Can GPT-4 Replace Human Examiners? A Competition on Checking Open-Text Answers

Abdullah Al Zubaer, Michael Granitzer, Stephan Geschwind, Johann Graf Lambsdorff, and Deborah Voss

S1 GPT Answers

For each question, we generate five variants of GPT answers. These variants differ only in terms of the information that GPT receives to base its answer on:

- Variant 1: No information.

- Variant 2: Relevant textbook excerpt.

- Variant 3: Irrelevant textbook excerpt.

- Variant 4: Relevant and irrelevant textbook excerpt.

- Variant 5: 10 randomly selected student answers.

Prompt (translated to English):

You are a student of macroeconomics and answer an open question by entering a text. Pay attention to precise wording and logical presentation.

Open question:
[OPEN QUESTION]

For variants 2, 3, and 4:

Please base your answer on the following textbook excerpt: [TEXTBOOK EXCERPT]

For variant 5:
Please base your answer on the following ten student answers: [STUDENT ANSWERS]

Answer:

Run Prompt

Shorten your answer to a maximum of [number based on average length of student answers] words:

Run Prompt

Add a typo to your answer:

Run Prompt

Prompt (Original Version in German):

Du bist eine Studentin der Makroökonomik und beantwortest eine offene Fragestellung durch Eingabe eines Textes. Achte auf präzise Formulierung und logische Darstellung.

Offene Fragestellung:

[OPEN QUESTION]

For variants 2, 3, and 4:

Bitte basiere deine Antwort auf folgendem Lehrbuchauszug:

[TEXTBOOK EXCERPT]

For variant 5:

Bitte basiere deine Antwort auf folgenden 10 Studierendenantworten:

[STUDENT ANSWERS]

Antwort:

Run Prompt

Kürze deine Antwort auf maximal [Zahl basierend auf der Länge der Studierendenantworten] Wörter:

Run Prompt

Füge deiner Antwort einen Tippfehler bei:

Run Prompt

S2 Instructions for Human Examiners

Instructions for Ranking Task (translated to English)

Please read the following instructions carefully before starting the task.

The goal of the task is to rank student responses to an open-ended question in macroeconomics based on a sample solution. It is important that you work on the task alone and independently of the others. You can of course talk during the task as usual, but please do not exchange ideas about the task and its contents. The others have received a similar task from us and are also asked to work on it alone and independently.

In the attachment of this e-mail you will find an Excel file. Once you open this file, you will see an Excel document with 10 spreadsheets: "Ranking1" to "Ranking10". At the top of each spreadsheet you will see an open question and the corresponding sample solution. These are the same on all spreadsheets. Below that, there are five randomly selected student answers. Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be assigned once. Please consider the following criteria when assigning ranks: Correctness, completeness, and relevance to the question. You can base your ranking on the closeness of the student answer to the sample solution. Please enter the rank as a number next to the respective student answer in the column "Rank". If you consider two answers to be equivalent, you must still give preference to one. Use style and spelling to help you decide.

Once you have assigned all five ranks on a spreadsheet, move to the next spreadsheet. Once you have edited all ten spreadsheets, save the document as "Rangfolge1_erledigt" and email it to us. Please include an approximate time estimate in the email. This is for research purposes only.

Instructions for Ranking Task (Original in German)

Bitte lies die folgenden Instruktionen sorgfältig durch, bevor Du mit der Aufgabe beginnst.

Ziel der Aufgabe ist es, Rangfolgen von Studierendenantworten auf eine offene Frage im Fach Makroökonomie basierend auf einer Musterlösung zu bilden. Es ist wichtig, dass Du die Aufgabe alleine und unabhängig von den anderen bearbeitest. Ihr könnt Euch während der Aufgabe natürlich wie immer unterhalten, aber tauscht Euch bitte nicht über die Aufgabe und ihre Inhalte aus. Die anderen haben eine ähnliche Aufgabe von uns erhalten und sind auch gebeten, diese alleine und selbstständig zu bearbeiten.

Im Anhang dieser E-Mail befindet sich eine Excel-Datei. Sobald Du diese Datei öffnest, siehst Du ein Excel-Dokument mit 10 Tabellenblättern: „Ranking1" bis „Ranking10". Oben auf jedem Tabellenblatt siehst Du eine offene Frage und die dazugehörige Musterlösung. Diese sind auf allen Tabellenblättern gleich. Darunter befinden sich fünf zufällig ausgewählte Studierendenantworten. Deine Aufgabe ist es, für jede Studierendenantwort einen Rang zwischen 1 (beste Antwort) und 5 (schlechteste Antwort) zu vergeben. Jeder Rang kann nur ein Mal vergeben werden. Bitte beachte bei der Vergabe der Ränge folgende Kriterien: Richtigkeit, Vollständigkeit und Relevanz für die Fragestellung. Dabei kannst Du deine Rangvergabe auf der jeweiligen Nähe der Studierendenantwort zur Musterlösung basieren. Bitte trage den Rang neben der jeweiligen Studierendenantwort in der Spalte „Rang" als Zahl ein. Falls Du zwei Antworten für gleichwertig hältst, musst Du dennoch eine bevorzugen. Ziehe Stil und Rechtschreibung als zusätzliche Entscheidungshilfen hinzu. Sobald Du alle fünf Ränge auf einem Tabellenblatt vergeben hast, wechsle auf das nächste Tabellenblatt.

Sobald Du alle zehn Tabellenblätter bearbeitet hast, speichere das Dokument unter „Rangfolge1_erledigt" und sende es uns per E-Mail. Bitte mache in der Email eine ungefähre Angabe über den zeitlichen Aufwand. Dies dient lediglich zu Forschungszwecken.

Instructions for Point Assessment (translated to English)

Please read the following instructions carefully before starting the task.

The goal of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

In the attachment of this e-mail you will find an Excel file. Once you open this file, you will see an Excel document with 10 spreadsheets: "Ranking1" to "Ranking10". At the top of each spreadsheet you will see an open question and the corresponding sample solution. These are the same on all spreadsheets. Below that, there are five randomly selected student answers. Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution. You can base your award of points on the proximity of the student's answer to the content of the sample solution. These are the relevant contents that are necessary for a correct and complete answer to the question. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. Enter only the total number of points for each student answer.

Once you have assigned all five ranks on a spreadsheet, move to the next spreadsheet. Once you have edited all ten spreadsheets, save the document as "Bepunktung1_erledigt" and email it to us. Please include an approximate time estimate in the email. This is for research purposes only.

Instructions for Point Assessment Task (Original in German)

Bitte lies die folgenden Instruktionen sorgfältig durch, bevor Du mit der Aufgabe beginnst.

Ziel dieser Aufgabe ist es, für Studierendenantworten auf eine offene Frage im Fach Makroökonomik basierend auf einer Musterlösung Punkte zu vergeben. In der Musterlösung sind diejenigen Inhalte wie folgt gekennzeichnet, für die es einen Punkt gibt: (1), (1 Punkt) oder (1P). Die Bepunktungen (1), (1 Punkt) oder (1P) beziehen sich jeweils auf den vorausgegangenen Inhalt.

Im Anhang dieser E-Mail befindet sich eine Excel-Datei. Sobald Du diese Datei öffnest, siehst Du ein Excel-Dokument mit 10 Tabellenblättern: „Ranking1" bis „Ranking10". Oben auf jedem Tabellenblatt siehst Du eine offene Frage und die dazugehörige Musterlösung. Diese sind auf allen Tabellenblättern gleich. Darunter befinden sich fünf zufällig ausgewählte Studierendenantworten. Deine Aufgabe ist es, für jede Studierendenantwort eine Punktzahl zu vergeben. Auch 0 Punkte sind möglich. Die Musterlösung enthält alle für die Fragestellung relevante und richtige Inhalte. Sie ist vollständig. Es können nicht mehr Punkte vergeben werden als in der Musterlösung. Dabei kannst Du deine Vergabe der Punkte auf der jeweiligen Nähe der Studierendenantwort zu den bepunkteten Inhalten der Musterlösung basieren. Dies sind diejenigen relevanten Inhalte, die für die inhaltlich korrekte und vollständige Beantwortung der Frage notwendig sind. Eine Punktzahl in Höhe der von der Musterlösung vergebenen Punktzahl bedeutet, dass die Studierendenantwort alle relevanten Inhalte der Musterlösung enthält. Es können nur ganze Punkte vergeben werden. Addiere die jeweiligen Punkte, die du für eine Studierendenantwort vergeben hast. Gib für jede Studierendenantwort nur die jeweilige Gesamtpunktzahl an.

Sobald Du alle zehn Tabellenblätter bearbeitet hast, speichere das Dokument unter „Bepunktung1_erledigt" und sende es uns per E-Mail. Bitte mache in der Email eine ungefähre Angabe über den zeitlichen Aufwand. Dies dient lediglich zu Forschungszwecken.

Note: The number in "Rangfolge1_erledigt" and "Bepunktung1_erledigt" indicates the question and is counted up to 6 with respect to the question at hand.

S3 Original Prompt

In this section, we present our original prompt, which is originally in German, alongside its English translation.

English Translation

Supplementary Table 1. Zero shot prompt template utilized in our experiments for investigating H1 w.r.t ranking.

You are a student assistant for macroeconomics and are given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 2. Zero shot prompt template utilized in our experiments for investigating H1 w.r.t point assessment.

You are a student assistant for macroeconomics and are given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

German (original)

Du bist eine studentische Hilfskraft für das Fach Makroökonomik und bekommst eine Aufgabe gestellt. Ziel dieser Aufgabe ist es, eine Rangfolge von Studierendenantworten auf eine offene Frage im Fach Makroökonomik basierend auf einer Musterlösung zu bilden.

Dies ist die offene Frage:

[INSERT_FRAGE]

Dies ist die Musterlösung:

[INSERT_SAMPLE_SOLUTION]

Dies sind fünf zufällig ausgewählte Studierendenantworten in zufälliger Reihenfolge:

Studierendenantwort_1: [STUDENT_ANSWER_1]

Studierendenantwort_2: [STUDENT_ANSWER_2]

Studierendenantwort_3: [STUDENT_ANSWER_3]

Studierendenantwort_4: [STUDENT_ANSWER_4]

Studierendenantwort_5: [STUDENT_ANSWER_5]

Deine Aufgabe ist es, für jede Studierendenantwort einen Rang zwischen 1 (beste Antwort) und 5 (schlecht-este Antwort) zu vergeben. Jeder Rang kann nur ein Mal vergeben werden. Bitte beachte bei der Vergabe der Rängefolgende Kriterien: Richtigkeit, Vollständigkeit und Relevanz für die Fragestellung. Dabei kannst Du deine Rangvergabeauf der jeweiligen Nähe der Studierendenantwort zur Musterlösung basieren. Falls Du zwei Antworten für gleichwer-tig hältst, musst Du dennoch eine bevorzugen. Ziehe Stil und Rechtschreibung als zusätzliche Entscheidungshilfen hinzu.

Stelle deine Rangvergabe wie folgt dar (nenne nur die Antwort und den dazugehörigen Rang):

Studierendenantwort_X: Rang Y

Output:\n

Supplementary Table 4. Zero shot prompt template utilized in our experiments for investigating H1 w.r.t point assessment.

Du bist eine studentische Hilfskraft für das Fach Makroökonomik und bekommst eine Aufgabe gestellt. Ziel dieser Aufgabe ist es, für Studierendenantworten auf eine offene Frage im Fach Makroökonomik basierend auf einer Musterlösung Punkte zu vergeben. In der Musterlösung sind diejenigen Inhalte wie folgt gekennzeichnet, für die es einen Punkt gibt: (1), (1 Punkt) oder (1P). Die Bepunktungen (1), (1 Punkt) oder (1P) beziehen sich jeweils auf den vorausgegangenen Inhalt.

Dies ist die offene Frage:

[INSERT_FRAGE]

Dies ist die Musterlösung:

[INSERT_SAMPLE_SOLUTION]

Dies sind fünf zufällig ausgewählte Studierendenantworten in zufälliger Reihenfolge:

Studierendenantwort_1: [STUDENT_ANSWER_1]

Studierendenantwort_2: [STUDENT_ANSWER_2]

Studierendenantwort_3: [STUDENT_ANSWER_3]

Studierendenantwort_4: [STUDENT_ANSWER_4]

Studierendenantwort_5: [STUDENT_ANSWER_5]

Deine Aufgabe ist es, für jede Studierendenantwort eine Punktzahl zu vergeben. Auch 0 Punkte sind möglich. Die Musterlösung enthält alle für die Fragestellung relevante und richtige Inhalte. Sie ist vollständig. Es können nicht mehr Punkte vergeben werden als in der Musterlösung.

Dabei kannst Du deine Vergabe der Punkte auf der jeweiligen Nähe der Studierendenantwort zu den bepunk-teten Inhalten der Musterlösung basieren. Dies sind diejenigen relevanten Inhalte, die für die inhaltlich korrekte undvollständige Beantwortung der Frage notwendig sind. Eine Punktzahl in Höhe der von der Musterlösung vergebenenPunktzahl bedeutet, dass die Studierendenantwort alle relevanten Inhalte der Musterlösung enthält. Es können nurganze Punkte vergeben werden. Addiere die jeweiligen Punkte, die du für eine Studierendenantwort vergeben hast. Gibfür jede Studierendenantwort nur die jeweilige Gesamtpunktzahl an.

Stelle deine Rangvergabe wie folgt dar (nenne nur die Antwort und die vergebene Punktzahl):

Studierendenantwort_X: Punktzahl

Output:\n

S4 Extended results for rank and point assessment addressing hypothesis 1

This section presents the extended result for rank assessment in Supplementary Table 5 and point assessment in Table 6.

Supplementary Table 5. Summary statistics for Inter-rater reliability using Kendall's W for all questions in rank assessment

Questions	Human team	AI-human team 1	AI-human team 2	AI-human team 3
Question 1	$0.791^{\pm0.107}$	$0.838^{\pm0.070}$	$0.858^{\pm0.073}$	$0.824^{\pm0.078}$
Question 2	$0.899^{\pm0.051}$	$0.859^{\pm0.085}$	$0.858^{\pm0.093}$	$0.879^{\pm0.096}$
Question 3	$0.793^{\pm0.176}$	$0.716^{\pm0.168}$	$0.771^{\pm0.135}$	$0.707^{\pm0.153}$
Question 4	$0.824^{\pm0.104}$	$0.856^{\pm0.067}$	$0.869^{\pm0.053}$	$0.811^{\pm0.104}$
Question 5	$0.776^{\pm0.149}$	$0.780^{\pm0.123}$	$0.804^{\pm0.158}$	$0.787^{\pm0.108}$
Question 6	$0.671^{\pm0.191}$	$0.684^{\pm0.199}$	$0.676^{\pm0.150}$	$0.716^{\pm0.181}$
Mean	$0.792^{\pm0.149}$	$0.789^{\pm0.141}$	$0.806^{\pm0.131}$	$0.787^{\pm0.134}$

Standard deviation provided as superscript

Supplementary Table 6. Summary statistics for Inter-rater reliability using Cronbach's Alpha for all tasks in point assessment

Questions	Human team	AI-human team 1	AI-human team 2	AI-human team 3
Question 1	0.935 [0.896, 0.961]	0.914 [0.862, 0.948]	0.896 [0.833, 0.937]	0.908 [0.852, 0.945]
Question 2	0.938 [0.901, 0.963]	0.943 [0.909, 0.966]	0.943 [0.908, 0.966]	0.936 [0.898, 0.962]
Question 3	0.925 [0.880, 0.955]	0.690 [0.504, 0.814]	0.724 [0.558, 0.834]	0.735 [0.576, 0.841]
Question 4	0.918 [0.869, 0.951]	0.924 [0.878, 0.954]	0.936 [0.897, 0.961]	0.920 [0.873, 0.952]
Question 5	0.904 [0.847, 0.943]	0.874 [0.798, 0.924]	0.927 [0.883, 0.956]	0.873 [0.797, 0.924]
Question 6	0.903 [0.844, 0.941]	0.895 [0.832, 0.937]	0.844 [0.751, 0.906]	0.869 [0.791, 0.921]
Question 1 to 6	0.944 [0.932, 0.954]	0.933 [0.918, 0.945]	0.932 [0.917, 0.944]	0.933 [0.919, 0.945]

Confidence Interval at 95% provided in square brackets

S5 Robustness checks and extensions

This section offers the English translation of the prompt in the Robustness Checks and extensions section regarding rank and point assessment. As $P_{v6.1}$ and $P_{v6.2}$, for rank and point assessment, deal with spelling errors in the original German prompt, we cannot provide a translation since the spelling errors cannot be reproduced through translation. We provide the German prompt as it is for rank assessment in Supplementary Table 11 and 12 and for point assessment in Supplementary Table 23 and 24.

S5.1 Prompts for rank assessment (English translation)

Supplementary Table 7. Prompt P_{v2} - Rank assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 8. Prompt P_{v3} - Rank assessment

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

You are a student assistant for macroeconomics and are given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 9. Prompt P_{y4} - Rank assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

This is the open question:

[INSERT_QUESTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 10. Prompt P_{v5} - Rank assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the open question:

[INSERT_QUESTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 11. Prompt $P_{v6.1}$ - Rank assessment

Du bhist eiene studäntische Hilfskraft führ dass Fach Makroökonomik unt bekommst einne Aufgabbe gestellet. Zihl dieser Aufgabee ist es, einne Rangvolge von Studirendenantworten auf einne offene Ffrage im Fach Makroökonomek basirend auf einer Musterlözung zu bilden.

Diss ist dei offenne Ffrage:

[INSERT_FRAGE]

Diss ist die Musterlössung:

[INSERT_SAMPLE_SOLUTION]

Diss sind fünv zufälig ausgevvälte Studierendenantworten in zufäliger Reihenfolgee:

Studierendenantwort_1: [STUDENT_ANSWER_1]

Studierendenantwort_2: [STUDENT_ANSWER_2]

Studierendenantwort_3: [STUDENT_ANSWER_3]

Studierendenantwort_4: [STUDENT_ANSWER_4]

Studierendenantwort_5: [STUDENT_ANSWER_5]

Deinne Aufgabe ist es, für jede Studierendennantwort einen Rrang zwisschen 1 (bäste Antwort) und 5 (schlechteste Antwort) zu vergeben. Jeder Rang kann nur ein Mall vergeben werdden. Bittte beachte bei der Vergabe der Ränge folgende Kriterrien: Richtigkheit, Vollständigkheit und Relevanz für die Fragestelung. Dabei kkanst Du deine Rangvergabe auf der jeweiligen Nähe der Studierendennantwort zur Musterlösung basieren. Falls Du zwei Antwotten für gleichwertig hältst, must Du dennoch eine bevorzugen. Ziee Stil und Rechtschreibung als zusätzliche Entscheidungshilfen hinzuh.

Stelle deine Rangvergabe wie folgt dar (nenne nur die Antwort und den dazugehörigen Rang):

Studierendenantwort_X: Rang Y

Output:\n

Supplementary Table 12. Prompt $P_{v6.2}$ - Rank assessment

Du bist eine studdentische Hilfskraft für das Fach Makroökonomiek und bekommst eine Aufgabe gestelt. Ziel dieser Aufgaabe ist es, eine Rangfolge von Studierendenantworten auf eine offene Frge im Fach Makroökonmik basierend auf einer Musterlösung zu bilden.

Dies ist die offene Fage:

[INSERT_FRAGE]

Dies ist die Musterlösug:

[INSERT_SAMPLE_SOLUTION]

Dies sind fünf zufällig ausgewählte Studierendenantworten in zufäliger Reihenfolge:

Studierendenantwort_1: [STUDENT_ANSWER_1]

Studierendenantwort_2: [STUDENT_ANSWER_2]

Studierendenantwort_3: [STUDENT_ANSWER_3]

Studierendenantwort_4: [STUDENT_ANSWER_4]

Studierendenantwort_5: [STUDENT_ANSWER_5]

Deine Aufgabe ist es, für jede Studierendenatwort einen Rang zwischen 1 (beste Antwort) und 5 (schlechteste Antowrt) zu vergeben. Jeder Rang kann nur ein Mal vergeben werden. Bitte beachte bei der Vergabe der Ränge folgende Kriterien: Richtigkeit, Vollständigkiet und Relevanz für die Fragestellung. Dabei kannst Du deine Rangvergabe auf der jeweiligen Nähe der Studierendenatwort zur Musterlösung basieren. Falls Du zwei Antworten für gleichwertig hältst, musst Du dennoch eine bevorzugen. Ziehe Stil und Rechtschreibung als zusätzliche Entscheidungshilfen hinzu.

Stelle deine Rangvergabe wie folgt dar (nenne nur die Antwort und den dazugehörigen Rang):

Studierendenantwort_X: Rang Y

Output:\n

Supplementary Table 13. Prompt P_{v7} - Rank assessment

Suppose you find yourself in the role of a student assistant for the subject of Macroeconomics and you are assigned an interesting, albeit somewhat challenging, task. Your goal? It involves creating a sort of ranking of student responses to an open-ended question in Macroeconomics, based on a given model answer.

Imagine the open question would be:

[INSERT_QUESTION]

And the model answer to it reads:

[INSERT_SAMPLE_SOLUTION]

Now, here are five randomly selected responses from students, in no specific order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task, or rather, your challenge, is to assign a rank to each of these responses between 1 (as the best answer) and 5 (as the least convincing answer). It is important that each rank remains unique. In assigning the ranks, you might consider the following criteria: the correctness of the answer, its completeness, and how relevant it is to the posed question. It could be helpful to base your decisions on how closely each answer matches the model solution. Should you find yourself considering two answers as equivalent, it would still be necessary to show a slight preference. Perhaps style and spelling could serve as additional factors in making a decision.

The way you present your ranking might look something like this (by only naming the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 14. Prompt P_{v8} - Rank assessment

You are a professor of Macroeconomics and have been given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 15. Prompt P_{v9} - Rank assessment

You are an experienced student assistant for the subject of Macroeconomics and have been given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 16. Prompt P_{v10} - Rank assessment

You are a beginner-level student assistant with limited knowledge in Macroeconomics and have been given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 17. Prompt P_{v11} - Rank assessment

You are an assistant for the subject of Macroeconomics and have been given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Present your ranking as follows (state only the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 18. Prompt P_{v12} - Rank assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to rank student answers to an open question in macroeconomics based on a sample solution.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to assign a rank between 1 (best answer) and 5 (worst answer) for each student answer. Each rank can only be awarded once. Please consider the following criteria when assigning ranks: Correctness, completeness and relevance to the question. You can base your ranking on the proximity of the student's answer to the sample solution. If you consider two answers to be of equal value, you must still give preference to one. Use style and spelling as additional decision-making criteria.

Think step by step. You must explain your ranking, and then present your ranking allocation as follows (only name the answer and the corresponding rank):

Student answer_X: Rank Y

Output:\n

Supplementary Table 19. Prompt P_{v2} - Point assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Supplementary Table 20. Prompt P_{v3} - Point assessment

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

You are a student assistant for macroeconomics and are given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Supplementary Table 21. Prompt P_{v4} - Point assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

This is the open question:

[INSERT_QUESTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Supplementary Table 22. Prompt P_{v5} - Point assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the open question:

[INSERT_QUESTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Du bhist eihne studäntishe Hilfskraft führ das Faach Makroökönomik unt bekommst aine Aufgahbe gestellt. Zihl deiser Aufgahbe is es, führ Studirendenantworten auf eine offähne Frage im Faach Makroökönomik basierent auf ainer Musterlösunk Punkte zu vergeben. In der Musterlösunk sünd diejenigen Inhalte wie folkt gekänzeichnet, führ die es ainen Punkt gibt: (1), (1 Punkd) odher (1P). Die Bepunktungen (1), (1 Punkd) oder (1P) beziähen sisch jewäils auf den vorausgegahnen Inhalt.

Dihes isst die offähne Frage:

[INSERT_FRAGE]

Dihes isst die Musterlösunk:

[INSERT_SAMPLE_SOLUTION]

Dihes sünd fünf zufällik ausgewählte Studirendenantworten in zufälliker Reihänfolge:

Studierendenantwort_1: [STUDENT_ANSWER_1]

Studierendenantwort_2: [STUDENT_ANSWER_2]

Studierendenantwort_3: [STUDENT_ANSWER_3]

Studierendenantwort_4: [STUDENT_ANSWER_4]

Studierendenantwort_5: [STUDENT_ANSWER_5]

Deine Aufgahbe isst es, für jede Studirendenantwort aine Punktzahl zu vergeben. Auch 0 Punkte sünd möglich. Die Musterlösunk enthält alle führ die Fragestellung relevante unt richtige Inhalte. Sie ist vollstängig. Es können nihct mehr Punkte vergeben wärden als in der Musterlösunk.

Dahbei kanst Du deine Vergabe der Punkte auf der jeweiligen Nähe der Studirendenantwort zu den bepunkteten Inhalten der Musterlösunk basieren. Dies sünd diejenigen relevanten Inhalte, die führ die inhaltlich korrekte unt vollständige Beantwortung der Frage notwendig sünd. Aine Punktzahl in Höhe der von der Musterlösunk vergebenen Punktzahl bedeutet, dass die Studirendenantwort alle relevanten Inhalte der Musterlösunk enthält. Es können nur ganze Punkte vergeben wärden. Addiere die jeweiligen Punkte, die du für aine Studirendenantwort vergeben hast. Gib für jede Studirendenantwort nur die jeweilige Gesamtpunktzahl an.

Stelle deine Rangvergabe wie folgt dar (nenne nur die Antwort unt die vergebene Punktzahl):

Studierendenantwort_X: Punktzahl

Output:\n

Supplementary Table 24. Prompt $P_{v6.2}$ - Point assessment

Du birst eine studentische Hilfskraft für das Fach Makroökonmie und bekommst eine Aufgabe gestelt. Ziel deiser Aufgabe ist es, für Studierendenantworten auf eine offene Frge im Fach Makroökonomik basirend auf einer Musterlösung Punkte zu vergebn. In der Musterlösug sind diejenigen Inhalte wie folgt gekenzeichnet, für die es einen Punk gibt: (1), (1 Punkt) oder (1P). Die Bepunktungen (1), (1 Punkt) oder (1P) beziehen sich jewils auf den vorausgehenden Inhalt.

Dies ist die offene Frge:

[INSERT_FRAGE]

Dies ist die Musterlösung:

[INSERT_SAMPLE_SOLUTION]

Dies sind fünf zufällig ausgewählte Studierendenantworten in zufälliger Reihenfolge:

Studierendenantwort_1: [STUDENT_ANSWER_1]

Studierendenantwort_2: [STUDENT_ANSWER_2]

Studierendenantwort_3: [STUDENT_ANSWER_3]

Studierendenantwort_4: [STUDENT_ANSWER_4]

Studierendenantwort_5: [STUDENT_ANSWER_5]

Deine Aufgabe ist es, für jede Studierendenantwort eine Punktzahl zu vergeben. Auch 0 Punkte sind möglich. Die Musterlösung enthält alle für die Fragestellung relevante und richtige Inhalte. Sie ist vollständig. Es können nicht mehr Punkte vergebn werden als in der Musterlösung.

Dabei kannst Du deine Vergabe der Punkte auf der jeweiligen Nähe der Studierendenantwort zu den bepunkteten Inhalten der Musterlösung basieren. Dies sind diejenigen relevanten Inhalte, die für die inhaltlich korrekte und vollständige Beantwortung der Frage notwendig sind. Eine Punktzahl in Höhe der von der Musterlösung vergebenen Punktzahl bedeutet, dass die Studierendenantwort alle relevanten Inhalte der Musterlösung enthält. Es können nur ganze Punkte vergeben werden. Addiere die jeweiligen Punkte, die du für eine Studierendenantwort vergebn hast. Gib für jede Studierendenantwort nur die jeweilige Gesamtpunktzahl an.

Stelle deine Rangvergabe wie folgt dar (nenne nur die Antwort und die vergebene Punktzahl):

Studierendenantwort_X: Punktzahl

Output:\n

Suppose you find yourself in the role of a student assistant for the subject of Macroeconomics and you are assigned a task. The goal of this task might be to assign points to student responses to an open question in Macroeconomics, based on a type of model solution. In this model solution, certain content might be marked in a specific way, for example with (1), (1 point), or (1P), which could indicate that one point can be assigned for these specific contents. These markings – (1), (1 point), or (1P) – might refer to the content immediately preceding them.

Let's imagine this is the open question:

[INSERT_QUESTION]

And this could be the model solution:

[INSERT_SAMPLE_SOLUTION]

Let's further assume there are five randomly selected student responses in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task could be to assign a score to each of these student responses, where even 0 points could be a possibility. The model solution could contain all the content relevant and correct for the question and be considered complete. It might be conceivable that no more points can be awarded than are provided for in the model solution.

You could base your allocation of points on how closely each student response matches the scored content of the model solution. These relevant contents could be those deemed necessary for a substantively correct and complete answer to the question. A score corresponding to the score given in the model solution could mean that the student response contains all the relevant contents of the model solution. It might be possible that only whole points can be awarded. You could be asked to add up the points you have assigned for each student response and to indicate only the total score for each student response.

Your ranking could then be presented like this (mention only the response and the score assigned):

Student answer_X: Points

Output:\n

Supplementary Table 26. Prompt P_{v8} - Point assessment

You are a professor of Macroeconomics and have been given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Supplementary Table 27. Prompt P_{v9} - Point assessment

You are an experienced student assistant for the subject of Macroeconomics and have been given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Supplementary Table 28. Prompt P_{v10} - Point assessment

You are a beginner-level student assistant with limited knowledge in Macroeconomics and have been given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Supplementary Table 29. Prompt P_{v11} - Point assessment

You are an assistant for the subject of Macroeconomics and have been given a task The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Present your point assessment as follows (state only the answer and the number of points awarded):

Student answer_X: Points

Output:\n

Supplementary Table 30. Prompt P_{v12} - Point assessment

You are a student assistant for macroeconomics and are given a task. The aim of this task is to award points for student answers to an open question in macroeconomics based on a sample solution. In the sample solution, the content for which a point is awarded is marked as follows: (1), (1 point) or (1P). The points (1), (1 point) or (1P) refer to the previous content.

This is the open question:

[INSERT_QUESTION]

This is the sample solution:

[INSERT_SAMPLE_SOLUTION]

These are five randomly selected student answers in random order:

Student answer_1: [STUDENT_ANSWER_1]

Student answer_2: [STUDENT_ANSWER_2]

Student answer_3: [STUDENT_ANSWER_3]

Student answer_4: [STUDENT_ANSWER_4]

Student answer_5: [STUDENT_ANSWER_5]

Your task is to award a number of points for each student answer. 0 points are also possible. The sample solution contains all relevant and correct content for the question. It is complete. No more points can be awarded than in the sample solution.

You can base your award of points on the respective proximity of the student's answer to the points awarded for the content of the sample solution. This is the relevant content that is necessary to answer the question correctly and completely. A score equal to the number of points awarded by the sample solution means that the student's answer contains all the relevant content of the sample solution. Only whole points can be awarded. Add up the points you have awarded for each student answer. State only the total number of points for each student answer.

Think step by step. You need to explain your points, and then present your ranking as follows (mention only the response and the score assigned):

Student answer_X: Points

Output:\n

S6 Robustness checks and extensions' extended result in rank and point assessment

Supplementary Table 31. Summary statistics of IRR w.r.t prompt design effect using Kendall's W in rank assessment across all questions.

Prompt Version	Human team	AI-human team 1	AI-human team 2	AI-human team 3
P_{v1}: Original$_{\text{RUN 1}}$		$0.789^{\pm 0.141}$	$0.806^{\pm 0.131}$	$0.787^{\pm 0.134}$
P_{v1}: Original$_{\text{RUN 2}}$	$0.792^{\pm 0.149}$	$0.774^{\pm 0.176}$	$0.802^{\pm 0.154}$	$0.761^{\pm 0.175}$
P_{v2}: Structure-1		$0.776^{\pm 0.177}$	$0.796^{\pm 0.158}$	$0.762^{\pm 0.177}$
P_{v3}: Structure-2		$0.756^{\pm 0.177}$	$0.788^{\pm 0.162}$	$0.748^{\pm 0.174}$
P_{v4}: Structure-3		$0.761^{\pm 0.185}$	$0.788^{\pm 0.176}$	$0.742^{\pm 0.192}$
P_{v5}: Structure-4		$0.753^{\pm 0.167}$	$0.786^{\pm 0.159}$	$0.755^{\pm 0.155}$
$P_{v6.1}$: Major Spelling Error		$0.761^{\pm 0.171}$	$0.787^{\pm 0.162}$	$0.749^{\pm 0.172}$
$P_{v6.2}$: Minor Spelling Error	N/A	$0.769^{\pm 0.176}$	$0.797^{\pm 0.162}$	$0.759^{\pm 0.179}$
P_{v7}: Suggestive and ambiguous instruction		$0.765^{\pm 0.180}$	$0.797^{\pm 0.169}$	$0.754^{\pm 0.184}$
P_{v8}: Role $\rightarrow$ Professor		$0.767^{\pm 0.182}$	$0.797^{\pm 0.160}$	$0.756^{\pm 0.180}$
P_{v9}: Role $\rightarrow$ Expert student assistant		$0.774^{\pm 0.169}$	$0.801^{\pm 0.156}$	$0.762^{\pm 0.167}$
P_{v10}:Role $\rightarrow$ Non-expert student assistant		$0.774^{\pm 0.166}$	$0.804^{\pm 0.154}$	$0.763^{\pm 0.163}$
P_{v11}:Role $\rightarrow$ Not given		$0.770^{\pm 0.176}$	$0.798^{\pm 0.160}$	$0.760^{\pm 0.178}$
P_{v12}:Advanced Prompting $\rightarrow$ Chain of thought		$0.770^{\pm 0.181}$	$0.798^{\pm 0.165}$	$0.760^{\pm 0.177}$

Standard deviation provided as superscript

Supplementary Table 32. Summary statistics of IRR w.r.t prompt design effect using Cronbach's Alpha in point assessment across all questions.

Prompt Version	Human team	AI-human team 1	AI-human team 2	AI-human team 3
P_{v1}: Original$_{\text{RUN 1}}$	0.944 [0.932, 0.954]	0.933 [0.918, 0.945]	0.932 [0.917, 0.944]	0.933 [0.919, 0.945]
P_{v1}: Original$_{\text{RUN 2}}$		0.930 [0.915, 0.943]	0.930 [0.915, 0.942]	0.928 [0.913, 0.941]
P_{v2}: Structure-1		0.927 [0.911, 0.940]	0.926 [0.910, 0.939]	0.926 [0.910, 0.939]
P_{v3}: Structure-2		0.921 [0.904, 0.935]	0.920 [0.903, 0.934]	0.920 [0.902, 0.934]
P_{v4}: Structure-3		0.928 [0.913, 0.941]	0.928 [0.912, 0.941]	0.924 [0.908, 0.938]
P_{v5}: Structure-4		0.920 [0.902, 0.934]	0.921 [0.904, 0.935]	0.918 [0.901, 0.933]
$P_{v6.1}$: Major Spelling Error		0.931 [0.916, 0.943]	0.930 [0.916, 0.943]	0.929 [0.914, 0.942]
$P_{v6.2}$: Minor Spelling Error	N/A	0.931 [0.916, 0.943]	0.931 [0.916, 0.944]	0.930 [0.916, 0.943]
P_{v7}: Suggestive and ambiguous instruction		0.927 [0.911, 0.940]	0.926 [0.910, 0.939]	0.926 [0.910, 0.940]
P_{v8}: Role $\rightarrow$ Professor		0.932 [0.917, 0.944]	0.933 [0.919, 0.945]	0.932 [0.918, 0.944]
P_{v9}: Role $\rightarrow$ Expert student assistant		0.930 [0.916, 0.943]	0.928 [0.913, 0.941]	0.929 [0.914, 0.942]
P_{v10}: Role $\rightarrow$ Non-expert student assistant		0.934 [0.920, 0.946]	0.934 [0.920, 0.946]	0.934 [0.920, 0.946]
P_{v11}:Role $\rightarrow$ Not given		0.930 [0.915, 0.943]	0.929 [0.914, 0.942]	0.928 [0.913, 0.941]
P_{v12}:Advanced Prompting $\rightarrow$ Chain of thought		0.921 [0.904, 0.935]	0.922 [0.906, 0.936]	0.918 [0.900, 0.932]

Confidence Interval at 95% provided in square brackets