.000	0.000
Rural settlement	-0.266 ***	-0.260 ***	-0.258 ***	-0.263 ***	-0.258 ***	-0.260 ***
Sex (male as a reference)	-0.522 ***	-0.523 ***	-0.520 ***	-0.519 ***	-0.513 ***	-0.515 ***
Central region	-1.110 ***	-1.095 ***	-1.093 ***	-1.111 ***	-1.094 ***	-1.086 ***
Eastern region	-1.256 ***	-1.244 ***	-1.243 ***	-1.254 ***	-1.240 ***	-1.238 ***
North region	-0.933 ***	-0.927 ***	-0.931 ***	-0.933 ***	-0.932 ***	-0.930 ***
South region	-0.809 ***	-0.788 ***	-0.791 ***	-0.810 ***	-0.795 ***	-0.788 ***
Western region	-1.035 ***	-1.028 ***	-1.031 ***	-1.040 ***	-1.038 ***	-1.029 ***
Education — high	1.070 ***	1.106 ***	1.100 ***	1.070 ***	1.090 ***	1.095 ***
Education — secondary	0.286	0.317	0.314	0.283	0.304	0.311
Education — vocational	0.626 **	0.660 **	0.657 **	0.626 **	0.650 **	0.655 **
House tenure — other	-0.388 **	-0.413 **	-0.430 **	-0.384 **	-0.433 **	-0.449 ***
House tenure — rented	-0.438 ***	-0.448 ***	-0.454 ***	-0.439 ***	-0.454 ***	-0.457 ***
House tenure — social	-0.037	-0.037	-0.061	-0.032	-0.057	-0.069

Notes: *** denotes significant coefficients at 99% confidence level, ** at 95%, and * at 90% confidence level.

Source: estimated by the author.

Table 3. Log-level wage equation

Coefficients	Human expertise			GPT-4 classifications		
	α	β	γ	α	β	γ
Constant	9.618***	9.545***	9.520***	9.597***	9.517***	9.518***
Experience	0.019***	0.019***	0.018***	0.019***	0.018***	0.017***
Experience ² /100	-0.052***	-0.051***	-0.049***	-0.052***	-0.048***	-0.047***
Rural settlement	-0.169***	-0.166***	-0.164***	-0.168***	-0.160***	-0.159***
Sex (male as a reference)	-0.307***	-0.324***	-0.322***	-0.304***	-0.321***	-0.330***
Central region	-0.602***	-0.588***	-0.584***	-0.604***	-0.587***	-0.581***
Eastern region	-0.579***	-0.572***	-0.573***	-0.579***	-0.575***	-0.574***
North region	-0.473***	-0.467***	-0.465***	-0.470***	-0.463***	-0.464***
South region	-0.407***	-0.401***	-0.399***	-0.403***	-0.395***	-0.396***
Western region	-0.475***	-0.467***	-0.467***	-0.473***	-0.465***	-0.465***
Education — high	0.451***	0.410***	0.396***	0.447***	0.386***	0.375***
Education — secondary	0.042	0.040	0.042	0.041	0.039	0.040
Education — vocational	0.182*	0.174*	0.173*	0.179*	0.168*	0.168*
Sector — accommodation and food services	-0.041	-0.022	-0.015	-0.044	-0.019	-0.009
Sector — administrative and support services	-0.080**	-0.079**	-0.082***	-0.092***	-0.087***	-0.071**
Sector — agriculture	-0.086***	-0.057**	-0.056*	-0.079***	-0.054*	-0.054*
Sector — arts, sports, entertainment, and recreation	-0.053	-0.031	-0.032	-0.055	-0.040	-0.032
Sector — construction	0.006	0.050*	0.058**	0.018	0.058**	0.057**

The end of Table 3

Coefficients	Human expertise			GPT-4 classifications		
	α	β	γ	α	β	γ
Sector — education	-0.195***	-0.183***	-0.188***	-0.194***	-0.174***	-0.169***
Sector — financial and insurance activities	0.121**	0.095**	0.073*	0.119***	0.076*	0.069
Sector — human health and social work	-0.075***	-0.069**	-0.071***	-0.071**	-0.062**	-0.063**
Sector — industry/manufacturing	0.011	0.034	0.033	0.012	0.033	0.038
Sector — information and telecommunications	0.094***	0.054	0.030	0.037	-0.005	0.048
Sector — other activities	-0.075***	-0.049**	-0.048**	-0.076***	-0.045**	-0.034
Sector — professional, scientific, and technical activities	-0.130**	-0.145**	-0.162***	-0.161***	-0.179***	-0.148***
Sector — public administration and defense	0.062	0.061	0.056	0.049	0.063	0.081**
Sector — real estate activities	0.005	-0.018	-0.030	0.006	-0.031	-0.042
Sector — transportation and storage	-0.030	-0.008	-0.019	-0.039	-0.019	-0.003
LLM exposure	-0.013	0.304***	0.282***	0.196***	0.421***	0.262***
σ_1	0.487***	0.484***	0.481***	0.487***	0.479***	0.478***
$\rho_{1,2}$	0.893***	0.894***	0.889***	0.896***	0.886***	0.881***

Notes: *** denotes significant coefficients at 99% confidence level, ** at 95%, and * at 90% confidence level; $\rho_{1,2}$ is the correlation between ϵ_{i1} and ϵ_{i2} error terms.

Source: estimated by the author.

creases but at decreasing rate. People living in rural settlements earn 15% lower wages (effect of coefficient β_j^1 on the wage is calculated as $e^{\beta_j^1} - 1$). Conditional gender wage gap is also significant and amounts to 28%. The highest wages are in Kyiv City (reference category) followed by South and North macro regions. The lowest incomes are in Central (lower level of industrialization) and in Eastern (likely affected by the war) macro regions. In comparison to the household members with primary or no education, the presence of vocational and higher education increases average salary at least by 18% and 46%, respectively, with robust effect across all specifications.

When comparing wage distribution by sectors of economy relative to wholesale and retail trade plus repair of motor vehicles and motorcycles, the lowest household members' income is in education. Also lower statistically significant values are observed in administrative and support services, agriculture, health, and other activities. Larger compensation is present in selected specifications for construction, finance, and information and telecommunication. Other sectors do not have statistically different income.

LLM exposure is positively correlated with the wage level in all cases except for α -exposure measured by human experts. On average, if LLM-based automation increases by 10%, wage increases by 2.0% for GPT-4 α -exposure, by 3.1% and 4.3% for β -exposure, by 2.9% and 2.7% for γ -exposure for human and GPT-4, respectively.

Based on the findings, the number of children, location of the household, gender, education level, and house ownership significantly impact the probability of employment. Experience, location, education, and LLM exposure are also significantly correlated with the income level. Positive association between LLM exposure and household members' earnings is consistent with the results received by Q. Chen et al. (2025) but is the opposite to B. Klein Teeselink (2025).

CONCLUSION

In this article, socio-economic survey of household income and expenditures conducted in late 2023 and early 2024 in Ukraine has been used for the analysis of association of LLM exposure with reported individual income. Reported occupations were mapped to the O*NET categories for which LLM exposure was defined by human experts and GPT-4 model. Heckit sample selection model was used to address selection bias when individuals are not employed and their income is not observed. Positive correlation with LLM exposure was found in the majority of cases which rejects the hypothesis of routinization. While it may seem to support augmentation idea when new technology increases the productivity of labor leading to higher wages, automation alternative is also possible. LLM impact is the largest on cognitive, creative, and technological industries which are usually characterized by above average earnings. LLM may not have been integrated largely into business processes as of the survey date which would imply its effect is delayed into the future. If that is the case, it can have detrimental impact on wages and employment, and the government should prepare reskilling and upskilling programs to mitigate negative shock on the economy and people's welfare.

The distinction between augmentation and automation can be assessed more accurately based on the panel data. Specifically, the data from HBS planned to be conducted during 2026 can be used for future research to calculate the impact of LLMs on individual income dynamically. In addition, taking into account the importance of military occupations, further studies can be also dedicated to the impact of LLMs on the military sector where automation will play a crucial role in the near future.

Funding. *The article was prepared on our own initiative, on a free basis.*

Conflict of interest. *There are no financial or personal conflicts of interest that could have influenced the writing of this article.*

Involvement of AI. *Generative AI tools (OpenAI GPT-5.2 and Google Gemini 3 Pro) were used as auxiliary tools for occupational mapping, consistency checks, and improving the clarity and readability of the text. The author reviewed, validated, and edited all AI-generated outputs. The scientific content, methodology, statistical analysis, interpretation of results, and conclusions remain entirely the responsibility of the author.*

REFERENCES

- Eloundou, T., Manning, S., Mishkin, P., Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*. Vol. 384. P. 1306—1308. <https://doi.org/10.1126/science.adj0998>
- Heyets, V., Blyzniuk, V., Nykyforuk, O. (2022). Topical aspects of social quality policy in the post-conflict economy of Ukraine. *Economy of Ukraine*. 65. 6(727). 3—22. <https://doi.org/10.15407/economyukr.2022.06.003> [in Ukrainian].
- Shyshkin, V. (2024). Changes in the level and structure of household income in Ukraine under the influence of a full-scale war. *Economy and Society*. No. 62. <https://doi.org/10.32782/2524-0072/2024-62-166> [in Ukrainian].
- Blyzniuk, V., Burlai, T., Huk, L. (2025). The state and possibilities of human capital development in Ukraine in the military and recovery context. *Economic theory*. No. 3. P. 31—68. <https://doi.org/10.15407/etet2025.03.031> [in Ukrainian].
- Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., Mueller J. et al. (2025). Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations. *arXiv preprint*. arXiv: 2503.04761. P. 1—38. <https://doi.org/10.48550/arXiv.2503.04761>
- Chen, D., Kane, C., Kozlowski, A., Kunievsky, N., Evans, J. (2025). The (Short-Term) Effects of Large Language Models on Unemployment and Earnings. *arXiv preprint*. arXiv:2509.15510. P. 1—22. <https://doi.org/10.48550/arXiv.2509.15510>
- Klein Teeselink, B. (2025). Generative AI and Labor Market Outcomes: Evidence from the United Kingdom. *SSRN Working Paper*. P. 1—46. <https://doi.org/10.2139/ssrn.5516798>
- Engberg, E., Koch, M., Lodefalk, M., Schroeder, S. (2025). Artificial intelligence, tasks, skills, and wages: Worker-level evidence from Germany. *Research Policy*. Vol. 54. Iss. 8. P. 1—14. <https://doi.org/10.1016/j.respol.2025.105285>
- Engberg, E., Görg, H., Lodefalk, M., Javed, F., Långkvist, M., Monteiro, N., Kyvik-Nordas, H. et al. (2024). AI unboxed and jobs: a novel measure and firm-level evidence from three countries. *IZA Discussion Paper*. No. 16717. P. 1—47. URL: <https://ssrn.com/abstract=4691172>
- Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. *SSRN Electronic Journal*. P. 1—61. <https://doi.org/10.2139/ssrn.3482150>

- Chen, Q., Ge, J., Xie, H., Xu, X., Yang, Y. (2025). Large language models at work in China's labor market. *China Economic Review*. Vol. 92. P. 1—18. <https://doi.org/10.1016/j.chieco.2025.102413>
- Verbeek, M. (2004). A Guide to Modern Econometrics. 2nd ed. Erasmus University Rotterdam, John Wiley & Sons Ltd., Hoboken. P. 228—232.
- Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*. Vol. 47. Iss. 1. P. 153—161. <https://doi.org/10.2307/1912352>
- Toomet, O., Henningsen, A. (2008). Sample Selection Models in R: Package sampleSelection. *Journal of Statistical Software*. Vol. 27. Iss. 1. P. 1—23. <https://doi.org/10.18637/jss.v027.i07>
- Blau, F., Kahn, L., Boboshko, N., Comey, M. (2024). The Impact of Selection into the Labor Force on the Gender Wage Gap. *Journal of Labor Economics*. Vol. 42. Iss. 4. P. 1093—1133. <https://doi.org/10.1086/725032>

Received on February 19, 2026

Reviewed on March 13, 2026

Revised on March 26, 2026

Signed for printing on April 02, 2026

Published on July 17, 2026

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- Eloundou, T., Manning, S., Mishkin, P., Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*. Vol. 384. P. 1306—1308. <https://doi.org/10.1126/science.adj0998>
- Геєць, В., Близнюк, В., Никифору, О. (2022). Актуальні аспекти політики соціальної якості в постконфліктній економіці України. *Економіка України*. 65. 6(727). 3—22. <https://doi.org/10.15407/economyukr.2022.06.003>
- Шишкін, В. (2024). Зміни рівня та структури доходів домогосподарств України під впливом повномасштабної війни. *Економіка та суспільство*. № 62. <https://doi.org/10.32782/2524-0072/2024-62-166>
- Близнюк, В., Бурлай, Т., Гук, Л. (2025). Стан і можливості розвитку людського капіталу України у мілітарному та відбудовчому контексті. *Економічна теорія*. № 3. С. 31—68. <https://doi.org/10.15407/etet2025.03.031>
- Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., Mueller, J. et al. (2025). Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations. *arXiv preprint*. arXiv:2503.04761. P. 1—38. <https://doi.org/10.48550/arXiv.2503.04761>
- Chen, D., Kane, C., Kozlowski, A., Kunievsky, N., Evans, J. (2025). The (Short-Term) Effects of Large Language Models on Unemployment and Earnings. *arXiv preprint*. arXiv:2509.15510. P. 1—22. <https://doi.org/10.48550/arXiv.2509.15510>
- Klein Teeselink, B. (2025). Generative AI and Labor Market Outcomes: Evidence from the United Kingdom. *SSRN Working Paper*. P. 1—46. <https://doi.org/10.2139/ssrn.5516798>
- Engberg, E., Koch, M., Lodefalk, M., Schroeder, S. (2025). Artificial intelligence, tasks, skills, and wages: Worker-level evidence from Germany. *Research Policy*. Vol. 54. Iss. 8. P. 1—14. <https://doi.org/10.1016/j.respol.2025.105285>
- Engberg, E., Görg, H., Lodefalk, M., Javed, F., Långkvist, M., Monteiro, N., Kyvik-Nordas, H. et al. (2024). AI unboxed and jobs: a novel measure and firm-level evidence from three countries. *IZA Discussion Paper*. No. 16717. P. 1—47. URL: <https://ssrn.com/abstract=4691172>
- Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. *SSRN Electronic Journal*. P. 1—61. <https://doi.org/10.2139/ssrn.3482150>

- Chen, Q., Ge, J., Xie, H., Xu, X., Yang, Y. (2025). Large language models at work in China's labor market. *China Economic Review*. Vol. 92. P. 1—18. <https://doi.org/10.1016/j.chieco.2025.102413>
- Verbeek, M. (2004). *A Guide to Modern Econometrics*. 2nd ed. Erasmus University Rotterdam, John Wiley & Sons Ltd., Hoboken. P. 228—232.
- Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*. Vol. 47. Iss. 1. P. 153—161. <https://doi.org/10.2307/1912352>
- Toomet, O., Henningsen, A. (2008). Sample Selection Models in R: Package sampleSelection. *Journal of Statistical Software*. Vol. 27. Iss. 1. P. 1—23. <https://doi.org/10.18637/jss.v027.i07>
- Blau, F., Kahn, L., Boboshko, N., Comey, M. (2024). The Impact of Selection into the Labor Force on the Gender Wage Gap. *Journal of Labor Economics*. Vol. 42. Iss. 4. P. 1093—1133. <https://doi.org/10.1086/725032>

Надійшла 19.02.2026

Прорецензована 13.03.2026

Доопрацьована 26.03.2026

Підписана до друку 02.04.2026

Оприлюднена 17.07.2026

Роман Семко, канд. екон. наук, доцент кафедри фінансів
Національний університет «Києво-Могилянська академія»
вул. Григорія Сковороди, 2, 04070, Київ, Україна

ВПЛИВ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ ТА ІНДИВІДУАЛЬНІ
ДОХОДИ В УКРАЇНІ В УМОВАХ ВІЙНИ (ЗА ДАНИМИ
ОПИТУВАННЯ ДОМОГОСПОДАРСТВ У 2023—2024 рр.)

Великі мовні моделі (ВММ) мають значний потенціал для автоматизації і підсилення багатьох когнітивних завдань. Вони здатні істотно впливати на рівень зайнятості та індивідуальні доходи, водночас створюючи виклики й відкриваючи нові можливості для української економіки, ринок праці якої перебуває під значним тиском через агресію РФ, масову міграцію і майбутній дефіцит робочої сили в період відбудови. Використовуючи існуючі дослідження, здійснено порівняння професій, указаних у соціально-економічному опитуванні домогосподарств України у 2023—2024 рр., з професіями США, для яких оцінено потенціал автоматизації за допомогою ВММ і побудованих на їх основі програмних продуктів. Для усунення потенційного зміщення вибірки, що виникає через неможливість спостерігати доходи непрацевлаштованих осіб, застосовано регресійну модель відбору Хекмана (Heckit), яка спочатку моделює ймовірність зайнятості, а на другому кроці визначає ключові чинники, пов'язані з індивідуальним доходом. Встановлено, що дохід членів домогосподарств статистично значущо корелює з рівнем впливу ВММ на виконання професійних завдань. Отримані результати можуть свідчити про те, що технології ВММ у майбутньому здатні підвищувати продуктивність і радше доповнювати, ніж автоматизувувати (замінювати) робочі місця. Водночас результати можна інтерпретувати інакше: у разі широкої інтеграції ВММ у бізнес-процеси в майбутньому це може впливати на доходи багатьох працівників, що вимагає розроблення державних програм перекваліфікації і підвищення кваліфікації, щоб запобігти негативним соціально-економічним наслідкам на ринку праці.

Ключові слова: доходи членів домогосподарств; зайнятість; великі мовні моделі; автоматизація завдань; модель відбору Хекмана.