ECONSTOR
Make Your Publications Visible.

A Service of

ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Heikkilä, Jussi T. S.

Article — Published Version

Human intelligence versus artificial intelligence in classifying economics research articles: exploratory evidence

Journal of Documentation

Suggested Citation: Heikkilä, Jussi T. S. (2024) : Human intelligence versus artificial intelligence in classifying economics research articles: exploratory evidence, Journal of Documentation, ISSN 1758-7379, Emerald, Bingley, Vol. 81, Iss. 7, pp. 18-30, https://doi.org/10.1108/JD-05-2024-0104

This Version is available at:
https://hdl.handle.net/10419/307995

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by/4.0/legalcode

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0022-0418.htm

Received 8 May 2024
Revised 20 October 2024
Accepted 24 October 2024

Human intelligence versus artificial intelligence in classifying economics research articles: exploratory evidence

Jussi T.S. Heikkilä
LUT University, Lahti Campus, Lahti, Finland and
Jyväskylä University School of Business and Economics, Jyväskylä, Finland

Abstract

Purpose – We compare human intelligence to artificial intelligence (AI) in the choice of appropriate Journal of Economic Literature (JEL) codes for research papers in economics.

Design/methodology/approach – We compare the JEL code choices related to articles published in the recent issues of the *Journal of Economic Literature* and the *American Economic Review* and compare these to the original JEL code choices of the authors in earlier working paper versions and JEL codes recommended by various generative AI systems (OpenAI's ChatGPT, Microsoft's Copilot, Google's Gemini) based on the abstracts of the articles.

Findings – There are significant discrepancies and often limited overlap between authors' choices of JEL codes, editors' choices as well as the choices by contemporary widely used AI systems. However, the observations suggest that generative AI can augment human intelligence in the micro-task of choosing the JEL codes and, thus, save researchers time.

Research limitations/implications – Rapid development of AI systems makes the findings quickly obsolete.

Practical implications – AI systems may economize on classification costs and (semi-)automate the choice of JEL codes by recommending the most appropriate ones. Future studies may apply the presented approach to analyze whether the JEL code choices between authors, editors and AI systems converge and become more consistent as humans increasingly interact with AI systems.

Originality/value – We assume that the choice of JEL codes is a micro-task in which boundedly rational decision-makers rather satisfice than optimize. This exploratory experiment is among the first to compare human intelligence and generative AI in choosing and justifying the choice of optimal JEL codes.

Keywords JEL codes, Artificial intelligence, Large language models, Search costs, Bounded rationality

Paper type Research paper

Introduction

The Journal of Economic Literature (JEL) classification codes system maintained and published by the American Economic Association is the de facto standard for classifying research papers in economics (Cherrier, 2017; Heikkilä, 2021, 2022; Bornmann and Wohlrabe, 2024). Data on JEL classification codes has been utilized to analyze the evolution of published economics papers by fields and styles (e.g. Card and DellaVigna, 2013; Angrist *et al.*, 2017; Bornmann and Wohlrabe, 2024). Kosnik (2018) has documented that there can be differences in the author-assigned and editor-assigned JEL codes (in the American Economic Review journal) and Heikkilä (2022) further illustrated that JEL codes of economics working papers can differ from those of the final peer-reviewed and published articles. Concurrently, the micro task of classifying economics research papers has become increasingly complex. For instance, Bornmann and Wohlrabe (2024) report that the average number of JEL codes per paper has increased steadily from about 1.9 in 1991 to 4.3 in 2021.

Using "human intelligence" may lead to subjective and boundedly rational choices (Artinger *et al.*, 2022) and different researchers choose different JEL codes that in their

© Jussi T.S. Heikkilä. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

I thank two anonymous reviewers for their helpful comments. Financial support from the Päijät-Häme Regional Fund of the Finnish Cultural Foundation is gratefully acknowledged.

Journal of Documentation
Vol. 81 No. 7, 2025
pp. 18-30
Emerald Publishing Limited
e-ISSN: 1758-7379
p-ISSN: 0022-0418
DOI 10.1108/JD-05-2024-0104

subjective opinion given their evolving beliefs and expectations are the most appropriate ones. It is an empirical question, to what extent researchers in reality familiarize themselves with the JEL Classification Codes guide when choosing JEL codes (Heikkilä, 2022) for their articles and how much there is rational inattention (Maćkowiak *et al.*, 2023). The guide is available online [1] and it has quite detailed instructions for the use of specific JEL codes so that it requires a costly and time-consuming effort to learn and follow the guide's recommendations. Presumably, there is a non-negligible amount of rational inattention around the use of JEL codes: researchers may think that the expected benefits of choosing appropriate JEL codes may not exceed the costs of learning the detailed instructions provided by the JEL codes guide.

It is important to keep in mind that originally, JEL codes classification system and its predecessors were developed to decrease search costs in the "paper era" in the 20th century (Cherrier, 2017), but now that we live in the era of digitized information, researchers can easily conduct their searches using online search engines that enable keyword searches from full texts of articles which has tremendously decreased search costs (cf. Goldfarb and Tucker, 2019). As in recent years, we have seen significant progress in the field of large language models (LLM) [2] and generative AI [3], we are expecting an increasing number of tasks to be automated (Brynjolfsson *et al.*, 2023; Eloundou *et al.*, 2024). Presumably, search costs will further decrease and time savings related to literature reviews increase as continuously improving AI systems can review and process whole databases of research papers. Recently, Korinek (2023) argued that "economists can reap significant productivity gains by taking advantage of generative AI to automate micro-tasks" and it seems that choosing JEL codes for research papers is one such task. Thus, the simple research question that is explored in this paper is the following: Can generative AI systems augment human intelligence in the choice of appropriate JEL codes? The answer is positive and this is demonstrated with a simple experiment next.

Human versus artificial intelligence in the choice of JEL codes: an experiment

Do researchers study AEA's JEL codes guide when submission systems of journals require them to assign appropriate JEL codes to their articles? Presumably, in this micro task, many researchers do not optimize but rather search JEL codes until they are satisfied (meeting satisficing aspiration level, cf. Simon, 1955; Artinger *et al.*, 2022) with their "good enough" JEL code choices and stop searching (cf. Caplin *et al.*, 2011). Due to technological progress, AI can nowadays be utilized either to replace human choice of JEL codes altogether or complement and augment human intelligence in such a classification task as illustrated in Figure 1. On the one hand, AI systems can base their choices on a huge and accumulating amount of training data which processing is beyond human capacity. On the other hand, AI systems may not be able to adapt to changes (e.g. new classification codes) in the JEL codes classification system as flexibly as humans, for instance when no existing articles in the training data are associated with novel JEL codes. As both human intelligence and artificial intelligence have their pros and cons, augmented intelligence might be the preferred option – at least for now.

However, there is no agreed concept of "optimal choice of JEL codes" – as there is no clear and consistent guidelines for the choice of keywords for articles (cf. Lu *et al.*, 2020). It is not unreasonable to assume that there is no common knowledge of JEL code choice criteria. Therefore, we intentionally do not define the concept "optimal" (or appropriate) here - that is, what and whose preferences ought to define what should be optimized in the choice of JEL codes. However, we prompted selected AI systems (OpenAI's ChatGPT, Microsoft's Copilot, Google's Gemini) to explain what they consider to be important using the following prompt:

> What should researchers consider when choosing the optimal JEL codes for their research articles? What should they optimize?

Table A1 in the Appendix summarizes the answers of AI systems. To summarize, they generally note the goal of the JEL code choice to be to maximize visibility, discoverability and impact

Human intelligence	Augmented intelligence	Artificial intelligence
• Choosing JEL codes manually with limited cognitive abilities and under bounded rationality. • Learning of the JEL codes guide is a costly and time-consuming task. • Subjective beliefs and expectations may matter in the choice of JEL codes. • Can flexibly adapt to changes in the JEL classification codes system (e.g., introduction of new JEL codes).	• AI systems can complement and augment researchers' abilities in choosing appropriate JEL codes. • Possibility to semi-automate better-informed choices of JEL codes with the synergy of human intelligence and AI systems.	• Choosing JEL codes objectively based on patterns of training data and machine learning. • Possibility to automate the choice of JEL codes (what should an AI system optimize in the choice of JEL codes)? • Generative AI systems (chatbots) can provide reasoning why specific JEL codes are appropriate. • May not be as flexible as human intelligence in adapting to changes in the JEL classification codes system (e.g., training data related to novel JEL codes accumulates over time).

Source(s): Author's illustration

Figure 1. Comparing human intelligence, augmented intelligence and AI in the choice of JEL codes

(ChatGPT 3.5 and 4 refer explicitly to citations) among appropriate and intended academic audiences. All selected AI systems list relevance and accuracy - JEL codes should accurately reflect the content (topics, fields, themes, focus, scope, methodologies) of the articles - as key factors to consider. Specificity is also listed, but its definition is slightly ambiguous. ChatGPT 4 highlights that choosing specific JEL codes "can increase the visibility of the article among researchers who are working on the same niche" while Gemini recommends to "avoid overly general secondary codes". Copilot and Gemini recommend consulting AEA's JEL codes guide. While Copilot advises not to overuse JEL codes, a bit surprisingly Gemini recommends limiting the choice to a maximum of two JEL codes (however, it recommends more codes itself as demonstrated in Table 1a and 2a below). ChatGPT 3.5 considers consistency with the JEL codes used in the existing literature and maintaining coherence within the academic discourse important and Gemini recommends considering journal audience and look at the typical JEL codes used in your target journal (even discuss the JEL code choice with colleagues). Copilot links the choice of JEL codes to the choice of keywords by noting that "look for keywords and phrases within the article that match the JEL code descriptions". Other factors listed include interdisciplinarity, current trends and emerging topics.

In order to compare the JEL code choices between human intelligence (by researchers) and artificial intelligence, we chose a set of articles, the latest published issue of Journal of Economic Literature as of 12 April 2024. This is the March 2024 issue, issue 1 of volume 62. It includes seven articles that are shown in Table 1. The Journal of Economic Literature journal is particularly appropriate case to study the choice of JEL codes as it is the outlet where the JEL classification system was introduced in 1969 and has been publishing the official classification system ever since [4]. We proceeded by prompting three generative AI systems (OpenAI's ChatGPT, Microsoft's Copilot, Google's Gemini) to assign JEL codes to the articles using the following prompt [5]:

Table 1. The choices of JEL codes based on human intelligence and artificial intelligence, Journal of Economic Literature 62(1)

(a)

	Assigned JEL codes						
		Based on abstract only					
Article published in Journal of Economic Literature Volume 62 Issue 1 (March 2024)	**Peer-reviewed Journal of Economic literature publication**	**ChatGPT3.5**	**ChatGPT4**	**Copilot**	**Gemini**	Average overlap with the final JEL codes	Average difference in the number of JEL codes
Alesina, Alberto, and Marco Tabellini. "The Political Effects of Immigration: Culture or Economics?"	D72, H23, J11, J15, K37, R23, Z13	D72, J61, J15, Z13, O17 (42.86%, -2)	F22, D72, Z13, H53, R23 (42.86%, -2)	D72, J15, J61 (28.57%, -4)	Primary: F22, J15, Secondary: D72, H12 (28.57%, -3)	35.72 %	-2.75
Zhuravskaya, Ekaterina, Sergei Guriev and Andrei Markevich. "New Russian Economic History."	N10, N30, N40, N60, P20, P30, Z13	N4, P2, O1, O57 (0%, -3)	N43, N44, P26, P48 (0%, -3)	N60, P20, P30, Z13 (57.14%, -3)	Primary: N13, O10 Secondary: P11, P27 (0%, 0)	14.29 %	-2.25
Capraro, Valerio, Joseph Y. Halpern, and Matjaž Perc. "From Outcome-Based to Language-Based Preferences."	C70, C90, D11, D90, Z13	D03, C72, D64, D82, Z13 (20%, 0)	C72, D91, Z13 (20%, -2)	D03, D83, C72 (0%, -2)	Primary: C72 Secondary: D03, A12 (0%, -2)	10.00 %	-1.5
Rohner, Dominic. "Mediation, Military, and Money: The Promises and Pitfalls of Outside Interventions to End Armed Conflicts."	C78, D74, D82, F13, F51, F52	F51, H56, D74, O19, P48 (33.33%, -1)	F51, F52, O19, D74 (50%, -2)	F51, H56, O19 (16.67%, -3)	Primary: C72 Secondary: F51, O15 (16.67%, -3)	29.17 %	-2.25
Goldin, Ian, Pantelis Koutroumpis, François Lafond, and Julian Winkler. "Why Is Productivity Slowing Down?"	E23, E24, J24, L16, O33, O47	O47, E22, F14, D24, O30 (16.67%, -1)	O47, E22, F43, O33 (16.67%, -2)	E23, E24, J24 (50%, -3)	Primary: O33, Secondary: O14, F41 (16.67%, -3)	16.75 %	-2.25
Hadavand, Aboozar, Daniel S. Hamermesh, and Wesley W. Wilson. "Publishing Economics: How Slow? Why Slow? Is Slow Productive? How to Fix Slow?"	A11, B29	A14, D23, G14, O30, M21 (0%, +3)	A11, A14, O33, J44 (50%, +2)	D83, L82, M37 (0%, +1)	Primary: A11 Secondary: O34, I23 (50%, +1)	25.00 %	1.75
Woodford, Michael. "Beyond the Natural Rate: Stephen Marglin on the Instability of Market Economies."	E12, E23, E24, E32, E41, E52	B22, E12, E32, E52, E62 (50%, -1)	E12, E32, E61 (33.33%, -3)	B22, E12, E32 (33.33%, -3)	Primary: B22, Secondary: E11, B52 (0%, -3)	29.17 %	-2.5
Average number of JEL codes	5.571	4.857	3.857	3.143	3.286		
Average overlap with the final JEL codes		23.27 %	30.41 %	26.53 %	15.99 %		
Average difference in the number of JEL codes		-0.714	-1.714	-2.429	-1.857		

(b)

	Assigned JEL codes			
			Based on whole article	
Article published in Journal of Economic Literature Volume 62 Issue 1 (March 2024)	**Working paper (presumably chosen by authors)**	**Peer-reviewed Journal of Economic literature publication**	**ChatGPT 4, working paper**	**ChatGPT4, Final publication**
Alesina, Alberto, and Marco Tabellini. "The Political Effects of Immigration: Culture or Economics?"	D72, J11, J15, J61, Z1 (42.86%, -2)	D72, H23, J11, J15, K37, R23, Z13	D72, J15, Z13, F22, J61 (42.86%, -2)	D72, H23, J11, J15, K37, R23, Z13 (100%, 0)
Zhuravskaya, Ekaterina, Sergei Guriev and Andrei Markevich. "New Russian Economic History."	N00 (0%, -6)	N10, N30, N40, N60, P20, P30, Z13	N43, N44, P26, P36, Q18 (0%, -2)	N10, N30, N40, N60, P20, P30, Z13 (100%, 0)
Capraro, Valerio, Joseph Y. Halpern, and Matjaž Perc. "From Outcome-Based to Language-Based Preferences."	C70, C91, D01, D03, D63 (20%, 0)	C70, C90, D11, D90, Z13	C70, C91, D01, D03, D63 (20%, 0)	C70, C90, D11, D90, Z13 (100%, 0)
Rohner, Dominic. "Mediation, Military, and Money: The Promises and Pitfalls of Outside Interventions to End Armed Conflicts."	D74, H56, N40 (16.67%, -3)	C78, D74, D82, F13, F51, F52	D74, F51, H56, N40, O19 (33.33%, -1)	F51, F52, D74, C78, O19 (66.67%, -1)
Goldin, Ian, Pantelis Koutroumpis, François Lafond, and Julian Winkler. "Why Is Productivity Slowing Down?"	O40, E66, D24 (0%, -3)	E23, E24, J24, L16, O33, O47	O40, E66, D24 (0%, -3)	E23, E24, J24, L16, O33, O47 (100%, 0)
Hadavand, Aboozar, Daniel S. Hamermesh, and Wesley W. Wilson. "Publishing Economics: How Slow? Why Slow? Is Slow Productive? How to Fix Slow?"	A11, B31 (50%, 0)	A11, B29	A11, A14, C80, J44, Y90 (50%, +3)	A11, A14, B29, C89, J01 (100%, +3)
Woodford, Michael. "Beyond the Natural Rate: Stephen Marglin on the Instability of Market Economies."	NA NA	E12, E23, E24, E32, E41, E52	B22, E12, E32, E52, J31 (50%, -1)	E12, E24, E32, E41, E52 (83.33%, -1)
Average number of JEL codes	3.167	5.571	4.714	5.714
Average overlap with the final JEL codes	21.59 %		28.03 %	92.86 %
Average difference in the number of JEL codes	-2.333		-0.857	0.143

Note(s): The listed articles appeared in the March 2024 issue (Vol. 62 No. 1) of the Journal of Economic Literature which was the current issue as of April 2024, available at: https://www.aeaweb.org/issues/755. The first number in parentheses under assigned JEL codes is the overlap with the final JEL codes and the second number is the difference in the number of assigned JEL codes compared to the final ones. Used AI systems: ChatGPT, https://chat.openai.com/; Copilot, copilot.microsoft.com; Gemini, https://gemini.google.com/. JEL codes are presented in the sequence provided in papers and recommended by AI systems. Data (incl. information about selected working papers) and full responses of AI systems are available in the Appendix/Supplementary Material

Source(s): Table by the author

Table 2. The choices of JEL codes based on human intelligence and artificial intelligence, American Economic Review 114(4)

(a)

Article published in American Economic Review Volume 114 Issue 4 (April 2024)	Assigned JEL codes: Peer-reviewed American Economic Review publication, website	Based on abstract only: ChatGPT 3.5	ChatGPT 4	Copilot	Gemini	Average overlap with the final JEL codes	Average difference in the number of JEL codes
Angelucci, Charles, and Andrea Prat. "Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News."	D72, D83, L82	D83, D72 (66.67%, -1)	D83, D72, Z13 (66.67%, 0)	D83, D72, D91, Z13, C83, D81 (66.67%, +3)	Primary: D83 Secondary: A12, D72 (66.67%, 0)	66.67 %	0.50
Engelmann, Jan B., Maël Lebreton, Nahuel A. Salem-Garcia, Peter Schwardmann, and Joël J. van der Weele. "Anticipatory Anxiety and Wishful Thinking."	C91, D12, D83, D91	D81, D91, C91 (50%, -1)	C91, D81, D91 (50%, -1)	C91, D83, D12, D91 (100%, 0)	Primary: D81 Secondary: C91, D03 (25%, -1)	56.25 %	-0.75
Exley, Christine L., and Judd B. Kessler. "Motivated Errors."	C91, D12, D64, D81, D83, D91	D03, D83, C91, C99 (33.33%, -2)	C91, D81, D91 (50%, -3)	C91, D12, D64, D81, D83 (100%, 0)	Primary: D83 Secondary: C72, D02 (16.67%, -3)	50.00 %	-2.00
Grossman, Gene M., Elhanan Helpman, and Stephen J. Redding. "When Tariffs Disrupt Global Supply Chains."	D72, F13, F14, L14, O19, P33	F13, F14, L14, D83 (50%, -2)	F13, F23, L14, C63 (33.33%, -2)	F12, F13 (16.67%, -4)	Primary: F12 Secondary: D22, L11, O11 (0%, -2)	25.00 %	-2.50
Baum-Snow, Nathaniel, Nicolas Gendron-Carrier, and Ronni Pavan. "Local Productivity Spillovers."	D22, D24, G32, L25, R11, R32	D24, R11, R12, R32 (50%, -2)	R12, L25, D24, R23 (33.33%, -2)	D22, D24, G32, L25 (66.67%, -2)	Primary: R11 Secondary: C61, O40 (16.67%, -3)	41.67 %	-2.25
Greenberg, Kyle, Parag A. Pathak, and Tayfun Sönmez. "Redesigning the US Army's Branching Process: A Case Study in Minimalist Market Design."	D47, H56, J45	D02, D71, H56 (33.33%, 0)	D47, M21, J45, C78 (66.67%, +1)	D47, D02, D71 (33.33%, 0)	Primary: D72 Secondary: D82, J41 (0%, 0)	16.42 %	0.25
Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski. "The Political Development Cycle: The Right and the Left in People's Republic of China from 1953."	D72, N15, N45, N55, O21, P21, P24	O11, O47, P26 (0%, -4)	O11, P16, Q15, O53 (0%, -3)	D71, N15, O11, P21 (42.86%, -3)	Primary: O10 Secondary: N16, P16 (0%, 0)	10.72 %	-2.50
Martínez-Marquina, Alejandro, and Mike Shi. "The Opportunity Cost of Debt Aversion."	C91, D91, G51	D14, D91, G11 (33.33%, 0)	D14, D81, D91, G11 (33.33%, +1)	C91, D91, G51 (100%, 0)	Primary: G21 Secondary: D81, D03 (0%, 0)	41.67 %	0.25
Alan, Sule, and Ipek Mumcu. "Nurturing Childhood Curiosity to Enhance Learning: Evidence from a Randomized Pedagogical Intervention."	D83, I21, I26, J13, O15	I21, I28, O15 (40%, -2)	I21, I28, D91, C93 (20%, -1)	D83, I21, I26 (60%, -2)	Primary: I21 Secondary: O15, C9 (40%, -2)	40.00 %	-1.75
Average number of JEL codes	4.778	3.222	3.667	3.778	3.111		
Average overlap with the final JEL codes		39.63 %	39.26 %	65.13 %	18.33 %		
Average difference in the number of JEL codes		-1.556	-1.111	-0.889	-1.222		

(b)

Article published in American Economic Review Volume 114 Issue 4 (April 2024)	Assigned JEL codes: Working paper (presumably chosen by authors)	Peer-reviewed American Economic Review publication, website	Based on whole article: ChatGPT 4, working paper	ChatGPT4, final article
Angelucci, Charles, and Andrea Prat. "Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News."	L82, D72, D83, D90 (100%, +1)	D72, D83, L82	D72, D83, L82 (100%, 0)	D72, D83, L82 (100%, 0)
Engelmann, Jan B., Maël Lebreton, Nahuel A. Salem-Garcia, Peter Schwardmann, and Joël J. van der Weele. "Anticipatory Anxiety and Wishful Thinking."	C91, D83 (50%, -2)	C91, D12, D83, D91	D83, D91, C91 (75%, -1)	C91, D12, D83, D91 (100%, 0)
Exley, Christine L., and Judd B. Kessler. "Motivated Errors."	C91, D64, D91 (50%, -3)	C91, D12, D64, D81, D83, D91	D03, D91, C91 (33.33%, -3)	D12, D81, D83 (50%, -3)
Grossman, Gene M., Elhanan Helpman, and Stephen J. Redding. "When Tariffs Disrupt Global Supply Chains."	F1, F13 (16.67%, -4)	D72, F13, F14, L14, O19, P33	F13, F23, D21 (16.67%, -3)	F13, F23, D23 (16.67%, -3)
Baum-Snow, Nathaniel, Nicolas Gendron-Carrier, and Ronni Pavan. "Local Productivity Spillovers."	-	D22, D24, G32, L25, R11, R32	R12, D24, C23 (16.67%, -3)	D22, D24, R12 (33.33%, -3)
Greenberg, Kyle, Parag A. Pathak, and Tayfun Sönmez. "Redesigning the US Army's Branching Process: A Case Study in Minimalist Market Design."	D47 (33.33%, -2)	D47, H56, J45	C78, D47, D82, J45 (66.67%, +1)	D47, H56, J45 (100%, 0)
Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski. "The Political Development Cycle: The Right and the Left in People's Republic of China from 1953."	-	D72, N15, N45, N55, O21, P21, P24	O11, P16, P26, N55 (14.29%, -3)	D72, N15, N45, N55, O21, P21, P24 (100%, 0)
Martínez-Marquina, Alejandro, and Mike Shi. "The Opportunity Cost of Debt Aversion."	C91, D14, D91 (66.67%, 0)	C91, D91, G51	C91, D14, D91 (66.67%, 0)	C91, D14, D91 (66.67%, 0)
Alan, Sule, and Ipek Mumcu. "Nurturing Childhood Curiosity to Enhance Learning: Evidence from a Randomized Pedagogical Intervention."	NA NA	D83, I21, I26, J13, O15	I21, I28, C93 (20%, -2)	I21, I28, D83, J13 (60%, -1)
Average number of JEL codes	2.500	4.778	3.222	3.667
Average overlap with the final JEL codes	52.78 %		45.48 %	69.63 %
Average difference in the number of JEL codes	-1.667		-1.556	-1.111

Note(s): The listed articles appeared in the April 2024 issue (Vol. 114 No. 4) of the American Economic Review which was the current issue as of April 2024, available at: https://www.aeaweb.org/issues/757. The first number in parentheses under assigned JEL codes is the overlap with the final JEL codes and the second number is the difference in the number of assigned JEL codes compared to the final ones. Used AI systems: ChatGPT, https://chat.openai.com/; Copilot, copilot.microsoft.com; Gemini, https://gemini. google.com/. Data (incl. information about selected working papers) and full responses of AI systems are available in the Appendix/Supplementary Material
Source(s): Table by the author

Please, assign optimal JEL codes to the following abstract and explain why they are optimal: [Abstract here]

Table 1 compares the JEL codes assigned to the articles (shaded columns) to the JEL codes that AI systems assigned to the articles based on their abstracts only (Table 1a) and based on full texts (Table 1b) of earlier working paper versions of the same articles as well as the final versions. The option to provide the AI system with the full text is not available for all the used AI systems and the full texts of the articles are mainly available for subscribers only. Thus, to enable comparisons between different AI systems we used only abstracts in Table 1a since the abstracts are publicly available online for anyone.

Since there are often multiple versions of earlier working papers, in Table 1b we chose the earliest working paper versions that we found from IDEAS RePEc and Google Scholar and prioritized major established working paper series in economics (e.g. NBER, CEPR, cf. Baumann and Wohlrabe, 2020) [6]. This illustrates how the set of JEL codes assigned to a working paper – presumably, typically by authors themselves – can be significantly different from the ones assigned to the peer-reviewed final paper.

Several patterns can be observed even from this limited set of articles. First, both AI systems and authors assign typically systematically less JEL codes compared to the final set and authors often assign less JEL codes than AI. In the case of AI systems, this could be explained by the fact that we prompted the AIs to suggest JEL codes based on the abstract only and not based on the whole article. For instance, Gemini typically suggests three JEL codes, one primary and two secondary ones. The justification for the choice of the specific JEL codes by the AIs is reasonable and Gemini even provides “justification for excluding other JEL codes” and explains why some selected JEL codes would not be appropriate.

Second, the overlap between the ones suggested by AI systems based on abstracts and final JEL codes varies in the range of 10%–30% (Table 1a). This may seem low, but it should be noted that AI systems (as well as authors) suggest systematically less JEL codes. Unlike other selected AI systems, Copilot provides by default the information sources underlying its reasoning for the JEL codes. The investigation of these sources reveals that in multiple cases Copilot refers to the publicly available working papers or the website of the final version of the underlying article where the abstract is available. Sometimes Copilot ends its answer: “For more details on the paper, you can refer to the [link to the article here].”

Third, we also tested which JEL codes ChatGPT 4 would assign to the selected working paper version and the final peer-reviewed articles based on the text of the whole article (Table 1b). Despite the fact that the differences between the content, focus and scope of the working paper version and the final article are relatively minor, ChatGPT 4 assigns quite different JEL codes to them. It seems that in most cases ChatGPT 4 recommends exactly the JEL codes listed in the articles and argues why they are appropriate.

To conclude, generative AI systems can augment human intelligence in choosing the JEL codes by providing reasoned suggestions based on the article abstracts only.

Next, in order to test the robustness of our observation regarding discrepancies of JEL code choices between human and artificial intelligence in the Journal of Economic Literature, we applied the same method to the nine articles published in the latest (as of April 2024) issue of the American Economic Review (AER).

Again, as Table 2 presents, we find that the generative AI can suggest reasonable JEL codes and provide the reasoning why these could be the optimal set of JEL codes (see Supplementary material). Again, in line with Kosnik’s (2018) and Heikkilä’s (2022) observations, we find that author-assigned JEL codes to the working paper versions (Table 2b) often differ from the final assigned JEL codes.

Again, authors (Table 2b) and AI systems assign typically less JEL codes and Table 2a indicates that the overlap between the JEL codes assigned by AI systems based on abstract and the final JEL codes ranges between ca. 18% (Gemini) and 65% (Copilot). For both ChatGPT 3.5 and 4 the overlap is about 40%. While Copilot sometimes refers to the websites where the abstract of the article - and related JEL codes - are publicly available, it often does not pick those exact JEL codes but rather assigns a smaller number of JEL codes. For this set of articles, the recommended JEL codes by ChatGPT4 based on the whole articles have lower overlap (ca. 70%) with the final ones compared to the overlap reported in Table 1b (ca. 93%).

In the prior literature it is common to focus on more aggregated levels of JEL codes instead of the most granular ones as we did in Tables 1 and 2. For instance, Card and Della Vigna (2013) classified articles based on their own classification where JEL codes were aggregated into 14 field categories and recently Bornmann and Wohlrabe (2024) focused in their analyses on the level of 20 primary JEL code categories (see also Kosnik, 2018). In Tables A2 and A3 in the Appendix, we apply this aggregated approach based on the 20 primary JEL code categories to the articles presented in Tables 1 and 2. These further analyses show a higher level of overlap as expected (ranging between 50–100%) indicating that while the AI systems may not recommend exactly the same JEL codes, they at least recommend JEL codes from the same JEL code categories in most of the cases. It seems that the imperfect overlap stems mainly from the fact that AI system as well as authors select fewer JEL codes compared to the final editor-assigned JEL codes of the articles.

The exploratory evidence presented here indicates that generative AI may help researchers and augment human intelligence in the choice of appropriate JEL codes for their articles. At minimum, generative AI can be used to cross-check the choices of researchers if not to fully automate the micro task. Thus, augmented intelligence may make the use of JEL classification codes more efficient and consistent and save researchers' time.

Discussion

Since learning the nuances of the AEA's JEL codes guide requires costly and time-consuming effort, it seems economic to utilize AI systems in (partially) automating the micro-task of choosing JEL codes. Our simple experiment indicates that generative AI may help researchers and augment their boundedly rational human intelligence in choosing appropriate JEL codes for their articles based on abstracts. However, we also documented that there are significant discrepancies between the chosen JEL codes by humans and the selected generative AI systems as well as between the AI systems.

If the scientific community and research publishers want to continue classifying research papers using JEL codes, then the use of AI may help make the human choice of JEL codes less boundedly rational and more consistent (cf. Kosnik, 2018). While it remains an open question how to define "appropriate" or "optimal" choice of JEL codes, AI can save time and help in finding more "satisficing" (Simon, 1955; Caplin *et al.*, 2011; Artinger *et al.*, 2022) sets of JEL codes. More consistent use of JEL codes improves the training data of AI systems. When this is complemented with automated recommendation systems that suggest JEL codes best describing research content, it could further decrease the search costs of the audiences as well as promote the analysis of research trends based on JEL codes (cf. Card and DellaVigna, 2013; Angrist *et al.*, 2017; Bornmann and Wohlrabe, 2024).

We acknowledge that the presented preliminary observations have several limitations. First, the analysis focuses on only two recent issues of leading economics journals, so the external validity of the observations is limited. More extensive analyses of larger numbers of articles across a more diverse set of journals would lead to more credible and generalizable observations.

Second, there is a hallucination problem with large language models – that is, they may generate text that is not true (Zhai, 2024). In this analysis we did not try to detect hallucination, but a more rigorous analysis of JEL code choices with larger sets of articles should be accompanied with the check of reasoning for each selected JEL code (to confirm that the AI systems do not come up with any hallucinated JEL codes).

Third, "model collapse" (Shumailov *et al.*, 2024) is another detrimental phenomenon which refers to the degenerative recursive process where AI systems trained with polluted (incl. hallucinated) data end up training the next generation of AI systems with model-generated polluted data. Similarly, in the context of JEL codes, if AI systems are again and again trained with data were inappropriate or hallucinated JEL codes are statistically linked to specific articles, this probably compounds the biases dynamically.

Fourth, the analysis used only a very limited set of AI systems. As the development of AI systems continues, authors can consult an increasing number of continuously improving AI systems to cross-check their recommendations of optimal JEL codes.

While the preliminary findings presented here will become quickly obsolete as AI systems (incl. training data) improve and are increasingly utilized, future studies may apply the presented approach to analyze whether the JEL code choices between authors, editors and AI systems converge and become more consistent over time.

Notes

1. See https://www.aeaweb.org/jel/guide/jel.php Accessed 10 April 2024. The guide also lists "Caveats" which, for instance, in the case of JEL code D82 "Asymmetric and Private Information: Mechanism Design" mention that "Studies about information in general not asymmetric or private should be classified under D83 [Search: Learning; Information and Knowledge; Communication; Belief; Unawareness]. Theoretical studies about contract theory should be classified under D86 [Economics of Contracts: Theory]". Then, D86 lists additional caveats that should be taken into account in classification.

2. See, e.g. Zhai (2024) for a recent review of the opportunities and challenges in the context of large language models and information retrieval. Zhai (2024, p. 481) notes that "While statistical language models have been applied to information retrieval (IR) since many decades ago, these new LLMs go far beyond traditional language models in their representation learning capacity, which enabled them to both understand natural language semantically and generate fluent meaningful natural language text."

3. OpenAI launched ChatGPT 3.5 in November 2022 and GPT4 in March 2023. Google launched Bard in February 2023 and rebranded it as Gemini in February 2024. Microsoft launched Copilot in February 2023 (formerly Bing Chat). There are also several other generative AI systems and new ones are launched, but we intentionally focus on this limited set of popular ones (as of April 2024).

4. See Cherrier (2017) and Heikkilä (2022) for additional information about the history of the JEL codes classification system.

5. We also experimented by slightly changing the prompts and arrived at similar sets of JEL codes that best fit the papers. For instance, asking the generative AI systems to "choose JEL codes that minimize search costs" leads, generally, to a smaller number of recommended JEL codes.

6. IDEAS is a large bibliographic database (more than 4.7 million items) dedicated to Economics based on Research Papers in Economics (RePEc) data. Available at: https://ideas.repec.org/. A significant share of working papers in the field of economics are indexed in RePEc (cf. Baumann and Wohlrabe, 2020).

References

Angrist, J., Azoulay, P., Ellison, G., Hill, R. and Lu, S. (2017), "Economic research evolves: fields and styles", *The American Economic Review*, Vol. 107 No. 5, pp. 293-297, doi: 10.1257/aer.p20171117.

Artinger, F., Gigerenzer, G. and Jacobs, P. (2022), "Satisficing: integrating two traditions", *Journal of Economic Literature*, Vol. 60 No. 2, pp. 598-635, doi: 10.1257/jel.20201396.

Baumann, A. and Wohlrabe, K. (2020), "Where have all the working papers gone? Evidence from four major economics working paper series", *Scientometrics*, Vol. 124 No. 3, pp. 2433-2441, doi: 10.1007/s11192-020-03570-x.

Bornmann, L. and Wohlrabe, K. (2024), "Recent temporal dynamics in economics: empirical analyses of annual publications in economic fields", *Journal of Documentation*, Vol. 80 No. 4, pp. 824-856, doi: 10.1108/JD-10-2023-0201.

Brynjolfsson, E., Li, D. and Raymond, L. (2023), "Generative AI at work", NBER Working Paper No. 31161, available at: http://www.nber.org/papers/w31161

Caplin, A., Dean, M. and Martin, D. (2011), "Search and satisficing", *The American Economic Review*, Vol. 101 No. 7, pp. 2899-2922, doi: 10.1257/aer.101.7.2899.

Card, D. and DellaVigna, S. (2013), "Nine facts about top journals in economics", *Journal of Economic Literature*, Vol. 51 No. 1, pp. 144-161, doi: 10.1257/jel.51.1.144.

Cherrier, B. (2017), "Classifying economics: a history of the JEL codes", *Journal of Economic Literature*, Vol. 55 No. 2, pp. 545-579, doi: 10.1257/jel.20151296.

Eloundou, T., Manning, S., Mishkin, P. and Rock, D. (2024), "GPTs are GPTs: labor market impact potential of LLMs", *Science*, Vol. 384 No. 6702, pp. 1306-1308, doi: 10.1126/science.adj0998.

Goldfarb, A. and Tucker, C. (2019), "Digital economics", *Journal of Economic Literature*, Vol. 57 No. 1, pp. 3-43, doi: 10.1257/jel.20171452.

Heikkilä, J. (2021), "Classifying economics for the common good: a note on the links between sustainable development goals and JEL codes", *Journal of Documentation*, Vol. 77 No. 2, pp. 305-319, doi: 10.1108/JD-08-2020-0146.

Heikkilä, J. (2022), "Journal of economic literature codes classification system (JEL)", *Knowledge Organization*, Vol. 49 No. 5, pp. 352-370, doi: 10.5771/0943-7444-2022-5-352.

Korinek, A. (2023), "Generative AI for economic research: use cases and implications for economists", *Journal of Economic Literature*, Vol. 61 No. 4, pp. 1281-1317, doi: 10.1257/jel.20231736.

Kosnik, L. (2018), "A survey of JEL codes: what do they mean and are they used consistently?", *Journal of Economic Surveys*, Vol. 32 No. 1, pp. 249-272, doi: 10.1111/joes.12189.

Lu, W., Liu, X., Huang, Y., Bu, T., Li, X. and Cheng, Q. (2020), "How do authors select keywords? A preliminary study of author keyword selection behavior", *Journal of Informetrics*, Vol. 14 No. 4, 101066, doi: 10.1016/j.joi.2020.101066.

Maćkowiak, B., Matějka, F. and Wiederholt, M. (2023), "Rational inattention: a review", *Journal of Economic Literature*, Vol. 61 No. 1, pp. 226-273, doi: 10.1257/jel.20211524.

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R. and Gal, Y. (2024), "AI models collapse when trained on recursively generated data", *Nature*, Vol. 631 No. 8022, pp. 755-759, doi: 10.1038/s41586-024-07566-y.

Simon, H. (1955), "A behavioral model of rational choice", *Quarterly Journal of Economics*, Vol. 69 No. 1, pp. 99-118, doi: 10.2307/1884852.

Zhai, C. (2024), "Large language models and future of information retrieval: opportunities and challenges", *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 481-490, doi: 10.1145/3626772.3657848.

Appendix

Table A1. Criteria for choosing JEL codes according to selected AI systems

ChatGPT3.5	ChatGPT4	Copilot	Gemini
Researchers should consider several factors when choosing the optimal JEL (Journal of Economic Literature) codes for their research articles: **Relevance**: The JEL codes should accurately reflect the primary focus of the research. Researchers should select codes that best represent the main topics, themes, and methodologies covered in their study. **Specificity**: Choose codes that provide a detailed description of the research content. Specific JEL codes help to ensure that the article reaches the appropriate audience interested in that particular subfield or topic. **Interdisciplinarity**: Consider whether the research spans multiple disciplines or subfields within economics. In such cases, researchers may choose multiple JEL codes to accurately capture the interdisciplinary nature of their work. **Consistency**: Ensure consistency with the JEL codes used in related literature or previous works. This helps maintain coherence within the academic discourse and facilitates easier identification and classification of the research within the broader literature. **Visibility**: Opt for JEL codes that are commonly used and recognized within the academic community. This increases the visibility of the research and enhances its discoverability by potential readers and collaborators. **Future Impact**: Anticipate how the chosen JEL codes may influence the dissemination and citation of the research. Opting for codes that align with emerging trends or areas of interest within the discipline can potentially increase the impact and relevance of the article. Ultimately, researchers should optimize the selection of JEL codes to accurately represent the content of their research while maximizing its visibility, relevance, and impact within the academic community.	Choosing the right Journal of Economic Literature (JEL) codes for a research article is crucial because it helps in accurately classifying and indexing the research, which in turn aids in reaching the appropriate academic audience. Here are several key considerations researchers should keep in mind when selecting JEL codes: **Relevance**: The primary factor in choosing JEL codes should be the relevance of the code to the core topics and themes of the paper. Each code should directly reflect the subject matter being discussed. Researchers need to ensure that the selected codes accurately represent their study's focus and scope. **Specificity**: While it is important to choose codes that correctly reflect the broad area of the research, opting for more specific codes can help in targeting readers who are most likely to be interested in the particular aspects of the research. Specificity can increase the visibility of the article among researchers who are working in the same niche. **Balance Between Broad and Narrow Codes**: There should be a balance between broad and narrow JEL codes. Broad codes can capture the general area of research, ensuring the work is discoverable by a wider audience, while narrow codes target specialists in particular subfields. Using a combination of both can maximize visibility and relevance. **Current Trends and Emerging Topics**: Researchers should be aware of current trends and emerging topics within their field. Selecting JEL codes that are aligned with these trends can increase the likelihood of the research being recognized as timely and relevant by peers, potentially leading to more citations and greater impact. **Code Limitations**: Most journals limit the number of JEL codes that can be assigned to an article. Therefore, it's important to prioritize and select the codes that are most relevant. This requires careful consideration to ensure that all aspects of the paper are adequately represented within the given constraints. **Optimizing for Audience and Indexing Services**: Choosing the right JEL codes also helps in ensuring that the paper is easily discoverable through economic literature indexing services. This enhances the research's visibility to the intended academic audience, which can include both generalists in economics and experts in specific sub-fields. By focusing on these considerations, researchers can effectively use JEL codes to enhance the visibility and impact of their work within the academic community. The selection of JEL codes should ultimately aim to optimize for accuracy in classification, relevance to potential readers, and alignment with current research trends.	When selecting **JEL (Journal of Economic Literature) codes** for research articles, researchers should take into account several factors to ensure accurate classification and effective communication within the field of economics. Here are some considerations: **Relevance to Research Content**: • Researchers should choose JEL codes that closely align with the primary focus of their research. These codes serve as a way to categorize scholarly literature, so accuracy is crucial. • Consider the main themes, topics, and methodologies covered in the article. For example, if the research primarily deals with macroeconomics, relevant JEL codes would fall under the "E" category. **Specificity and Granularity**: • JEL codes are organized hierarchically, with broader categories (e.g., "E" for macroeconomics) and more specific subcategories (e.g., "E32" for business fluctuations). • Researchers should aim for a balance between specificity and generality. Too broad a code may not accurately capture the nuances of the research, while overly specific codes may limit visibility. **Multiple Codes (if applicable)**: • Some articles cover multiple topics or use interdisciplinary approaches. In such cases, researchers can assign multiple JEL codes to reflect the various aspects of their work. • Be cautious not to overuse codes; select the most relevant ones. **Consult the JEL Classification Guide**: • The **American Economic Association (AEA)** provides a comprehensive JEL Classification Codes Guide that outlines guidelines, keywords, and examples for each classification. • Researchers can refer to this guide to understand how JEL codes are used and to find appropriate codes for their work. **Keywords and Phrases**: • Look for keywords and phrases within the article that match the JEL code descriptions. These can guide the selection process. • For instance, if the research involves health economics, relevant keywords might include "healthcare," "medical costs," or "public health." **Avoid Misclassification**: • Misclassifying an article can lead to it being overlooked by researchers interested in specific topics. • Double-check the definitions and descriptions of JEL codes to avoid errors. In summary, researchers should optimize their JEL code selection by ensuring accuracy, relevance, and appropriate granularity. By following these guidelines, they can effectively communicate the focus of their research within the economics community[12].	When choosing JEL codes, researchers are aiming for a sweet spot between **accuracy** in representing their research and **discoverability** for potential readers. Here's what to consider: **Accuracy:** • **Core Focus:** The primary JEL code should reflect the main methodology or field of study. Use the JEL Classification Codes Guide (https://www.aeaweb.org/jel/guide/jel.php?class=G) to find the best fit based on your topic and approach. • **Specificity:** Balance is key. A primary code captures the broad area, while a well-chosen secondary code highlights a more specific aspect not fully covered by the first. Avoid overly general secondary codes. **Discoverability:** • **Keywords:** Think about search terms people might use to find similar research. Aligning JEL codes with these keywords increases the chances of readers finding your work. • **Journal Audience:** Look at the typical JEL codes used in your target journal. Matching your codes with the journal's focus areas can improve acceptance and readership within that community. **Additional Tips:** • **Keep it Concise:** Limit yourself to a maximum of two JEL codes. • **Seek Input:** Discuss your research with colleagues familiar with your field to get their insights on the most appropriate JEL codes. By carefully considering these factors, researchers can optimize their JEL code selection to ensure their work reaches the intended audience and gains the recognition it deserves.

Note(s): Prompts conducted in May 2024

Source(s): Table by the author

Table A2. The choices of JEL categories based on human intelligence and artificial intelligence, Journal of Economic Literature 62(1)

(a)

Article published in Journal of Economic Literature Volume 62 Issue 1 (March 2024)	Assigned JEL codes: Peer-reviewed Journal of Economic literature publication	Based on abstract only: ChatGPT3.5	ChatGPT4	Copilot	Gemini	Average overlap with the final JEL categories	Average difference in the number of JEL categories
Alesina, Alberto, and Marco Tabellini. "The Political Effects of Immigration: Culture or Economics?"	D, H, J, K, R, Z	D, J, Z, O (50%, -2)	F, D, Z, H, R (66.67%, -1)	D, J (33.33%, -4)	Primary: F, J Secondary: D, H (50%, -2)	50.00 %	-2.25
Zhuravskaya, Ekaterina, Sergei Guriev and Andrei Markevich. "New Russian Economic History."	N, P, Z	N, P, O (66.67%, 0)	N, P (66.67%, -1)	N, P, Z (100%, 0)	Primary: N, O Secondary: P (66.67%, 0)	75.00 %	-0.25
Capraro, Valerio, Joseph Y. Halpern, and Matjaž Perc. "From Outcome-Based to Language-Based Preferences."	C, D, Z	D, C, Z (100%, 0)	C, D, Z (100%, 0)	D, C (66.67%, -1)	Primary: C Secondary: D, A (66.67%, 0)	83.34 %	-0.25
Rohner, Dominic. "Mediation, Military, and Money: The Promises and Pitfalls of Outside Interventions to End Armed Conflicts."	C, D, F	F, H, D, O, P (66.67%, +2)	F, O, D (66.67%, 0)	F, H, O (33.33%, 0)	Primary: C Secondary: F, O (66.67%, 0)	58.34 %	0.5
Goldin, Ian, Pantelis Koutroumpis, François Lafond, and Julian Winkler. "Why Is Productivity Slowing Down?"	E, J, L, O	O, E, F, D (50%, 0)	O, E, F (50%, -1)	E, J (50%, -2)	Primary: O, Secondary: O, F (25%, -2)	43.75 %	-1.25
Hadavand, Aboozar, Daniel S. Hamermesh, and Wesley W. Wilson. "Publishing Economics: How Slow? Why Slow? Is Slow Productive? How to Fix Slow?"	A, B	A, D, G, O, M (50%, +3)	A, O, J (50%, +1)	D, L, M (0%, +1)	Primary: A Secondary: O, I (50%, +1)	37.50 %	1.5
Woodford, Michael. "Beyond the Natural Rate: Stephen Marglin on the Instability of Market Economies."	E	B, E (50%, +1)	E (100%, 0)	B, E (100%, +1)	Primary: B, Secondary: E, B (100%, +1)	87.50 %	0.75
Average number of JEL categories	3.143	3.714	2.857	2.286	3.143		
Average overlap with the final JEL categories		61.91 %	71.43 %	54.76 %	60.72 %		
Average difference in the number of JEL categories		0.571	-0.286	-0.714	-0.286		

(b)

Article published in Journal of Economic Literature Volume 62 Issue 1 (March 2024)	Assigned JEL codes: Working paper (presumably chosen by authors)	Peer-reviewed Journal of Economic literature publication	Based on whole article: ChatGPT 4, working paper	ChatGPT4, Final publication
Alesina, Alberto, and Marco Tabellini. "The Political Effects of Immigration: Culture or Economics?"	D, J, Z (50%, -2)	D, H, J, K, R, Z	D, J, Z, F (50%, -2)	D, H, J, K, R, Z (100%, 0)
Zhuravskaya, Ekaterina, Sergei Guriev and Andrei Markevich. "New Russian Economic History."	N (0%, -2)	N, P, Z	N, P, Q (66.67%, 0)	N, P, Z (100%, 0)
Capraro, Valerio, Joseph Y. Halpern, and Matjaž Perc. "From Outcome-Based to Language-Based Preferences."	C, D (66.67%, -1)	C, D, Z	C, D (66.67%, -1)	C, D, Z (100%, 0)
Rohner, Dominic. "Mediation, Military, and Money: The Promises and Pitfalls of Outside Interventions to End Armed Conflicts."	D, H, N (33.33%, 0)	C, D, F	D, F, H, N, O (66.67%, +2)	F, D, C, O (100%, +1)
Goldin, Ian, Pantelis Koutroumpis, François Lafond, and Julian Winkler. "Why Is Productivity Slowing Down?"	O, E, D (50%, -1)	E, J, L, O	O, E, D (50%, -1)	E, J, L, O (100%, 0)
Hadavand, Aboozar, Daniel S. Hamermesh, and Wesley W. Wilson. "Publishing Economics: How Slow? Why Slow? Is Slow Productive? How to Fix Slow?"	A, B (100%, 0)	A, B	A, C, J, Y (50%, +2)	A, B, C, J (100%, +2)
Woodford, Michael. "Beyond the Natural Rate: Stephen Marglin on the Instability of Market Economies."	NA NA	E	B, E, J (100%, +2)	E (100%, 0)
Average number of JEL categories	2.333	3.143	3.429	3.571
Average overlap with the final JEL categories	50.00 %		64.29 %	100.00 %
Average difference in the number of JEL categories	-1.000		0.286	0.429

Note(s): The listed articles appeared in the March 2024 issue (Vol. 62 No. 1) of the Journal of Economic Literature which was the current issue as of April 2024, available at: https://www.aeaweb.org/issues/755. The first number in parentheses under assigned JEL categories is the overlap with the final JEL categories and the second number is the difference in the number of assigned JEL categories compared to the final ones. Used AI systems: ChatGPT, https://chat.openai.com/; Copilot, copilot.microsoft.com; Gemini, https://gemini.google.com/. Data (incl. information about selected working papers) and full responses of AI systems are available in the Appendix/Supplementary Material

Source(s): Table by the author

Table A3. The choices of JEL categories based on human intelligence and artificial intelligence, American Economic Review 114(4)

(a)

Article published in American Economic Review Volume 114 Issue 4 (April 2024)	Assigned JEL codes: Peer-reviewed American Economic Review publication, website	Based on abstract only: ChatGPT 3.5	ChatGPT 4	Copilot	Gemini	Average overlap with the final JEL categories	Average difference in the number of JEL categories
Angelucci, Charles, and Andrea Prat. "Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News."	D, L	D (50%, -1)	D, Z (50%, 0)	D, Z, C (33.33%, +1)	Primary: D Secondary: A, D (50%, 0)	45.83 %	0.00
Engelmann, Jan B., Maël Lebreton, Nahuel A. Salem-Garcia, Peter Schwardmann, and Joël J. van der Weele. "Anticipatory Anxiety and Wishful Thinking."	C, D	D, C (100%, 0)	C, D (100%, 0)	C, D (100%, 0)	Primary: D Secondary: C, D (100%, 0)	100.00 %	0.00
Exley, Christine L., and Judd B. Kessler. "Motivated Errors."	C, D	D, C (100%, 0)	C, D (100%, 0)	C, D (100%, 0)	Primary: D Secondary: C, D (100%, 0)	100.00 %	0.00
Grossman, Gene M., Elhanan Helpman, and Stephen J. Redding. "When Tariffs Disrupt Global Supply Chains."	D, F, L, O, P	F, L, D (60%, -2)	F, L, C (40%, -2)	F (20%, -4)	Primary: F Secondary: D, L, O (80%, -1)	50.00 %	-2.25
Baum-Snow, Nathaniel, Nicolas Gendron-Carrier, and Ronni Pavan. "Local Productivity Spillovers."	D, G, L, R	D, R (50%, -2)	R, L, D 75%, -1)	D, G, L (75%, -1)	Primary: R Secondary: C, O (25%, -1)	56.25 %	-1.25
Greenberg, Kyle, Parag A. Pathak, and Tayfun Sönmez. "Redesigning the US Army's Branching Process: A Case Study in Minimalist Market Design."	D, H, J	D, H (66.67%, -1)	D, M, J, C (50%, +1)	D (33.33%, -2)	Primary: D Secondary: D, J (66.67%, -1)	54.17 %	-0.75
Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski. "The Political Development Cycle: The Right and the Left in People's Republic of China from 1953."	D, N, O, P	O, P (50%, -2)	O, P, Q (50%, -1)	D, N, O, P (100%, 0)	Primary: O Secondary: N, P (75%, -1)	68.75 %	-1.00
Martínez-Marquina, Alejandro, and Mike Shi. "The Opportunity Cost of Debt Aversion."	C, D, G	D, G (66.67%, -1)	D, G (66.67%, -1)	C, D, G (100%, 0)	Primary: G Secondary: D (66.67%, -1)	75.00 %	-0.75
Alan, Sule, and Ipek Mumcu. "Nurturing Childhood Curiosity to Enhance Learning: Evidence from a Randomized Pedagogical Intervention."	D, I, J, O	I, O (50%, -2)	I, D, C (50%, -1)	D, I (50%, -2)	Primary: I Secondary: O, C (50%, -1)	50.00 %	-1.50
Average number of JEL categories	3.222	2.000	2.667	2.333	2.556		
Average overlap with the final JEL categories		65.93 %	64.63 %	67.96 %	68.15 %		
Average difference in the number of JEL categories		-1.222	-0.556	-0.889	-0.667		

(b)

Article published in American Economic Review Volume 114 Issue 4 (April 2024)	Assigned JEL codes: Working paper (presumably chosen by authors)	Peer-reviewed American Economic Review publication, website	Based on whole article: ChatGPT 4, working paper	ChatGPT4, final article
Angelucci, Charles, and Andrea Prat. "Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News."	L, D (100%, 0)	D, L	D, L (100%, 0)	D, L (100%, 0)
Engelmann, Jan B., Maël Lebreton, Nahuel A. Salem-Garcia, Peter Schwardmann, and Joël J. van der Weele. "Anticipatory Anxiety and Wishful Thinking."	C, D (100%, 0)	C, D	D, C (100%, 0)	C, D (100%, 0)
Exley, Christine L., and Judd B. Kessler. "Motivated Errors."	C, D (100%, 0)	C, D	D, C (100%, 0)	D (50%, -1)
Grossman, Gene M., Elhanan Helpman, and Stephen J. Redding. "When Tariffs Disrupt Global Supply Chains."	F (20%, -4)	D, F, L, O, P	F, D (40%, -3)	F, D (50%, -3)
Baum-Snow, Nathaniel, Nicolas Gendron-Carrier, and Ronni Pavan. "Local Productivity Spillovers."	-	D, G, L, R	R, D, C (50%, -1)	D, R (50%, -2)
Greenberg, Kyle, Parag A. Pathak, and Tayfun Sönmez. "Redesigning the US Army's Branching Process: A Case Study in Minimalist Market Design."	D (33.33%, -2)	D, H, J	C, D, J (66.67%, 0)	D, H, J (100%, 0)
Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski. "The Political Development Cycle: The Right and the Left in People's Republic of China from 1953."	-	D, N, O, P	O, P, N (75%, -1)	D, N, O, P (100%, 0)
Martínez-Marquina, Alejandro, and Mike Shi. "The Opportunity Cost of Debt Aversion."	C, D (66.67%, -1)	C, D, G	C, D (66.67%, -1)	C, D (66.67%, -1)
Alan, Sule, and Ipek Mumcu. "Nurturing Childhood Curiosity to Enhance Learning: Evidence from a Randomized Pedagogical Intervention."	NA NA	D, I, J, O	I, C (25%, -2)	I, D, J (75%, -1)
Average number of JEL categories	1.667	3.222	2.333	2.333
Average overlap with the final JEL categories	70.00 %		69.26 %	76.85 %
Average difference in the number of JEL categories	-1.167		-0.889	-0.889

Note(s): The listed articles appeared in the March 2024 issue (Vol. 62 No. 1) of the Journal of Economic Literature which was the current issue as of April 2024, available at: https://www.aeaweb.org/issues/755. The first number in parentheses under assigned JEL categories is the overlap with the final JEL categories and the second number is the difference in the number of assigned JEL categories compared to the final ones. Used AI systems: ChatGPT, https://chat.openai.com/; Copilot, copilot.microsoft.com; Gemini, https://gemini.google.com/. Data (incl. information about selected working papers) and full responses of AI systems are available in the Appendix/Supplementary Material

Source(s): Table by the author

Supplementary material

The supplementary material for this article can be found online.

Corresponding author

Jussi T.S. Heikkilä can be contacted at: jussi.heikkila@lut.fi

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**