

How does ChatGPT score in a Public Economics exam?

Carlos Contreras

jcontreg@yahoo.es

Universidad Complutense de Madrid (en excedencia). Departamento de Economía Aplicada, Pública y Política. España.

Recibido: 20 de noviembre de 2023

Acceptado: 27 de junio de 2024

Abstract

Nowadays it is almost impossible to discuss the future of the education industry without considering the influence of artificial intelligence (AI). This article examines the opportunities and challenges associated with the use of AI in education, with a particular focus on the ability of ChatGPT-3.5 to address a Public Economics exam that includes theoretical questions and practical exercises at the level of a course for undergraduate students. The algorithm has been demonstrated to be capable of attaining a B- grade evaluation in accordance with the established criteria. The study's conclusion is that this AI-based chatbot demonstrates remarkable capabilities in assisting educators in preparing materials for lectures, classes, and seminars, as well as providing tutoring support to students in Public Economics.

Key words: artificial intelligence, ChatGPT, Public Economics.

JEL Codes: A22; I20; H89

1. INTRODUCTION

The term "artificial intelligence" (AI) is used to describe technologies that are capable of performing tasks or functions that would usually require human intelligence or sensory abilities. This branch of computing science is concerned with the development of software algorithms and techniques that enable computers and machines to simulate human perception and decision-making processes, thereby enabling them to successfully complete tasks. In essence, AI is a technology that enables the construction of systems capable of thinking and acting in a manner analogous to humans, with the capacity to achieve defined objectives.

As in all economic and social areas, the advent of AI will have profound implications for higher education. The application of AI to education (AIEd) has been hailed as one of the most significant developments of the century¹, with the capacity to transform the educational landscape and influence the role of all involved stakeholders². Generative artificial intelligence algorithms with the ability to produce coherent and contextually relevant texts represents an important milestone in this process. The long-term implications of cognitive automation

driven by generative artificial intelligence extend beyond the field of education, with a revolutionary impact on research in most scientific disciplines.³

The objective of this paper is to assess the potential contribution of generative artificial intelligence to the teaching of Public Economics with a particular focus on the efficiency with which it can assist teachers in preparing lecture, class, and seminars materials. To this end, the paper assesses the performance of ChatGPT-3.5 by OpenAI in solving an undergraduate Public Economics exam. This study builds upon several previous papers that have demonstrated the ability of versions of ChatGPT to pass standardised tests in various fields. These include mathematics,⁴ medicine,⁵ law,⁶ physics,⁷ and economics.⁸

The algorithm was presented with ten theoretical questions and four practical exercises. The ChatGPT-3.5 version achieved a grade of B- in accordance with the selected evaluation criteria.

The remainder of the article is structured as follows: Section 2 examines the promises and challenges of the development of artificial intelligence in education. Section 3 addresses the implications of the emergence of generative AI. Section 4 outlines the methodology employed. Section 5 presents the results obtained. Appendix A shows the answers given by the algorithm. Appendix B provides a rationale for the scores given.

2. PROMISES AND CHALLENGES OF ARTIFICIAL INTELLIGENCE IN EDUCATION

Today it is almost impossible to think about the future of the education industry without considering the impact of AI. There is a consensus that AI will transform the industry by changing access to knowledge, teaching tools, approaches to learning and even the way teachers are trained⁹. AIED refers to the use of AI for learning, assessment of learning and other fundamental educational purposes, including management and administration. Several areas of educational practice have been penetrated by AI, including non-teaching tasks as timetabling, resource allocation, student tracking, or providing of reports on students to their parents/guardians.¹⁰

The integration of AI into various tasks in this industry is expected to result in an increasing complementarity between machine and human performance. Algorithms can be of significant utility in the provision of basic content and for the expeditious assessment of essays and tests. In this context, teaching assistants may be tasked with monitoring the grades assigned by AI algorithms and providing the human touch needed to motivate students and address non-academic learning issues. Experienced educators can concentrate on more complex pedagogical duties, such as devising novel instructional methodologies, offering feedback on oral and written communication, and cultivating an achievement-oriented classroom environment.¹¹

With regard to the direct interaction between IA algorithms and learners, it is worth mentioning a number of relevant arguments and assertions:

- i. The utilisation of AIED may facilitate a deeper comprehension of students' learning processes and enhance their learning experiences.
- ii. It is possible to utilise IA algorithms effectively in the context of student profiling for the purposes of admission and retention.

iii. In the context of Intelligent Tutoring Systems (ITS), educational robots are being employed in a variety of roles, including that of a teaching assistant, peer and co-learner, and companion. Furthermore, they are a fundamental component of some learning platforms.¹²

iv. ITS-type applications are employed both as the primary mode of instruction and integrated into teacher-led courses.¹³

v. Artificial intelligence (AI)-based predictive analytics algorithms can also assist in the identification of students who are at risk of not completing or passing their course. This allows universities to implement appropriate intervention strategies. The field of learning analytics employs the use of data mining techniques to identify patterns within large data sets, including those generated on the internet.

vi. The IEAd can provide just-in-time assessments that can be useful in facilitating and accelerating relevant adaptations of learning systems, with the aim of improving student performance and ensuring that curricula are aligned with the educational policies.¹⁴

vii. The utilisation of AIED applications has the potential to enhance the efficacy of educational processes, as it can facilitate the expeditious attainment of specified levels of proficiency by students.¹⁵

viii. AIED can facilitate the development of personalised learning systems. Given the diversity of learners in terms of their strengths and needs, AI-based adaptive learning platforms can be employed to enhance the learning process.¹⁶

ix. Automated AI-based Essay Scoring Systems can assess and provide feedback to students, while helping to manage teacher workload.¹⁷

x. The use of technology-based algorithms in the field of facial recognition has become a valuable tool in the proctoring of online examinations, as well as in the analysis of student behaviour.¹⁸

xi. AIED algorithms can also be used to facilitate connections between students and between students and their teachers. The consequence of this is an enhanced efficiency in the process of learning.¹⁹

There is a growing concern across society about the potential impact of the increasing use of artificial intelligence. Education is no exception to this trend. Among the main threats and vulnerabilities to the use of AI, the following have been highlighted:

i. The widespread use of AI applications can have a major impact on the education labour market. AI already has the capacity to replace many administrative and teaching assistants in higher education.²⁰ Artificial intelligence is also forcing a rethink of the role of the teacher, and even to consider the partial replacement of teachers by *virtual teacher-bots*.²¹

ii. The implementation of AI techniques in education could potentially have a detrimental impact on the relationship between teachers and students. It is essential to investigate AI approaches that empower teachers, in order to prevent the emergence of conflicting authority structures between staff, machines, companies, and students.²²

iii. AIED can result in students becoming overly reliant on online platforms and artificial teaching assistants.

iv. Intelligent Tutoring Systems have the potential to diminish the quality of higher education. Due to their complexity, the implemented applications may not perform as promised. Consequently, placing undue reliance on these systems could have the unintended consequence of undermining the level of excellence.

v. Furthermore, it is challenging for ITS algorithms to model and adapt to student behaviours, skills, and states of mind, which are often less structured and well-defined than those of traditional problem-solving.²³

vi. Confusing the appearance of intelligence with actual intelligence can lead to the mistaken belief that AI tools can do more than they can. No AI system is truly intelligent (including today's LLMs) because they don't really understand anything.²⁴

vii. In relation to AI-based predictive analytics, staff and students have raised concerns about the potential for misinterpretation of data, constant monitoring, lack of transparency, inadequate support, and the potential to inhibit active learning.²⁵

viii. It may be the case that the implementation of AI techniques results in an alteration of the content of programmes, leading to deviations from the intended educational policy.

ix. The widespread use of AI could result in an over-individualised approach to education. This would result in a lack of emotional intelligence and an inability to acquire social values, which are typically attained through conventional education.

x. For AI to be integrated into the university culture, stakeholders must be receptive to its adoption. This requires reliable algorithms and a widespread perception that the negative effects of its use will be mitigated.

xi. The existence of black boxes represents a structural weakness of artificial intelligence systems. Sometimes, it is difficult to explain the results generated in the context of AIED, which raises doubts about how the models make their inferences or the objective functions that they employ to evaluate educational performance.²⁶

The use of AI algorithms in higher education has the potential to deliver significant advances. However, there are also a number of challenges that must be overcome. As with other areas of AI, ethical considerations are of particular importance. It is imperative that universities develop robust policies and research agendas that take into account the ethical implications of implementing AIED systems. In this regard, the following key issues must be addressed:

i. The rise of AIED has led to a proliferation of interaction logs, which has generated a significant amount of data about students. While there is still debate about what should be considered personal data, it is evident that there are significant privacy concerns about the vast majority of the data being generated.²⁷

ii. Other areas of concern include the issue of ownership and rights over educational records, as well as consent for the use of the data.²⁸

iii. One of the key objectives of the AIED initiative is to enhance educational equality and narrow the attainment gap between different student groups. However, it is unclear to what extent this can be achieved, and there is a risk of perpetuating digital exclusion through algorithmic bias. Indeed, there is some debate as to whether the use of AIED may exacerbate inequalities.²⁹

iv. One challenge in the construction of AIED systems is the potential lack of transparency regarding the pedagogical assumptions, data on which their models are built, or the socio-cultural orientation of the curriculum.³⁰

v. Another concern regarding AIED systems is the potential lack of protection for human autonomy. This is defined as the user's ability to modify the system's operation or to be excluded from it when adverse effects may occur.³¹

vi. The presence of ideological and cultural biases in implemented AIED systems could, in practice, lead to changes in programmes that define what is important to know and how students should learn.³²

On another note, given the increasing dominance of AI in cognitive tasks in education, it is not unexpected that the question has arisen as to whether universities should change their methods. There is a consensus that not only mastery of content, but also the acquisition of non-cognitive skills play an important role in students' academic and future career outcomes.

Innovations in AI can adapt the classroom experience, allowing teachers to focus more on helping students develop these important skills.³³ In light of the future employability of their students, should universities focus their efforts on developing only human skills? Should universities pay less attention to the development of skills aimed at performing tasks that AI can do (even better than humans)? And if this is the case, one wonders what the uniquely human skills will be in the near future.

3. THE EMERGENCE OF GENERATIVE AI: DOES IT REPRESENT A CHANGE?

Generative artificial intelligence (GAI) systems can produce text, images, video, music, and other types of content, including machine code, as well as performing translations. One such system is the Generative Pre-trained Transformer (GPT).³⁴ GPT's strength lies in its ability to generate coherent and contextually relevant text in various natural language processing tasks, including language translation, text generation, and text completion. Using natural language processing (NLP) and a wealth of publicly available digital data, GPT models can read and produce human-like text in multiple languages and demonstrate creativity by writing anything from a paragraph to an entire article convincingly on countless topics.³⁵

Currently the versions of OpenAI are the most popular LLM on the market. There is a free version based on the Chat-GPT-3.5 model and a paid version called Chat-GPT-4. ChatGPT-4's public launch on 30 November 2022 has prompted reflection on the impact of generative pre-trained transformer systems on education.³⁶ Other chat engines include Microsoft's Bing (also based on GPT-3.5/4) which can browse the Internet in real time. Google's Bard, which is based on its PaLM-2 Bison model, is comparable to Bing in that it can search the web to include real-time information in its responses to user queries. It offers functionality at a similar level to ChatGPT-3.5. Other AI-based chat tools currently available on the market include Anthropic's Claude 2 and Meta's LLaMA 2 among others.

Although there is consensus that these systems have the capacity to transform current educational practice,³⁷ the educator community has expressed mixed feelings about their extraordinary ability to perform complex tasks. Those who are optimistic about the impact of GPT models argue that they have the potential to improve the quality of some of the various AI-based educational systems, such as those related to Personalized Tutoring, Automated Essay Grading, Interactive Learning, or Adaptive Learning. On the contrary, those who are skeptical about the prospect of using these models for educational purposes argue that their generalized use could lead to the appropriation of intellectual property and enable opportunistic students to engage in academic misconduct such as cheating and fraud. It has also been argued that this system does not solve the problem of the lack of human interaction that is common to algorithmic-based tutoring systems. They also remind us that GPT models (which are simply based on statistical patterns) still show a very limited understanding (if any at all). Similarly, a dialogue system based on GPT models has shown limited ability to generate contextually appropriate responses in a conversation and to personalize instruction. Furthermore, it has been pointed out that GPT-based tutoring systems still lack the ability to provide explanations adapted to the students' needs and, worse still, sometimes provide incorrect, inappropriate, or irrelevant answers (especially when the training data is not sufficiently relevant to the area being addressed). Finally, it has been argued that GPT does not solve the problem of biased results (as it still depends on the quality of training data).

There are essentially five types of potential deficiency that have been identified in the academic responses provided by the GPT, as shown in Table 1.

Table 1. Potential shortcomings in the responses provided by GPT

Lack of contextual understanding	Generative models, including GPT, show a lack of ability to understand context and situations. This can lead to the generation of contextually inappropriate responses.
Poor quality	Although the GPT system can provide correct answers that are well suited to the context of the question the quality of these resources may be poor due to outdated, superficial, or incomplete content.
Risk of hallucinations	The GPT tool occasionally generates fabricated answers that are irrelevant to the inquiry or founded on non-existent empirical data.
No response	Sometimes the GPT-based system is unable to provide quantitative answers to certain questions, either because of insufficient data or because the data has been processed incorrectly.
Fake references	GPT can generate references to fictitious articles, resulting in reduced reliability and robustness.

The objective of this article is not to provide information about whether a specific version of ChatGPT can pass a Public Economics examination. Instead, the aim is to utilise the insights gained from its responses to evaluate four key questions:

- a. Can teachers utilise this tool to enhance their efficiency when preparing material for lectures, exercises and seminars?
- b. How should the role of essays and non-classroom exams be reviewed in the context of IA tools?
- c. Can teachers rely on this AI tool for tutoring?
- d. Should students be encouraged to use AI-based chat in their learning process so that their ability to formulate prompts is improved?

One of the key responsibilities of a teacher is to prepare practical exercises, suggest essays, and correct students' work. If there is sufficient confidence in the capabilities of tools such as AI-based chat, their use should result in a reduction in the teaching load for these tasks, an increase in the quality and variety of materials, and an increase in the time available for face-to-face tutoring.

It has been demonstrated that if a question is entered into an AI-based chat room, an essay will be generated in a matter of seconds. This raises the question of whether the traditional approach to university writing is now obsolete³⁸. This is not necessarily the case. It is evident that essays will continue to be an integral component of students' academic work. However, the advent of AI-based chat rooms necessitates a re-evaluation of the nature of the questions posed. In light of these considerations, it is evident that the level of complexity of questions should be significantly higher. The objective of the tests is not merely to assess students' writing abilities, but also to ascertain their capacity to integrate theoretical concepts, resolve apparent contradictions, and apply their acquired knowledge to practical scenarios.

In the context of essay writing assignments and open-book examinations, the primary objective of monitoring has traditionally been to ensure that students are the authors of their work. The advent of AI chatbots has introduced a new risk factor into the equation, namely

the possibility of cheating. Nevertheless, the development of AI-based software also enables educators to utilise plagiarism detection software to ascertain whether submitted material has been copied.³⁹

It has been argued that the use of AI tools can make students too dependent on technology and prevent them from developing their own problem-solving skills, while at the same time stifling their creativity. However, we should not forget that students will inevitably enter a world where AI is commonplace.⁴⁰ According to the human capital theory developed by Gary Becker in the 1960s, investing in the education and training of individuals can help increase the productivity of workers. Given that increased productivity is precisely the most expected outcome of the use of artificial intelligence tools, training university students in AI systems seems a relevant goal. And this includes the acquisition of skills in how best to query AI-based chats. The critical analysis of texts generated by artificial intelligence (AI) tools can facilitate the enhancement of writing skills. Consequently, in addition to the continued requirement for students to write essays, a potential aspect of the university education process could be the evaluation of not only the outcome, but also the questions asked and the iterative process followed in the use of text-generative AI systems. As with any computer system, the quality of the output of an AI system depends on the quality of the input data. The process of developing effective instructions for generative AI systems (called 'prompt engineering') should be considered as part of the learning process. To improve the quality of content generated by LLM systems, it is essential to provide them not only with clear questions, but also with contextual and stylistic instructions. For optimal LLM results, it is recommended to be iterative and patient. When using GPT software to teach economics, teachers can think of this AI-based tool not as a database, but as a large collection of economists, historians and scientists to whom questions can be posed.⁴¹

The utilisation of artificial intelligence (AI) tools is set to become ubiquitous across the entire spectrum of professional life. Today's university students will soon enter a world of work in which the way they use AI tools will determine not only their productivity but will also affect the ethical standards of their work. The utilisation of AI tools in the educational setting represents an ideal opportunity to educate students about the ethical implications of such technology. All teachers, regardless of their subject area, have the potential to integrate these objectives into their educational methodology.

Intelligent tutoring systems have the potential to tailor lessons to the specific needs of learners. By offering learning experiences that are tailored to the unique needs and interests of each student, personalised learning platforms can enhance student engagement and motivation.⁴² However, in the current state of ChatGPT, educators may be reluctant to use this artificial intelligence tool in tutoring tasks because of the possibility that it may provide answers that contain hallucinations. While human-written texts show a strong correlation between authoritative style and reality-based content, LLM-generated texts can sometimes appear to be of high quality in style but lacking in reliability in content.⁴³ It is clear that this risk exists, but the use of AI-based chatbots for academic tasks can help students develop critical thinking skills and the ability to be alert to failures in the output of AI-based systems. Such abilities will prove invaluable in their future careers.

4. METHODOLOGY

The methodology employed involved posing a series of questions to Chat GPT-3.5⁴⁴ on topics of the subject Public Economics. This course is usually taught as an elective in the

second or third year of the Economics degree programme. The Public Economics course currently consists of analysing various decisions involving the public sector and their consequences using microeconomic tools. It could be said that it is a course in microeconomics applied to the field of taxation and public spending. The usual skill requirements for Public Economics include courses in Mathematics and Microeconomics. In some universities this subject is taught under names such as Economics of the Public Sector or Public Finance. Other universities offer part of the subject content under courses called Economic Policy Analysis, Economics of Tax policy, Economics of Inequity, etc. Regarding the syllabus, the subject of Public Economics typically covers topics such as Theories of Public Sector, Market failures (as Externalities, Public goods or Imperfect competition), Public Choice and voting, Commodity taxation, Income taxation, Optimal taxation, Tax evasion, Fiscal Federalism, Redistribution policies, Public expenditure, and Fiscal Policies among others. Some of the most commonly used textbooks at the undergraduate level include Stiglitz and Rosengard (2015), Kennedy (2012), Gruber (2005), Hyman (2005), or Rosen and Gayer (1995). Texts by Hindriks and Myles (2013) and Myles (2005) are usually used for intermediate level. In the Master of Science in Economics and PhD in Economics degrees, Atkinson and Stiglitz (2015) and Auerbach *et al.* (2013) are commonly used.

The test designed consists of ten theoretical questions and four exercises on a range of Public Finance topics. The questions were posed to ChatGPT-3.5 during the month of November 2023. The prompts that are common to all questions are presented in Table 2.

Table 2. Common part of prompts

Context	Instructions fine-tuning
The question pertains to a topic covered in a Public Economics examination.	I am seeking an answer in the style of a second- or third-year undergraduate student.

In the case of a human examinee, it would normally be requested that the theoretical questions posed be answered within an average response time of 15-20 minutes. The number of words typically written by an economics student depends on several factors, such as the student's writing speed, familiarity with the subject and the type of questions asked but is generally between 400 and 600 words. Taking this into account, an additional instruction was incorporated into the prompt provided to ChatGPT-3.5, specifying a maximum number of words for each question. This information is presented in Table 3.

Table 3. Instructions in relation to the length of answers

Question	1	2	3	4	5	6	7	8	9	10
Word limit	450	450	575	525	500	600	300	450	450	525

The prompts for the theoretical questions are presented in Table 4.

Table 4. Theoretical questions

	Topic	Question
Q1	Public sector growth theories	There has been a significant expansion of the public sector in several countries over the last century. Public economics has proposed several theories to clarify the reasons behind this increase. A comprehensive interpretation would include components from all hypotheses. Cite some of the prominent theories and their principal arguments.
Q2	Public goods and club goods	What similarities and differences exist between public goods, local public goods, and club goods?
Q3	Preference revelation mechanisms	A challenge for the government in making efficient decisions regarding public goods provision is their inadequate understanding of consumers' preferences and their willingness to pay for such goods. What mechanisms has economic theory developed to incentivise consumers to disclose their true valuation of public goods?
Q4	Tragedy of the commons	What is the phenomenon commonly referred to as the 'tragedy of the commons'? How can this issue be resolved?
Q5	The Coase theorem	Under what conditions can market forces generate an efficient allocation of resources in the presence of externalities without government intervention?
Q6	Asymmetric information	What strategies and actions can be implemented to alleviate inefficiencies in resource allocation resulting from information asymmetries?
Q7	Public Choice Voting	When faced with a choice between two options, can a collective decision-making process meet the criteria of Anonymity, Neutrality, Decisiveness, and Positive Responsiveness?
Q8	Ramsey Rule for optimal commodity taxation	The conventional understanding of an established optimal tax principle is that commodities ought to be taxed in a manner that decreases the compensated demand for each good to the same extent in relation to the pre-tax state. Who can be credited for creating this principle and what are the difficulties that arise from its immediate implementation?
Q9	Income & commodity taxation	When both income and commodity taxes can be used together as optimal taxation methods, how does the separability of the utility function impact the outcome?
Q10	Tiebout hypothesis	What assumptions are necessary to claim that consumers/tax payers, by selecting the jurisdiction in which to reside (i.e., voting with their feet), are demonstrating their preference for local public goods and thus achieving an efficient and stable equilibrium?

The questions of the exercises are shown in table 5.

Table 5. Exercises

	Topic	Statement
E1	Public goods provision and Lindahl prices	<p>The production cost for a public good (G) which benefits 3 individuals is given by $C_G = 50G + 20$. Marginal utilities for individuals A, B and C are given by: $MU(A)_G = 20 - G$; $MU(B)_G = 40 - G$; $MU(C)_G = 80 - G$.</p> <p>Considering that each consumer pays 1/3 of the marginal cost of G:</p> <p>a) Determine whether the use a <i>median voter rule</i> produces an efficient allocation for G.</p> <p>b) Calculate the <i>Lindahl prices</i> for the socially efficient level of G.</p> <p>d) What would be the optimal level of provision in the case of a private good X such that, $C_X = 50X + 20$, if marginal utility for individuals A, B and C are given by $MU(A)_X = 20 - X$; $MU(B)_X = 40 - X$; $MU(C)_X = 80 - X$.</p>
E2	Externalities	<p>To produce X, a given firm has total private costs given by the function $TPC = X^2 + 10X + 50$. The equilibrium price in the perfectly competitive market is $P_x^* = 250$. The firm also incurs external costs amounting to $EC = X^2 + 2X$ while producing X. Calculate: i) the quantity produced assuming no government intervention; ii) the optimal quantity produced when internalizing the negative externality with a Pigouvian tax; iii) the tax on each unit of output. Compare the profits generated by the firm in the scenario of government involvement and without it.</p>
E3	Vote trading	<p>Consider that three parliamentary groups, each representing one of the regions A, B, and C, are of equal size. There are currently two projects to decide on: G1, which is an airport that will be constructed in region A, and G2, a highway that will be built in region B. Funding for each project will be split evenly amongst the regions, with each providing one-third. Both projects will cost 900 million euros each but are predicted to generate 1.2 billion euros in benefits for the regions in which they are completed. Questions: a) What is the projected result of a vote if vote trading is prohibited? b) Can national welfare be augmented through vote trading? c) Is there a possibility of exploiting a minority? d) Propose a potential compensation strategy to avert this outcome.</p>
E4	Excess burden of an <i>ad valorem</i> tax	<p>The demand curve for good X is given by $X_D = 4000 - 4P$ and the supply curve by $X_S = -600 + 2P$. The government decides to introduce an <i>ad valorem tax</i> on the consumption of this good, so that $t = 5\%$ of the expenditure on good X. You are asked to calculate: a) Price and equilibrium quantity before and after tax; b) Who pays the tax?; c) The elasticities of demand and supply and their role in the distribution of the tax burden; d) The excess burden (using the Harberger triangle); e) The government revenue.</p>

The responses offered by ChatGPT-3.5 to the queries and exercises are detailed in Appendix A. The answers were scored according to the criteria shown in Table 6:

Table 6: Assessment criteria

	Criteria	Weighting of each criterion	Weighting of each block
Theoretical queries	Contextual understanding	25%	
	Quality	50%	
	Completeness	25%	
	Total	100%	40%
Exercises	Approach to the solution	50%	60%
	Accuracy of the numerical results	50%	
	Total	100%	
		Total	100%

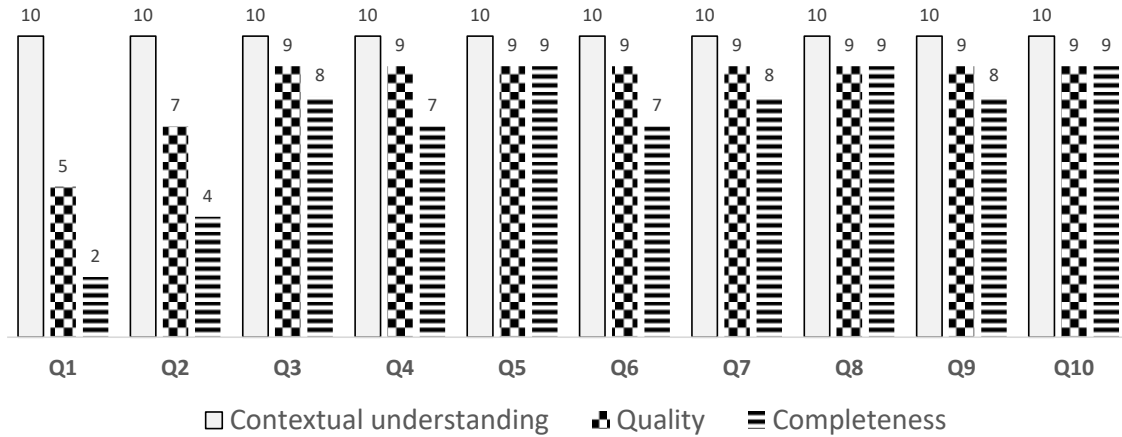
5. RESULTS

The provided answers demonstrated an exemplary fit to the theoretical questions in all instances. The quality of the responses was generally of a high standard, with scores ranging from A to A+, except for the first two questions. In the initial question pertaining to the explanatory theories of public expenditure growth, several hypotheses were omitted, including those related to the displacement effect, fiscal illusion, government agency, rent-seeking, budget-setting, and unbalanced growth. In the second question, regarding public goods and club goods, the text would require consideration of common property resources that display consumption rivalry but cannot be excluded. Furthermore, a comment on the *congestion effect*, which highlights the partial rivalry of consumption from a given level of demand, would have been appropriate. Additionally, commentary on the variable usage-based surcharge rule and the *two-part tariff* would be valuable. And finally, it would be beneficial to differentiate between small and large clubs.

The explanations justifying the assessment are given in Tables B1 to B10 in Appendix B.

Figure 1 shows how the theoretical questions scored according to the chosen criteria.

Figure 1. Marks for theoretical questions



Source: own elaboration.

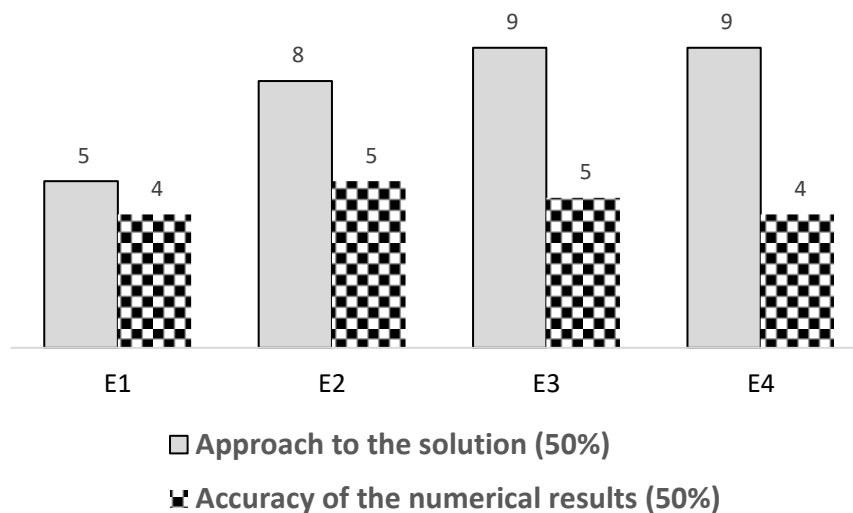
In addition to the examination mark, the extent to which answers contained hallucinations and/or false bibliographical references was also assessed. Note that the prompts deliberately did not require references to be given. The results were very encouraging in this respect. See Table 7.

Table 7. Was the response accompanied by hallucinatory content or false references?

Question	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Hallucination	No	No	No	No	No	No	No	No	No	No
False bibliographical reference	No	No	No	No	No	No	No	No	No	No

Regarding the Exercises, the distribution of marks was more heterogeneous. In general, the approach taken by ChatGPT-3.5 to address the Exercises was effective. However, in Exercise 1 on the optimal provision of public goods, the answer failed to take into account the fact that, according to the correctly revealed preferences, only two of the three individuals were willing to contribute to the production of the public good. Furthermore, the answer failed to consider the fact that, at the given marginal cost that determines the price of the private good, demand arises solely from one of the consumers. With regard to the numerical results, ChatGPT-3.5 provided an answer in two out of the four exercises (Exercises 1 and 2). In Exercise 1, some of the results were not entirely accurate due to an error in the approach previously outlined. In Exercise 2 the results were also erroneous due to an incorrect marginal cost derivative. Furthermore, not all the requisites' results were obtained. In Exercises 3 and 4, for reasons that are unclear ChatGPT-3.5 did not utilise the data provided to generate the numerical results. However the approach employed was correct in both cases.

Figure 2. Marks for Exercises

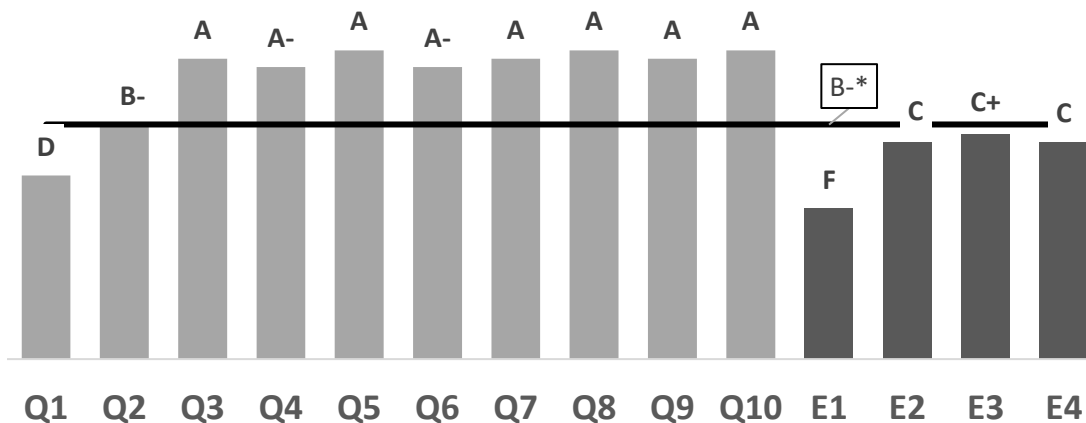


Source: own elaboration.

The explanations justifying the assessment are given in Tables B11 to B14 in Appendix B.

The weighted result of ChatGPT-3 responses, as shown in Figure 3 is a grade of B-.

Figure 3: Scoring of ChatGPT-3 responses



* Each theoretical question carries a weight of 4%, while each exercise is weighted at 15%.

6. CONCLUSIONS

ChatGPT-3.5 displays a good level of competence in the field of Public economics at the undergraduate level, to the extent that it is able to pass an exam with a grade of 'B-'. The algorithm was able to provide accurate and contextually appropriate responses to all theoretical questions. In the context of solving exercises, the results were less encouraging. Although ChatGPT-3.5 demonstrated an ability to provide correct approaches, the numerical results showed varying levels of quality. The algorithm demonstrated a consistent ability to interpret the questions posed and correctly apply the mathematical formulae. Nevertheless, in two exercises, the algorithm failed to utilise the provided data to generate results.

It has been argued that one of the current disadvantages of utilising the AI tool is that it lacks the capacity to comprehend context, read tone and emotions, address complex topics, or create personalised lesson plans. It has been pointed out that LLM models exhibit limitations in their ability to perform tasks requiring complex, multi-step reasoning.⁴⁵ In this context, it is accurate to state that the questions included in the examination were of a straightforward nature. It is possible that this is the reason for the satisfactory outcome.

The utilisation of generative AI can be a valuable asset in the automation of minor tasks, commonly referred to as 'micro-tasks', that lectures and teaching assistants have to perform in the preparation of their lectures and classes. Based on the ability of ChatGPT-3.5 to deal with basic questions of Public Economics, it could be argued that the integration of generative AI into teachers' operational procedures can significantly improve their productivity.

In a similar vein, the level of Public Economics demonstrated by ChatGPT-3.5 at the undergraduate level allows us to conclude that teachers can use this tool to gamify the classroom and propose games that can be solved individually or in teams. This strategy has the potential to foster students' engagement with the subject matter.

One aspect of the research into the use of AI tools is the extent to which they can assist teachers in implementing more personalised education, allowing students to learn at their own pace. The findings of this study indicate that the utilisation of ChatGPT-3.5 capabilities in the domain of Public Finance enables the expedient development of instructional materials that facilitate a more personalised educational experience.

It is the responsibility of contemporary educators to prepare students for the novel work environment that is characterised by the proliferation of artificial intelligence use. In the specific case of Economics education, the consideration of current events that can be analysed using ChatGPT-3.5 can assist students in developing a better understanding of the relationship between theoretical concepts and the real-world context of the global economy.

The utility of AI-based assistants in providing responses to exercises and essays in the field of Public Economics (as in many other scientific disciplines) will continue to expand in the future. This will permit educators to concentrate on tasks where they exhibit comparative advantage when preparing their materials (such as questioning, evaluating and reviewing the content generated by these algorithms). In this way, they will have more time to support students in more complex areas of the subject.

The development of students' comprehensive intellectual abilities, including critical thinking, fluent writing, sound logical reasoning, precise linguistic expression and agile thinking, has consistently been regarded as the fundamental mission of education. The advent

of AI-based tools in our daily lives does not negate this mission; rather, it compels educators to reinforce their students' critical thinking abilities and to cultivate their creativity as a means of differentiating themselves from robots.

Appendix A

Table A1: ChatGPT-3.5 answer to Question 1

The growth of the public sector in various countries over the past century can be attributed to a combination of factors and theories in public economics. While a comprehensive explanation would indeed involve elements from multiple theories, here are some key theories that have been advanced to explain this phenomenon:

Wagner's Law: Proposed by German economist Adolph Wagner, this theory suggests that public sector growth is a natural consequence of economic development. As a country's income and population grow, its citizens demand more public goods and services, such as education, healthcare, and infrastructure. Thus, the government's role in providing these services expands over time to meet these increasing demands.

Market Failure Theory: This theory argues that the government's involvement is necessary to address market failures, such as externalities, public goods, and imperfect competition. In cases where the private market cannot efficiently allocate resources or provide certain services, the public sector steps in to correct these failures and ensure economic stability.

Redistribution Theory: Public sector growth can also be explained by the desire to reduce income inequality and achieve wealth redistribution. Governments implement policies and programs, such as progressive taxation and social welfare systems, to transfer wealth from the affluent to the less privileged, thereby addressing societal disparities.

Public Choice Theory: Developed by economists like James Buchanan and Gordon Tullock, this theory focuses on the self-interested behavior of individuals within the government. It argues that public officials and bureaucrats may pursue policies that benefit themselves or special interest groups, leading to government expansion as they seek to maximize their own welfare.

Bureaucratic Expansion Theory: According to this theory, the growth of the public sector is driven by the inherent desire of bureaucracies to expand and increase their influence. Bureaucrats may seek to expand their responsibilities and budgets to secure their own job security and prestige.

Political Business Cycle Theory: This theory suggests that politicians may expand the public sector to gain political support or re-election. In this view, government spending and policies may be influenced by electoral cycles, with politicians enacting popular but potentially unsustainable programs to secure votes.

Globalization and Technological Change: In the modern context, globalization and rapid technological advancements have also played a role in the growth of the public sector. Governments have had to adapt to new challenges and opportunities presented by globalization, such as regulating international trade, addressing transnational issues, and responding to technological disruptions.

It's important to note that public sector growth is often influenced by a complex interplay of these theories and is context specific. Different countries may experience growth in the public sector for various reasons, and a comprehensive explanation typically involves considering a combination of these theories to understand the underlying dynamics.

Table A2: ChatGPT-3.5 answer to Question 2

Common Features:

Non-excludability: Public goods, local public goods, and club goods are all non-excludable. This means that it is difficult or impossible to exclude individuals from consuming these goods once they are provided. This is in contrast to private goods, which can be easily restricted to those who have paid for them.

Non-rivalry: These goods are also non-rivalrous. Consumption of these goods by one individual does not diminish their availability for others. In other words, one person's use of the good does not preclude someone else from using it simultaneously or in the future.

Differences:

Public Goods: Public goods are characterized by both non-excludability and non-rivalry. Examples include national defence, clean air, and street lighting. They are provided by the government because private markets often fail to supply them efficiently. The free-rider problem is a significant issue, as people can benefit from public goods without paying for them.

Local Public Goods (or Local Public Expenditures): Local public goods also exhibit non-excludability, meaning it is challenging to prevent individuals from benefiting once they are provided. However, they may exhibit rivalry at the local level. These goods are typically financed and provided by local governments or authorities and are meant to serve a specific geographic area. Examples include local parks, street cleaning, and public schools. Local public goods can lead to what's called the "spillover" problem, where people from neighbouring areas may also benefit from these goods even if they do not directly pay for them. Local governments must address the challenge of financing and providing these goods effectively.

Club Goods (or Excludable Commons): Club goods are excludable but non-rivalrous. This means that access to the good can be restricted, and consumption by one individual does not reduce its availability to others. Club goods are typically provided by private entities, and individuals must join or "pay to play." Examples include cable television, private recreational clubs, and toll roads. They are often associated with subscription or membership models, where individuals who pay can access the good while non-paying individuals are excluded.

In summary, while public goods, local public goods, and club goods all share the feature of non-excludability, they differ in terms of rivalry and the means by which they are provided and financed. Public goods are non-excludable and non-rivalrous, often necessitating government provision. Local public goods are non-excludable and may exhibit some level of rivalry, and they are typically provided by local governments. Club goods are excludable but non-rivalrous and are provided by private entities with exclusion mechanisms like memberships or subscriptions. These distinctions are important in public economics for understanding how goods are allocated, financed, and regulated.

Table A3: ChatGPT-3.5 answer to Question 3

In public economics, economists have developed several mechanisms and strategies to incentivize consumers to reveal their true preferences and willingness to pay for public goods. The primary challenge is to address the free-rider problem, where individuals may understate their preferences or willingness to pay for a public good to avoid contributing more. Here are some of the main mechanisms constructed by economic theory to tackle this issue:

Voting Mechanisms:

One of the most common ways to reveal preferences for public goods is through democratic processes like voting. Individuals cast their votes in favour of public projects or policies, which signals their support and willingness to contribute. However, voting mechanisms may have limitations, such as strategic voting or the "tyranny of the majority." To address these issues, mechanisms like supermajority rules or constitutional safeguards may be employed.

Revelation Mechanisms:

Revelation mechanisms, such as the Vickrey-Clarke-Groves (VCG) mechanism, are designed to elicit truthful information from individuals. In a VCG auction, participants reveal their true valuations for a public good or project, and the mechanism determines both the optimal allocation and the contributions required from each participant. VCG mechanisms are incentive-compatible, meaning that it is in the best interest of individuals to reveal their true valuations. They are theoretically efficient and can be used for public goods provision.

User Fees and Pricing Mechanisms:

By charging individuals for their use of a public good, governments can approximate consumers' willingness to pay. For example, tolls on roads or entrance fees to public parks allow users to reveal their preferences by choosing whether or not to use the service and how frequently. Properly designed user fees can help cover the costs of maintaining public goods and improve the allocation of resources.

Taxation and Subsidies:

Taxation and subsidies are powerful tools for governments to collect revenue for public goods. By designing tax systems that take into account individuals' ability to pay, governments can indirectly gauge

their willingness to support public goods.

Subsidies, on the other hand, can encourage the consumption of certain public goods by reducing the cost for users.

Public Opinion Surveys and Contingent Valuation:

Surveys and contingent valuation methods involve directly asking individuals about their preferences and willingness to pay for specific public goods. These approaches can provide valuable insights into public preferences. However, surveys are often limited by issues related to hypothetical bias and the accuracy of stated preferences.

Information Campaigns and Public Engagement:

Informing the public about the benefits and costs of public goods can influence their willingness to reveal their true valuations. Well-structured information campaigns can help individuals make more informed decisions regarding public goods. Engaging the public in the decision-making process through consultations, town hall meetings, or referendums can also elicit preferences and encourage transparency.

Bidding and Auction Mechanisms:

Auctions, such as the ascending-bid (English) and descending-bid (Dutch) auctions, can be used to allocate public goods like radio spectrum licenses or natural resource concessions. Participants reveal their valuations through their bids, and the highest bidder wins the right to use the resource. Auctions can promote efficiency and generate revenue for the government. Each of these mechanisms has its advantages and limitations, and the choice of mechanism often depends on the specific context, the nature of the public good, and the preferences of the stakeholders involved. Combining multiple mechanisms may also be a practical approach to improve the accuracy of revealed preferences for public goods.

Table A4: ChatGPT-3.5 answer to Question 4

The "Tragedy of the Commons" is a concept in economics and environmental science that describes a situation in which shared or common resources are overused and depleted due to the self-interest of individuals, ultimately leading to a negative outcome for all. This dilemma arises when multiple individuals, each pursuing their own self-interest, deplete a commonly held resource, resulting in environmental degradation or resource exhaustion. The tragedy of the commons is a classic example of a collective action problem.

The tragedy of the commons can be illustrated with the following example: Imagine a common pasture that multiple farmers can use to graze their livestock. Each farmer's incentive is to maximize their herd size, as the benefit of adding more animals goes directly to them. However, when all farmers do this and the pasture is overgrazed, it leads to soil erosion, reduced forage quality, and, ultimately, a degraded and unusable common resource.

Solving the tragedy of the commons typically involves implementing strategies and mechanisms to address the overuse and degradation of shared resources. Several approaches can be employed:

Property Rights and Privatization: One way to mitigate the tragedy of the commons is to establish clear property rights or ownership over the resource. When individuals have ownership or stewardship over a resource, they have an incentive to manage it sustainably, as they bear the costs of depletion. Privatization of common resources, when feasible, can lead to more responsible and sustainable resource management.

Regulation and Government Intervention: Governments can implement regulations, rules, and policies to manage and conserve common resources. This might include setting catch limits for fisheries, imposing quotas for resource use, and enforcing penalties for overuse. Government agencies can monitor and enforce compliance with these regulations.

Cooperative Management and Community-Based Solutions: Communities and resource users can form cooperatives or associations to collectively manage and protect common resources. By establishing rules and norms for resource use, participants can better coordinate and prevent overexploitation. Cooperative management often works well for small-scale, localized resources.

Market-Based Solutions: Tradable permits or cap-and-trade systems can be introduced to allocate resource usage rights. This approach allows individuals or organizations to buy, sell, or trade permits for resource use. By putting a price on resource use, it encourages efficient allocation and conservation.

Technological Innovation: Developing and implementing new technologies and techniques for sustainable resource management can help mitigate the tragedy of the commons. For example, advanced

monitoring systems, precision agriculture, and sustainable fishing practices can enhance resource sustainability.

Education and Public Awareness: Increasing awareness about the consequences of overuse and the benefits of resource conservation can lead to more responsible behavior by resource users. Environmental education and outreach programs can help individuals understand the long-term costs of resource depletion.

It's important to note that the solutions to the tragedy of the commons can vary depending on the specific resource, its scale, and the social, cultural, and economic context. In some cases, a combination of strategies may be necessary to effectively address the issue. Sustainable resource management is crucial to ensure that shared resources can be used and enjoyed by present and future generations without leading to their depletion or degradation.

Table A5: ChatGPT-3.5 answer to Question 5

In the presence of externalities, which are spillover effects of economic activities on third parties that are not directly involved in the transaction, market forces may not lead to an efficient allocation of resources. Externalities can be positive (benefits accruing to third parties) or negative (costs imposed on third parties). However, under certain conditions, market forces can generate an efficient allocation of resources without the need for government intervention. These conditions are often referred to as Coasean or Coasian solutions, named after the economist Ronald Coase, who developed this framework. The following conditions are essential for market forces to address externalities efficiently:

Clearly Defined Property Rights: In a Coasian solution, well-defined and transferable property rights are crucial. Property rights determine who has control over a resource or activity. When property rights are clearly established, individuals can negotiate and make voluntary agreements about the use of those rights.

Low Transaction Costs: Transaction costs refer to the costs associated with negotiating, reaching agreements, and enforcing contracts. To achieve efficient outcomes without government intervention, transaction costs should be low enough to facilitate bargaining among parties affected by the externality. Low transaction costs enable parties to negotiate and come to agreements that internalize the externalities.

Rationality and Information: The parties involved in the externality must be rational and well-informed. They should understand the consequences of their actions and be willing to engage in negotiations that benefit all parties involved.

Mutual Agreement: In the absence of government intervention, the parties affected by externalities must voluntarily agree on a solution. This means that negotiations and agreements should be consensual and not imposed by a third party.

Absence of Market Power: Market power, which is the ability of a single party or a few parties to influence market outcomes, should be minimal. Market power can hinder negotiations and lead to inefficient outcomes.

Symmetry and Reciprocity: To reach efficient agreements, there should be symmetry in the parties' ability to affect each other with externalities. That is, the parties should be equally able to cause harm or benefit to one another. Reciprocity in negotiations can help achieve equitable solutions.

When these conditions are met, individuals or organizations can negotiate and reach agreements that internalize the externalities. This means that they consider the costs and benefits of their actions on third parties and adjust their behavior accordingly. For instance, if a factory is emitting pollution that negatively affects a nearby community, the factory owner and the affected community can negotiate to reduce emissions to an optimal level that minimizes harm while maintaining the factory's operation.

It's important to note that the Coasean approach does not always work in practice, especially when transaction costs are high, property rights are unclear, or negotiations are complicated. In such cases, government intervention through regulations, taxes, or subsidies may be necessary to address externalities and achieve an efficient allocation of resources. The choice between relying on market solutions or government intervention depends on the specific circumstances and the relative effectiveness of each approach.

Table A6: ChatGPT-3.5 answer to Question 6

Information asymmetry refers to a situation in which one party in an economic transaction has more or

better information than the other party, leading to potential market inefficiencies and misallocation of resources. Several strategies and actions can be taken to mitigate the problems created by information asymmetries:

Transparency and Disclosure: Increase transparency by requiring the party with more information to disclose relevant details to the other party. For example, in financial markets, publicly traded companies are required to disclose financial statements, performance metrics, and other material information to investors.

Regulation and Standardization: Governments can impose regulations and industry standards that compel businesses to provide certain information to consumers or other market participants. For instance, nutritional labeling on food products provides consumers with standardized information.

Certification and Third-Party Verification: Independent third-party organizations can provide certification and verification services. These organizations evaluate and confirm the quality, safety, or reliability of products or services, providing consumers with trusted information. For example, the Fair Trade certification indicates that certain social and environmental standards have been met.

Consumer Education and Financial Literacy: Empower consumers with information and financial literacy to better understand products and services. This can help consumers make more informed decisions and reduce the impact of information asymmetry.

Market Competition: Encourage competition in markets, as it can be a self-regulating mechanism. When multiple firms offer similar products or services, consumers have more options and can choose the providers that offer better information and value.

Warranties and Guarantees: Businesses can offer warranties and guarantees to mitigate concerns arising from information asymmetry. These assurances provide consumers with a form of insurance against potential product defects or problems.

Licensing and Credentialing: In professional services and industries, requiring practitioners to obtain licenses or credentials can help ensure that consumers receive services from individuals with a certain level of competence and expertise.

Online Reviews and Ratings: Utilize online platforms and communities where consumers can share their experiences and rate products and services. This user-generated content can help prospective buyers make more informed decisions.

Escrow Services: In financial transactions, the use of escrow services can protect both buyers and sellers. Funds are held by a neutral third party until agreed-upon conditions are met, reducing the risk of fraud or non-delivery.

Government Intervention: In cases where information asymmetry is severe or where market forces alone are insufficient, government intervention may be necessary. This can take the form of regulations, consumer protection laws, and enforcement agencies to ensure fair and transparent practices.

Financial Intermediaries: Financial intermediaries, such as banks and insurance companies, often bridge the information gap between lenders and borrowers. They assess the risk and creditworthiness of borrowers and provide information to investors, reducing the adverse selection problem.

Contractual Safeguards: Use contractual arrangements, such as non-disclosure agreements, to protect proprietary information when entering into agreements with other parties. These contracts can help safeguard valuable information.

Insurance: Some risks associated with information asymmetry can be transferred through insurance. For example, errors and omissions insurance can protect professionals from liability in cases of unintentional errors or omissions in their services.

Data Analytics and Technology: Advances in data analytics and technology can help identify patterns and anomalies, which can be useful in detecting fraudulent or deceptive practices and managing information-related risks.

Mitigating the problems of information asymmetry often involves a combination of these strategies and actions, depending on the specific context and industry. The goal is to reduce the adverse effects of asymmetric information and facilitate more efficient and fair resource allocation in the market.

Table A7: ChatGPT-3.5 answer to Question 7

In public economics, Arrow's Impossibility Theorem, developed by economist Kenneth Arrow,

demonstrates that it is impossible to have a collective decision-making process that satisfies all of the following four requirements simultaneously when choosing among more than two options:

Anonymity: Anonymity means that each individual's preferences should be treated equally, without regard to their identity. In other words, the decision-making process should not favour or discriminate against any specific individual.

Neutrality: Neutrality implies that the decision-making process should be neutral or impartial regarding the available options. It should not favor one option over another based on any inherent characteristic of the options.

Decisiveness: Decisiveness means that the collective decision-making process should always yield a clear and unambiguous outcome. There should be a definitive choice among the available options, rather than a tie or indecisiveness.

Positive Responsiveness (or Positive Responsiveness to Individuals): Positive responsiveness states that if every individual prefers one option over another, the collective choice should also reflect this preference, i.e., the majority preference should be respected.

Arrow's Impossibility Theorem, which was formulated in 1951, proves that it is not possible to design a collective decision-making process that satisfies all four of these conditions simultaneously when choosing among more than two options. This theorem highlights the inherent challenges and complexities in aggregating individual preferences into a collective decision in a fair and consistent manner.

In practice, real-world decision-making processes, such as voting systems and democratic institutions, often involve trade-offs and compromises among these conditions. Various voting systems, like majority rule, ranked choice voting, and others, aim to strike a balance between these criteria, recognizing that perfect satisfaction of all four is unattainable. The choice of a decision-making process may depend on the specific context, objectives, and values of the society or organization involved.

Table A8: ChatGPT-3.5 answer to Question 8

The optimal tax rule you are referring to is associated with the work of Frank Ramsey, an early 20th-century British economist. It is known as Ramsey's "Inverse Elasticity Rule" or "Ramsey Rule" for optimal commodity taxation.

Ramsey's rule suggests that to achieve an efficient allocation of resources, taxes should be levied on goods in such a way that the compensated (or compensated price) demand for each good is reduced in the same proportion relative to the pre-tax position. In other words, the tax rates should be set inversely proportional to the price elasticities of demand for each good. This rule aims to minimize the deadweight loss (or excess burden) associated with taxation, which occurs when the tax system distorts consumer choices and market efficiency.

While Ramsey's rule provides valuable insights into optimal taxation, its direct application can lead to several problems and challenges:

Implementation Complexity: Determining the price elasticities of demand for various goods can be challenging, and it may require detailed data and analysis. In practice, obtaining accurate elasticity estimates for all goods is often difficult.

Distributional Consequences: Ramsey taxation, which aims for efficiency, may not address concerns related to income inequality or fairness. The tax burden under Ramsey's rule may fall disproportionately on goods that are necessities for low-income individuals.

Substitution Effects: Ramsey pricing can lead to significant substitution effects. As consumers respond to price changes, their preferences and consumption patterns may shift, making it challenging to predict the exact outcomes of a Ramsey tax system.

Revenue Adequacy: Ramsey taxation may not always generate sufficient revenue for the government to fund essential public services and programs. In practice, policymakers often need to consider revenue requirements in addition to efficiency concerns.

Practical Implementation: Implementing a Ramsey tax system in the real world can be complex, requiring constant monitoring and adjustment as consumer preferences and market conditions change. This can be administratively challenging.

Sensitivity to Elasticity Estimates: The outcomes of a Ramsey tax system are sensitive to the accuracy of price elasticity estimates. Small errors in elasticity estimation can lead to suboptimal tax rates and

inefficient resource allocation.

Tax Evasion and Tax Planning: When consumers and businesses anticipate the impact of Ramsey taxation, they may engage in tax planning and evasion to minimize their tax liabilities, potentially undermining the intended efficiency gains.

In practice, governments often balance multiple objectives, including revenue generation, equity, and political considerations when designing tax systems. While Ramsey's rule provides valuable guidance on how to minimize the efficiency costs of taxation, it may need to be supplemented with other tax policy considerations to address the practical complexities and distributional concerns of real-world tax systems.

Table A9: ChatGPT-3.5 answer to Question 9

The concept of separability in the context of optimal taxation rules is essential and can significantly affect the design of an optimal tax system when both income and commodity taxes can be employed simultaneously. Separability refers to the degree to which the utility function is separable between consumption and leisure (or labour supply), meaning that the marginal rate of substitution (MRS) between consumption goods is independent of labour supply. In mathematical terms, if the utility function is fully separable, then the MRS depends only on the utility derived from consumption goods and not on labor supply.

The role of separability in optimal taxation can be understood as follows:

Effect on Taxation of Income and Commodity Goods: When the utility function is fully separable, it implies that the consumer's preferences for consumption and leisure are entirely independent. In this case, the government can design the tax system more efficiently because it can tax labour income and consumption goods separately without distorting the labour-leisure choice. Specifically, the government can apply income taxation to labour earnings (tax on labour income) to generate revenue without affecting the individual's choices about how much to work. At the same time, it can also apply commodity taxes to various goods without influencing the labour-leisure trade-off.

Efficiency and Simplicity: The separability of the utility function simplifies the design of an optimal tax system, making it more efficient and straightforward. This can reduce administrative and compliance costs associated with complex tax structures. Consumers can make consumption choices based on their preferences for goods, while the government can focus on generating revenue through income taxation without disrupting these choices.

Income Redistribution and Equity: While separability simplifies the design of the tax system, it does not address issues of income redistribution and equity. In practice, policymakers often consider equity concerns when designing tax systems. Full separability may not account for these considerations.

Practical Realism: It is important to note that in practice, the utility function may not always be fully separable. Real-world utility functions can exhibit varying degrees of separability, and the extent of separability can depend on individual and household characteristics.

Policymakers must strike a balance between the theoretical ideal of full separability and the practical realities of complex utility functions and social objectives.

In summary, the separability of the utility function plays a significant role in the design of optimal tax systems that combine income and commodity taxes. When the utility function is fully separable, the government can tax labour income and consumption goods separately, which can lead to a more efficient and straightforward tax system. However, in practice, separability is a theoretical concept, and real-world tax systems often involve trade-offs between efficiency and equity considerations.

Table A10: ChatGPT-3.5 answer to Question 10

The concept of "voting with one's feet" is rooted in the idea that individuals can reveal their preferences for local public goods and achieve an efficient and stable equilibrium by choosing the jurisdiction in which to live. However, this concept relies on certain assumptions and conditions to be valid. The following are some of the key assumptions required to make this claim:

Mobility: Individuals must have the ability to move between jurisdictions or regions easily and at a relatively low cost. High mobility ensures that people can choose the jurisdiction that aligns with their preferences for public goods.

Perfect Information: Individuals must have perfect or near-perfect information about the quality and quantity of public goods provided in different jurisdictions. This means they should be aware of the services, infrastructure, and policies of various regions and how these affect their well-being.

Homogeneous Preferences: For the concept of voting with one's feet to work, it is assumed that individuals have relatively homogeneous preferences for public goods. In other words, the public goods provided by different jurisdictions should be suitable for a wide range of residents and not cater exclusively to a particular group.

No Externalities: The absence of externalities is crucial. If individual choices to live in a particular jurisdiction have significant negative spillover effects on others (e.g., congestion or pollution), it can undermine the efficiency of the equilibrium.

Competitive Jurisdictions: Jurisdictions should be in competition with one another to attract residents. If a region has a monopoly on providing public goods, individuals may not have meaningful choices, and the concept of voting with one's feet loses its effectiveness.

Ease of Relocation: The ease and cost of relocating should not be prohibitive. If moving is too costly or complicated, individuals may not be able to make location choices based on their preferences for public goods.

No Tax or Benefit Redistributions: Redistribution of wealth through taxation or benefits across regions can disrupt the efficiency and stability of the equilibrium. Regions should not engage in significant wealth transfers that might influence individuals' choices.

Stable Preferences: Preferences for public goods should be relatively stable over time. If preferences change frequently or are subject to sudden shifts, the effectiveness of voting with one's feet may be limited.

Size of Jurisdictions: Jurisdictions should be of an appropriate size. If they are too small, the provision of public goods might be inefficient due to economies of scale. If they are too large, individual preferences may not be adequately reflected in local public goods provision.

No Barriers to Entry: Individuals should not face discriminatory barriers to entry or discrimination based on factors like race, ethnicity, or socioeconomic status when choosing a jurisdiction to live in.

It's important to recognize that the real-world applicability of the "voting with one's feet" concept can be limited due to deviations from these assumptions. In practice, regions or jurisdictions often have diverse populations with varying preferences, and addressing the needs of all residents can be a complex challenge. Additionally, factors like inequality, political power, and historical legacies can influence the efficiency and equity of public goods provision across regions.

Table A11: ChatGPT-3.5 answer to Exercise 1

Let's go step by step through the questions:

a) Efficient Allocation using Median Voter Rule:

The median voter rule suggests that the level of the public good chosen will be the one preferred by the median voter. To find this, let's first find the median marginal utility. Arrange the marginal utilities in ascending order:

$$MU(A)_G = 20 - G; MU(B)_G = 40 - G; MU(C)_G = 80 - G.$$

Ordering them, we get: $MU(A)_G < MU(B)_G < MU(C)_G$.

So, $MU(B)_G$ is the median. Setting $MU(B)_G$ equal to zero and solving for G: $40 - G = 0 \Rightarrow G = 40$

So, according to the median voter rule, the efficient allocation for the public good is $G = 40$.

b) Lindahl Prices:

Lindahl prices are the prices at which individuals are willing to pay for an additional unit of the public good. The individual's willingness to pay is equal to their marginal utility.

$$p_A = MU(A)_G; p_B = MU(B)_G; p_C = MU(C)_G$$

Substitute in the given expressions:

$$p_A = 20 - G; p_B = 40 - G; p_C = 80 - G = 40$$

At the efficient level $G = 40$, the prices are:

$$p_A = 20 - 40 = -20; p_B = 40 - 40 = 0; p_C = 80 - 40 = 40$$

Since prices cannot be negative, we need to adjust them to be non-negative. We can add a constant to all prices to make them non-negative without changing the allocation. The sum of these prices must equal the

cost of the public good at the efficient level: $-20 + 0 + 40 + \text{constant} = 50(40) + 20$

Solving for the constant: $20 + \text{constant} = 200 + 20 \Rightarrow \text{constant} = 1,980$. So, the Lindahl prices are:

$$p_A = 20 - G + \text{constant} = 1,980;$$

$$p_B = 40 - G + \text{constant} = 1,980$$

$$p_C = 80 - G + \text{constant} = 2,020$$

c) Graphical Representation:

d) Optimal Level of a Private Good:

For a private good, we want to equate the marginal utilities to determine the efficient level. Equate the marginal utilities to zero:

$$MU(A)_X = 20 - X = 0 \Rightarrow X = 20; MU(B)_X = 40 - X = 0 \Rightarrow X = 40;$$

$$MU(C)_X = 80 - X = 0 \Rightarrow X = 80.$$

The efficient level is the one that satisfies all individuals, so $X = 40$ is the optimal level for the private good.

Table A12: ChatGPT-3.5 answer to Exercise 2

To solve this problem, we need to consider the external costs and find the socially optimal quantity and the Pigouvian tax. Let's break down the solution step by step:

i) Quantity produced in the case of no government intervention:

In a perfectly competitive market, the quantity produced is where the marginal cost equals the price. The marginal cost (MC) is the derivative of the total private cost function with respect to quantity (X):

$$MC = \frac{\partial TPC}{\partial x} = 2X + 10$$

Setting MC equal to the price P_x^*

$$2X + 10 = 250$$

Solving for X, $X = 120$. So, the quantity produced with no government intervention is **120**.

ii) Optimal quantity produced with Pigouvian tax (internalizing the externality):

The socially optimal quantity is where the marginal social cost (MSC) equals the price. The MSC includes both private and external costs. The MSC is the sum of the marginal private cost (MPC) and the marginal external cost (MEC):

$$MSC = MPC + MEC$$

$$MSC = 2X + 10 + 2X$$

Setting MSC equal to the price;

$$250 = 2X + 10 + 2X$$

So, the optimal quantity produced with internalizing the externality is $X=60$.

iii) Tax per unit of output:

The Pigouvian tax is the difference between the marginal social cost and the marginal private cost:

$$\text{Tax per unit} = MSC - MPC$$

Substitute the values:

$$\text{Tax per unit} = (2X+10+2X) - (2X+10) = 250 - 130 = 120$$

So, the tax per unit of output is 120.

Comparing profits:

To compare profits, we need to calculate the profits with and without the Pigouvian tax.

No government intervention:

$$\text{Total revenue (TR)} = 250 \times 120 = 30,000$$

$$\text{Total private cost (TPC)} = X^2 + 10X + 50 \text{ at } X = 120$$

$$\text{Profit without intervention} = TR - TPC$$

With Pigouvian tax:

$$\text{Total revenue (TR)} = 250 \times 60 = 15,000$$

$$\text{Total private cost (TPC)} = X^2 + 10X + 50 \text{ at } X = 60$$

$$\text{Tax revenue} = \text{Tax per unit} \times X = 120 \times 60 = 7,200$$

$$\text{Profit with intervention} = TR - TPC - \text{Tax revenue}$$

Compare the two profits to see the impact of government intervention.

Table A13: ChatGPT-3.5 answer to Exercise 3

(a) Expected Outcome without Vote Trading:

Without vote trading, each region will vote in favor of its own project, resulting in a tie. Therefore, neither

project (G1 nor G2) will be approved.

(b) Increasing Welfare through Vote Trading:

Vote trading could potentially increase overall welfare. In this case, regions could negotiate and agree to support each other's projects, leading to both projects being approved. For example, region A might agree to support G2 (highway in region B) in exchange for region B's support for G1 (airport in region A). This way, both projects get approved, and overall welfare increases.

(c) Risk of Exploitation of a Minority (Region C):

Yes, there is a risk of exploitation of a minority (region C) in this situation. Since regions A and B can form a coalition to approve both projects, region C might be left without its preferred project, even if it has significant benefits for the region.

(d) Possible Compensation for Region C:

To avoid the exploitation of region C, compensation or side payments can be considered. Regions A and B could offer some form of compensation to region C in exchange for its support. This compensation could be in the form of additional funding for another project in region C, future benefits, or other concessions.

For example, if regions A and B offer additional funding or support for a project in region C, it could incentivize region C to join the coalition and approve both projects. This way, all regions benefit, and there is a more equitable distribution of projects.

Negotiating fair compensation is crucial to achieving a cooperative outcome and avoiding the exploitation of the minority region.

Table A14: ChatGPT-3.5 answer to Exercise 4

Certainly! Let's go through each calculation:

a. Price and Equilibrium Quantity Before and After Tax:

Before tax: $X_D = 4000 - 4P$; $X_S = -600 + 2P$

Set $X_D = X_S$ to find the equilibrium price (P) before tax.

After tax:

The new demand curve with tax (X_{D_t}) is given by $X_{D_t} = 4000 - 4(P(1+t))$, where

$T = 0.05$ (5% tax rate). The supply curve remains the same.

Set $X_{D_t} = X_S$ to find the new equilibrium price (P') and then substitute it back to find the new equilibrium quantity.

b. Change in Price Experienced by Agents (Tax Incidence):

To find who pays the tax, compare the change in price before and after the tax. The side (demand or supply) with a more inelastic curve tends to bear more of the tax burden.

Change in price = $P' - P$

c. Elasticity of Demand and Supply Analysis:

Calculate the price elasticity of demand (E_d) and supply (E_s) using the equilibrium values before tax. The formula for elasticity is:

$$E = \frac{\% \text{ change in quantity}}{\% \text{ change in price}}$$

d. Excess Burden (Harberger Triangle):

The excess burden (deadweight loss) can be found by comparing the quantity traded before and after the tax. It represents the reduction in total surplus due to the tax.

Appendix B

The explanations justifying the assessment are given in tables B1 to B14.

Table B1: Assessment of question 1

Contextual Understanding	Quality	Completeness
A+	D	F

It is evident that several hypotheses are missing from the current analysis. The following hypotheses have been identified as missing from the literature:

1. Peacock-Wiseman's "displacement effect" hypothesis
2. The Fiscal Illusion hypothesis (Puviani, Goetz, Buchanan, Wagner, Oates and others)
3. The Government Agency hypothesis (Laffont and others)
4. The Rent-Seeking hypothesis (Tullock and others)
5. The Budget-Setting hypothesis (Wildavsky and others)
6. The Unbalanced Growth hypothesis (Baumol, Bowen, Beck and others)

In the context of the Theory of Bureaucratic Expansion, it would be beneficial to include references to Niskanen, Romer and Rosenthal.

In the field of political business cycle theory, references to Frey and Schneider would be greatly appreciated.

In the field of redistribution theory, references to Downs, Romer, Rosenthal, Richard and Schneider would be greatly appreciated.

Table B2: Assessment of question 2

Contextual Understanding	Quality	Completeness
A+	B-	E

- It would be beneficial to include a reference to Buchanan in this context.
- The case of common property resources that offer rivalry in consumption but are not excludable is not addressed.
- A comment on the *congestion effect*, which reflects partial rivalry in consumption above a certain level of demand, is missing.
- It would be beneficial to receive a comment on *variable utilization rule* and *two-part tariff*
- It would be beneficial to distinguish between small and large clubs .

Table B3: Assessment of question 3

Contextual Understanding	Quality	Completeness
A+	A	B+

Table B4: Assessment of question 4

Contextual Understanding	Quality	Completeness
A+	A	C+

- No comment is made on the case of common property *resources* that are rivalrous but not excludable.
- No comment is made on the congestion effect, which reflects partial rivalry in consumption above a certain level of demand.

Table B5: Assessment of question 5

Contextual Understanding	Quality	Completeness
A+	A	A

Table B6. Assessment of question 6

Contextual Understanding	Quality	Completeness
A+	A	B-
<ul style="list-style-type: none"> No commentary is provided on the concepts of <i>moral hazard</i>, <i>adverse selection</i>, <i>pooling equilibria</i>, <i>separating equilibria</i>, or the distinction between <i>hidden knowledge</i> and <i>hidden action</i>. No commentary is provided on the concepts of <i>signalling</i> and <i>screening</i>. It would be beneficial to include references to Akerlof and Spence in the text. 		

Table B7: Assessment of question 7

Contextual Understanding	Quality	Completeness
A+	A	A
No commentary is provided on May's theorem regarding majority rule.		

Table B8: Assessment of question 8

Contextual Understanding	Quality	Completeness
A+	A	A

Table B9: Assessment of question 9

Contextual Understanding	Quality	Completeness
A+	A+	A
<ul style="list-style-type: none"> It would be beneficial to include references to Mirrlees, Diamond, Myles, Atkinson and Stiglitz. 		

Table B10: Assessment of question 10

Contextual Understanding	Quality	Completeness
A+	A+	A

Table B11. Assessment of Exercise 1

Approach to the solution	<ul style="list-style-type: none"> The answer does not consider the fact that only individuals B and C would be willing to contribute to the production of G in accordance with their preferences. No calculation is made of what the median voter demand would be. The answer does not take into account the fact that at marginal cost = 50, only individual C would be willing to demand private good X.
Accuracy of numerical results	<ul style="list-style-type: none"> According to <i>Samuelson condition</i> for the optimal provision of public goods $(80 - G) + (40 - G) = 50$ then $G^* = 35$ (instead of $G^* = 40$). The demand of the median voter is $40 - G = (50/3) \Rightarrow G_B = 23.33$. Therefore $G^* = 35 \neq G_B = 23.33$. Lindahl Prices: $P_A = 0$ (0%); $P_B = 5$ (10%); $P_C = 45$ (90%); Optimal level of provision of X, $X_C^* = 80 - 50 = 30$.
Other comments	<ul style="list-style-type: none"> It would be beneficial to include references to Samuelson and Bowen's postulates.

Table II-12: Assessment of Exercise 2

Approach to the solution	<ul style="list-style-type: none"> The approach is correct.
Accuracy of numerical results	<ul style="list-style-type: none"> The quantity produced without internalising the external cost is correct ($X = 120$), but the optimal quantity calculated for the case where the external effect is taken into account is too high (60 instead of 59.5) The external marginal cost calculated by ChatGPT-3.5 is erroneous. This is because the derivative used is $2X$, rather than $2X+2$. Consequently, the numerical results obtained are also incorrect. Consequently, the result for the Pigouvian tax needed to correct the externality is also erroneous. The correct value is $t = 121$ (not 120). The revenue in the non-intervention case is correctly calculated, but the profit result is not arrived at because the total costs are not calculated beforehand. The market profit that should have been displayed is 14,350.

- In the scenario where a Pigouvian tax is introduced, the total revenue of the firm is incorrect. The correct figure is 14,875, not 15,000. The amount of tax revenue is almost correct (as the effects are compensated). The correct figure is 7,199.5, not 7,200. The profit figure is not provided, as the cost before the introduction of the Pigouvian tax (4,185.25) has not been calculated. In the case of intervention, the profit figure is 3,490.25, but this was not calculated.

Table B13. Assessment of Exercise 3

Approach to the solution	<ul style="list-style-type: none"> • The approach is correct
Accuracy of numerical results	<ul style="list-style-type: none"> • The numbers that are given in the question are not be used. No numerical solutions are provided. • Regional net utility and voting behaviour: Without vote trading $U(A)_{G1} = 1,200 - 300 = 900 \text{ (YES)}$ $U(B)_{G1} = 0 - 300 = -300 \text{ (NO)}$ $U(C)_{G1} = 0 - 300 = -300 \text{ (NO)}$ $U(A)_{G1} + U(B)_{G1} + U(C)_{G1} = 1,200 - 900 = 300 \text{ (NO)}$ $U(A)_{G2} = 0 - 300 = 900 \text{ (NO)}$ $U(B)_{G2} = 1,200 - 300 = -300 \text{ (YES)}$ $U(C)_{G2} = 0 - 300 = -300 \text{ (NO)}$ $U(A)_{G2} + U(B)_{G2} + U(C)_{G2} = 1,200 - 900 = 300 \text{ (NO)}$ • With vote trading without compensation $U(A)_{G1+G2} = 1,200 - 600 = 600 \text{ (YES)}$ $U(B)_{G1+G2} = 1,200 - 600 = 600 \text{ (YES)}$ $U(C)_{G1+G2} = 0 - 600 = -600 \text{ (NO)}$ $U(A)_{G1+G2} + U(B)_{G1+G2} + U(C)_{G1+G2} = 2,400 - 1,800 = 600 \text{ (YES)}$ • With vote trading and compensation (for example, a payment of 400 from each of the regions that benefit from the projects to region C. $U(A)_{G1+G2} = 1,200 - 400 - 600 = 200 \text{ (YES)}$ $U(B)_{G1+G2} = 1,200 - 400 - 600 = 200 \text{ (YES)}$ $U(C)_{G1+G2} = 0 + 800 - 600 = 200 \text{ (YES)}$ $U(A)_{G1+G2} + U(B)_{G1+G2} + U(C)_{G1+G2} = 2,400 - 1,800 = 600 \text{ (YES)}$
Other comments	<ul style="list-style-type: none"> • It would be beneficial to include references to Buchanan and Tullock.

Table B14. Assessment of Exercise 4

Approach to the solution	<ul style="list-style-type: none"> • The approach is correct
Accuracy of numerical results	<ul style="list-style-type: none"> • The numbers that are given in the question are not be used. No numerical solutions are provided. • Numerical solution would be: a) $X_D = X_S \rightarrow 4000 - 4P = -600 + 2P \rightarrow$ $P^* = 766.66; X^* = 933.33$ $X_{Dt} = X_S \rightarrow 4000 - 4(1.05)P = -600 + 2P \rightarrow$

$$P_{St} = 741,94; P_{Dt} = 741,935 (1.05) = 779.03; X_t = 883.88$$

b) Change in $P_S = -24.73$; Change in $P_D = +12.36$

Consumers pay: 33.32%; Producer pay: 66.68%

c) The role of elasticities:

$$E_S = \{[(933.33 - 883.88)/883.88]/[(766.67 - 741.94)/766.67]\}$$

$$E_D = \{[(933.33 - 883.88)/833.88]/[(779.03 - 766.67)/766.67]\}$$

$$E_S = 1.733 (33.32\%); E_D = 3.467 (66.68\%)$$

Alternatively:

$$E = \left(\frac{\partial X}{X}\right) / \left(\frac{\partial P}{P}\right) = \frac{\partial X}{\partial P} \frac{P}{X} \text{ and } \left|\frac{\partial X_D}{\partial P}\right| = 4; \left|\frac{\partial X_S}{\partial P}\right| = 2;$$

$$E_S = 2 (741.94/933.33) / |E_S + E_D| = 33.32\%$$

$$E_D = 4 (770.03/933.33) / |E_S + E_D| = 66.68\%$$

d) Harberger triangle:

Consumer surplus:

$$-(779.03 - 766.67)(933.33 - 883.88) 0.5 = -305.60 (33.32\%)$$

Producer surplus:

$$-(766.67 - 741.93)(933.33 - 883.88) 0.5 = -611.70 (66.68\%)$$

Total excess of burden:

$$(779.03 - 741.93)(933.33 - 883.88) 0.5 = 917.29 (100\%)$$

e) Tax revenue: $(883.88) (779.03 - 741.93) = 32,792$

Notas

¹ See Becker *et al.* (2018).

² See Nguyen *et al.* (2023).

³ See Korinek (2023).

⁴ See Leswing (2023). Gilson *et al.* (2023).

⁵ See Gilson *et al.* (2023).

⁶ See Choi *et al.* (2022) and Terwiesch (2023).

⁷ See West (2023).

⁸ See Geerling *et al.* (2023) or Trent (2023). OpenAI (2023) reports on the evaluation of a number of tests originally designed for humans and solved by GPT-4, including tests in Microeconomics and Macroeconomics.

⁹ See Ungerer and Slade (2022).

¹⁰ See Nguyen *et al.* (2023).

¹¹ See Arnett (2016) and Selwyn (2019).

¹² See Pandey and Gelin (2017).

¹³ See Murphy (2019).

¹⁴ See Aguiar *et al.* (2015), Lakkaraju *et al.* (2015), Luckin and Holmes (2016), Murphy (2019), and The Institute for Ethical AI in Education (2020).

¹⁵ See Du Boulay (1998).

¹⁶ See The Institute for Ethical AI in Education (2020), Klačnjana-Miličević and Ivanović (2021) or Tapalova *et al.* (2022).

¹⁷ See Foltz *et al.* (2013), Murphy (2019), Swauger (2020) and The Institute for Ethical AI in Education (2020).

¹⁸ See The Institute for Ethical AI in Education (2020).

¹⁹ See Goksel and Bozkurt (2019).

²⁰ See Popenici and Kerr (2017).

- ²¹ See Bayne (2015), Knox (2016), and Selwyn (2019).
²² See Holmes (2023).
²³ See Conati (2009).
²⁴ See Holmes (2023).
²⁵ See Braunack-Mayer *et al.* (2020).
²⁶ See Porayska-Pomsta and Holmes (2022).
²⁷ See Marković *et al.* (2019), Braunack-Mayer *et al.* (2020).
²⁸ See Marković *et al.* (2019) and Holmes *et al.* (2021) or Ungerer *et al.* (2022),
²⁹ See Nichols and Holmes (2018), Knox *et al.* (2019b), Vincent-Lancrin and Van der Vlies (2020), or Holstein and Doroudi (2021), Ungerer *et al.* (2022), Hong *et al.* (2022).
³⁰ See Porayska-Pomsta and Holmes (2022) or Nguyen *et al.* (2023).
³¹ See Porayska-Pomsta and Holmes (2022).
³² See Mohammed and Nell-Watson (2019) and Madaio *et al.* (2022).
³³ See Arnett (2016).
³⁴ Other GAI systems are the following: Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM); Transformers, Bayesian Networks, Conditional Language Models.
³⁵ See Aydin and Karaarslan (2023).
³⁶ See Holmes (2023).
³⁷ See Baidoo-Anu and Owusu-Ansah (2023).
³⁸ See McMurtrie (2022).
³⁹ See Farazouli *et al.* (2024) and Dawson (2024).
⁴⁰ See Pickell and Doak (2023).
⁴¹ See Cowen and Tabarrok (2023).
⁴² See Zhai (2022) and Adıgüzel *et al.* (2023).
⁴³ See Iqbal *et al.* (2022).
⁴⁴ Note that ChatGPT-3.5 does not incorporate the latest developments and references as it was trained on data only up until September 2021.
⁴⁵ See Dziri *et al.* (2023).

REFERENCES

- Adıgüzel, T., Kaya, M. H., and Cansu, F. K. (2023) “Revolutionizing education with AI: Exploring the transformative potential of ChatGPT,” *Contemporary Educational Technology*.
- Aguiar, E., Lakkaraju, H, Bhanpuri, N., Miller, D., Yuhas, B., and Addison, K. L. (2015). “Who, When, and Why: A Machine Learning Approach to Prioritizing Students at Risk of Not Graduating High School on Time,” in *Association for Computing Machinery, LAK '15, Proceedings of the Fifth International Conference on Learning Analytics and Knowledge, Poughkeepsie, N.Y.*: 93–102.
- Akgun, S. and Greenhow, C. (2022). “Artificial intelligence in education: Addressing ethical challenges in K-12 settings,”. *AI Ethics* 2: 431–440.
- Aoun, J. E. (2017): *Robot-proof: Higher education in the age of artificial intelligence*. MIT Press.
- Arnett, T. (2016) *Teaching in the Machine Age: How innovation can make bad teachers good and good teachers better*. Christensen Institute.
- Atkinson, A. B. and Stiglitz, J. E. (2015). *Lectures on public economics: Updated edition*. Princeton University Press.

- Auerbach, A. J., Chetty, R., Feldstein, M., and Saez, E. (Eds.). (2013). *Handbook of Public Economics*. Newnes.
- Aydın, Ö. and Karaarslan, E. (2023). "Is ChatGPT leading generative AI? What is beyond expectations?" *Academic Platform Journal of Engineering and Smart Systems*.
- Baidoo-Anu, D. and Owusu-Ansah, L. (2023). "Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning," *Journal of AI*, 7 (1): 52-62.
- Baker, R. S. (2016). "Stupid tutoring systems, intelligent humans," *International Journal of Artificial Intelligence in Education*, 26: 600-614.
- Bearman, M., Ryan, J., and Ajjawi, R. (2023). "Discourses of artificial intelligence in higher education: A critical literature review," *Higher Education*, 86(2): 369-385.
- Becker, S. A., Brown, M., Dahlstrom, E., Davis, A., DePaul, K., Diaz, V., and Pomerantz, J. (2018). *NMC Horizon Report: 2018 Higher Education Edition*. Educause.
- Braunack-Mayer, A. J., Street, J. M., Tooher, R., Feng, X., and Scharling-Gamba, K. (2020): "Student and staff perspectives on the use of big data in the tertiary education sector: A scoping review and reflection on the ethical issues," *Review of Educational Research*, 90 (6). 788-823.
- Choi, J. H., Hickman, K. E., Monahan, A., and Schwarcz, D. (2022) "ChatGPT Goes to Law School," *Journal of Legal Education* 387. Available at SSRN: <https://ssrn.com/abstract=4335905> or <http://dx.doi.org/10.2139/ssrn.4335905>.
- Cowen, T. and Tabarrok, A. T. (2023) "How to Learn and Teach Economics with Large Language Models, Including GPT," *GMU Working Paper in Economics* 23-18.
- Dawson, P., Nicola-Richmond, K. and Partridge, H. (2024) "Beyond open book versus closed book: A taxonomy of restrictions in online examinations," *Assessment & Evaluation in Higher Education*, 49. 2: 262-274.
- Du Boulay, B. (1998). "What does the 'AI' in AIED buy?" *IEE Colloquium on Artificial Intelligence in Educational Software*. IET Digital Library.
- Dziri, N., Lu, X., Sclar, M., Li, X. L., Jiang, L., Lin, B. Y., ... and Choi, Y. (2024) "Faith and fate: Limits of transformers on compositionality," *Advances in Neural Information Processing Systems*, 36.
- Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024) "Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, 49(3), 363-375.
- Foltz, P. W., Streeter, L. A., Lochbaum, K. E., and Landauer, T. K. (2013): "Implementation and applications of the intelligent essay assessor" in M. D. Shermis & J. Burstein (Eds.), *Handbook of automated essay evaluation*: 68–88. Routledge.

- Geerling, W., Mateer, G. D., Wooten, J., and Damodaran, N. (2023) “ChatGPT has Aced the Test of Understanding in College Economics: Now What?” *The American Economist*, 68 (2): 233-245. <https://doi.org/10.1177/05694345231169654>.
- Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., and Chartash, D. (2023) “How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment.” *JMIR Medical Education*, 9(1), e45312. <https://doi.org/10.2196/45312>.
- Goksel N, and Bozkurt, A. (2019). “A. Artificial intelligence in education: Current insights and future perspectives” in *Handbook of Research on Learning in the Age of Transhumanism*: 224–236. IGI Global.
- Gruber, J. (2005). Public finance and public policy. Macmillan.
- Hindriks, J., and Myles, G. D. (2013). Intermediate public economics. MIT press.
- Holmes, W. (2023). “AIED—Coming of Age?” *International Journal of Artificial Intelligence in Education*: 1-11.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., ... and Koedinger, K. R. (2021). “Ethics of AI in education: Towards a community-wide framework,” *International Journal of Artificial Intelligence in Education*: 1-23.
- Holstein, K., and Doroudi, S. (2021). “Equity and Artificial Intelligence in Education: Will “AIED” Amplify or Alleviate Inequities in Education?” arXiv preprint.
- Hong, Y., Nguyen, A., Dang, B., and Nguyen, B. P. T. (2022). “Data Ethics Framework for Artificial Intelligence in Education (AIED),” in *International Conference on Advanced Learning Technologies (ICALT)*: 297-301. IEEE.
- Hyman, D. N. (2005). Public finance: A contemporary application of theory to policy. Thomson/South Western.
- Hrastinski, S., Olofsson, A.D., Arkenback, C. *et al.* (2019). “Critical Imaginaries and Reflections on Artificial Intelligence and Robots in Postdigital K-12 Education,” *Postdigital Science and Education* 1: 427–445.
- Iqbal, N., Ahmed, H., and Azhar, K. A. (2022) “Exploring teachers' attitudes towards using ChatGPT”. *Glob. J. Manag. Adm. Sci.* 3, 97–111. doi: 10.46568/gjmas.v3i4.163.
- Kennedy, M. M. J. (2012). Public finance. PHI Learning Pvt. Ltd.
- Klašnja-Milićević, A. and Ivanović, M. (2021). “E-Learning personalization systems and sustainable education,” *Sustainability*, 13 (12): 6713.
- Knox, J. (2016). “Posthumanism and the MOOC: opening the subject of digital education,” *Studies in Philosophy and Education*, 35: 305-320.

- Knox, J., Wang, Y., and Gallagher, M. (2019a). “Introduction: AI, inclusion, and ‘everyone learning everything’,” *Artificial intelligence and inclusive education: Speculative futures and emerging practices*: 1-13.
- Knox, J., Yu, W., and Gallagher, M. (2019b). *Artificial intelligence and inclusive education*. Springer Singapore.
- Korinek, A. (2023) “Generative AI for economic research: Use cases and implications for economists”, *Journal of Economic Literature*, 61(4): 1281-1317.
- Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., and Addison, K. L. (2015). “A machine learning framework to identify students at risk of adverse academic outcomes” in *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*: 1909-1918.
- Leswing, K. (2023). March 26). OpenAI announces GPT-4, claims it can beat 90% of humans on the SAT. CNBC. <https://www.cnbc.com/2023/03/14/openai-announces-gpt-4-says-beats-90percent-of-humans-on-sat.htm>.
- Lin P.H., Wooders, A., Wang, J.T.Y., and Yuan W.M. (2018). “Artificial intelligence, the missing piece of online education?” *IEEE Engineering Management Review*, 46 (3): 25–38.
- Luckin, R. and Holmes, W. (2016): *Intelligence unleashed: An argument for AI in education*. UCL Knowledge Lab (University College London) and Pearson.
- McMutrie, B. (2023) “AI and the future of undergraduate writing,” *The Chronicle of Higher Education*. February 8). <https://www.chronicle.com/article/ai-and-the-future-of-undergraduate-writing>.
- Madaio, M., Blodgett, S. L., Mayfield, E., and Dixon-Román, E. (2022). “Beyond ‘fairness’: Structural (in) justice lenses on ai for education” in *The ethics of artificial intelligence in education*: 203-239. Routledge.
- Marković, M. G., Debeljak, S., and Kadoić, N. (2019). “Preparing students for the era of the General Data Protection Regulation (GDPR),” *TEM Journal: Technology, Education, Management, Informatics*, 8: 150–156.
- Mohammed, P. S. and Nell’Watson, E. (2019). “Towards inclusive education in the age of artificial intelligence: Perspectives, challenges, and opportunities” in *Artificial Intelligence and Inclusive Education: Speculative futures and emerging practices*: 17-37.
- Mollick, E. R. and Mollick, L. (2023) “Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts,” <https://dx.doi.org/10.2139/ssrn.4391243>.
- Murphy, R. F. (2019). “Artificial intelligence applications to support k–12 teachers and teaching: a review of promising applications, challenges, and risks,” *Perspective*: 1–20.
- Myles, G. D. (2005). *Public Economics*. Cambridge and New York: Cambridge University Press. Persson, Mats, and Agnar Sandmo.

- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., and Nguyen, B. P. T. (2023). “Ethical principles for artificial intelligence in education” in *Education and Information Technologies*, 28 (4): 4221-4241.
- Nichols, M. and Holmes, W. (2018). “Don't do Evil: Implementing Artificial Intelligence in Universities” in *EDEN Conference Proceedings*, 2: 110-118.
- OpenAI (2023) “GPT-4 Technical Report,” <https://arxiv.org/pdf/2303.08774>.
- Pickell, T. R. and Doak, B. R. (2023). Five Ideas for How Professors Can Deal with GPT-3... For Now.
- Popenici, S. A., and Kerr, S. (2017). “Exploring the impact of artificial intelligence on teaching and learning in higher education,” *Research and Practice in Technology Enhanced Learning*, 12(1): 1-13.
- Porayska-Pomsta, K., and Holmes, W. (2022). “Toward ethical AIED,” arXiv preprint.
- Rosen, H. S., & Gayer, T. (1995). Public Finance, Richard D. Irwin Inc.
- Selwyn, N. (2019): *Should robots replace teachers? AI and the future of education*. John Wiley & Sons.
- Sharma, S. and Yadav, R. (2022) “Chat GPT—A technological remedy or challenge for education system”, *Global Journal of Enterprise Information System*, 14(4): 46-51.
- Stiglitz, J. E., & Rosengard, J. K. (2015). Economics of the public sector: Fourth international student edition. WW Norton & Company.
- Swauger, S. (2020). “Our bodies encoded: Algorithmic test proctoring in higher education” in J. Stommel, C. Friend, & S. M. Morris (Eds.), *Critical Digital Pedagogy*. Pressbooks.
- Tapalova, O., Zhiyenbayeva, N., and Gura, D. (2022). “Artificial Intelligence in Education: AIED for Personalised Learning Pathways,” *The Electronic Journal of e-Learning*, 20 (5):639-653.
- Terwiesch, C. (2023) “Would Chat GPT3 Get a Wharton MBA?”, *Pennsylvania: Mack Institute for Innovation Management at the Wharton School, University of Pennsylvania*.
- The Institute for Ethical AI in Education (2020). *Interim report: towards a shared vision of ethical AI in education*. The University of Buckingham.
- Trent, C. (2023) “ChatGPT and Current Events in the Economics Classroom,” *Business Education Innovation Journal*, 15(1).
- Ungerer, L., and Slade, S. (2022). “Ethical considerations of artificial intelligence in learning analytics in distance education contexts” in *Learning Analytics in Open and Distributed Learning: Potential and Challenges*: 105-120. Singapore: Springer Nature Singapore.
- Vincent-Lancrin, S., and Van der Vlies, R. (2020). *Trustworthy artificial intelligence (AI) in education: Promises and challenges*. OECD Library.

West, C. G. (2023) “AI and the FCI: Can ChatGPT project an understanding of introductory Physics?,” arXiv preprint arXiv:2303.01067.

Zhai, X. (2022) “ChatGPT user experience: Implications for education”. Available at SSRN 4312418.

¿Qué nota obtiene ChatGPT en un examen de Economía Pública 01?

Resumen

Hoy en día es casi imposible discutir el futuro de la industria educativa sin considerar la influencia de la inteligencia artificial (IA). Este artículo examina las oportunidades y desafíos asociados con el uso de la IA en la educación, con especial atención en la capacidad de ChatGPT-3.5 para resolver un examen de Economía Pública que incluye preguntas teóricas y ejercicios prácticos, al nivel de un curso de Grado. El algoritmo ha demostrado ser capaz de alcanzar una notable con los criterios establecidos de puntuación. La conclusión del estudio es que este chatbot basado en IA demuestra capacidades suficientes para ayudar a los educadores a preparar materiales para lecciones, clases y seminarios, así como servir de apoyo en tutoría con estudiantes en esta materia.

Palabras clave: Inteligencia artificial, ChatGPT, Economía Pública.

Códigos JEL: A22; I20; H89