

Leveraging generative AI in research and policy analysis



Workshop on Empirical Research in the AI Era, 16th December 2024
Banco de Portugal Microdata Research Laboratory (BPLIM), Academia of Data Science

Dr. Jesse Lastunen
Research Associate, UNU-WIDER
lastunen@wider.unu.edu

Generative AI and LLMs

- **Traditional (“narrow”)** AI focuses on specific tasks using predefined rules or logic
- **Generative AI** enables the creation of new content, and includes Large Language Models (LLMs) that specialize in language – like Gemini, Claude and OpenAI’s o1
- **LLMs** are based on transformer models with deep neural networks, benefiting from extensive pre-training and fine-tuning
- **Limitations** include factual inaccuracies, privacy concerns, algorithmic bias, reproducibility, etc.
- Despite limitations, LLMs offer **insightful responses** that can be helpful in a wide range of applications

12h15 – 13h00 **Reproducibility and AI: Potential, Challenges, and Other Preliminary Thoughts**
Lars Vilhuber, Cornell University, and American Economic Association

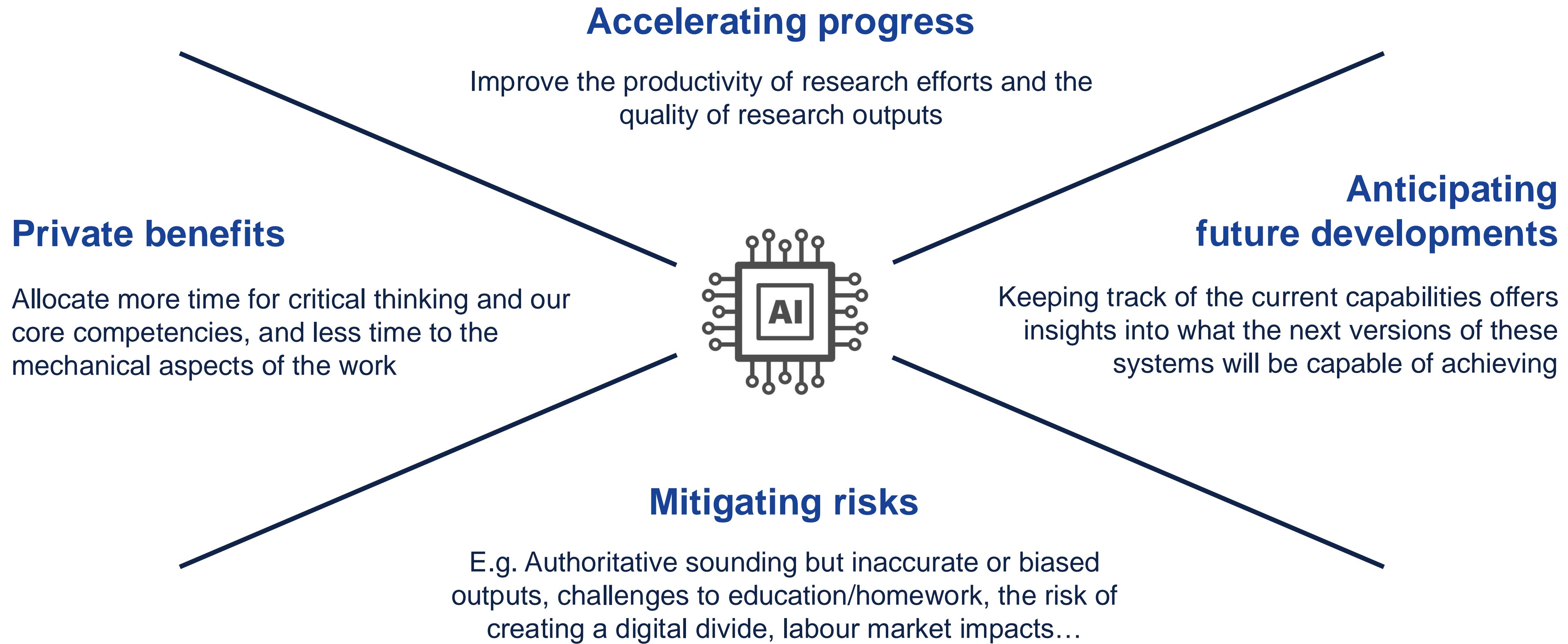
Prompts

- **Prompts are user-provided inputs** or queries that guide the model to generate specific, relevant responses
- **Prompt engineering** involves crafting detailed, context-rich queries to steer LLMs towards producing relevant content
- **Prompt generators** convert basic instructions describing a user's intent into a more specific set of instructions that induce LLMs to generate higher-quality output
 - Achieving good results with LLMs requires careful prompting and iterative refinement – users need to understand what they are asking

LLMs for researchers and policy analysts

- Recent advancements in GenAI, especially LLMs, have broadened the applicability of these tools in economics, policy analysis, and related work
- LLMs excel in understanding data structures and abstract concepts and thereby performing cognitive tasks relevant to researchers
- Commonly used tools/interfaces also come with built-in functionalities that allow users to upload files and perform data processing tasks
- Focus of this presentation: practical overview of LLMs in *research support tasks* instead of direct research applications, following [Korinek \(2023/2024\)](#) and using lessons from work at a UN-affiliated development economics research organization

Why should we care?



Commonly used LLMs as of December 2024

AI Lab	Model	Best version	Released	Tokens	Data cutoff	URL
Google	Gemini	2.0 Flash	Dec 2024	2M	Aug 2024	gemini.google.com
OpenAI	o1	o1	Dec 2024	128k	May 2024	chatgpt.com
Mistral	Mistral	24.11	Nov 2024	128k	Jan 2024	chat.mistral.ai
Meta	LLaMA	3.1	Jul 2024	128k	Dec 2023	meta.ai (not in Europe)
Anthropic	Claude	Sonnet 3.5	Jun 2024	200k	Apr 2024	claude.ai

- 2024 improvements with top tools (also beyond textual prompts/responses):
 - Growing context windows
 - Better recall
 - Improved visual information processing
 - Voice conversations
 - Thinking in steps
 - AI controlling your screen
 - Reasoning agents, etc.

Selected LLMs and interfaces

OpenAI's ChatGPT at chatgpt.com

- Use GPT-4o mini for free
- ChatGPT Plus subscription is around \$20/month, with access to more advanced models and tools
- Other benefits: Custom “GPTs”, image generation, plugins for expanded capabilities, advanced voice mode

Google's Gemini at gemini.google.com

- Use 1.5 Flash for free, with internet access, image generation, and photo uploads
- Premium subscription is around \$20/month, with access to 2.0 Flash experimental, 1.5 Pro with Deep Research, and others

Microsoft's Copilot at copilot.microsoft.com

- Sign up and use for free
- Access to OpenAI's GPT-4 Turbo model, web search, image generation (\$ for Office integration)
- Enterprise version protects your data

Anthropic's Claude at claude.ai

- Sign up and use the Sonnet 3.5 model for free, upload documents, internet access
- Claude Pro is \$20/month, with access to more advanced models, larger context window, etc.

Access to LLMs beyond chat interfaces

- **APIs:** integrate LLMs directly into your software applications
 - More advanced and customizable use cases, more flexibility and control
 - Pay-per-use basis, need to sign up for an API key, write code to interact with API endpoints
- **Real-time voice assistants**
- **Playgrounds** or experimentation platforms
 - Experiment with different model settings, like temperature; more control over formats
- + **Run open source LLMs locally on your own computer**
 - Offline access, private, cost-effective, customizable

Other tools relevant for researchers

Consensus at consensus.app

- A search engine that uses AI to find insights in +200M research papers
- ChatGPT integration through a GPT plugin e.g. to use academic work as source material in generated text with full citations

NotebookLM at notebooklm.google.com

- Upload and interact with Google Docs, PDFs, and web links
- Summarize, generate questions, and create notes with AI assistance

- **Other similar search engines:** elicit.com, scite.ai, ScholarAI.io, perplexity.ai
- **Computation, math:** gpt.wolfram.com OpenAI “GPT”

- **Some services allow you to use many different models, like** poe.com
- **Grammar and language:** trinka.ai, wordvice.ai

Ideation and feedback
Writing and editing
Background research
Promotion

Coding
Data analysis
Mathematical derivations

Key domains where LLMs help economists and policy analysts

This framework and some of the example prompts presented come from:

Korinek, Anton (2023)
“Generative AI for Economic Research: Use Cases and Implications for Economists”
Journal of Economic Literature
61 (4): 1281–1317

Fall 2024 Update: “Generative AI for Economic Research: LLMs Learn to Collaborate and Reason”

Ideation and feedback	<i>Brainstorming</i>	Coding	<i>Writing code</i>
	<i>Research strategy</i>		<i>Explaining code</i>
	<i>Feedback</i>		<i>Translating code</i>
	<i>Providing counterarguments</i>		<i>Debugging code</i>
Writing and editing	<i>Synthesizing text</i>	Data analysis	<i>Creating synthetic data</i>
	<i>Editing text</i>		<i>Creating scientific graphics</i>
	<i>Evaluating text</i>		<i>Creating figures</i>
	<i>Drafting academic text</i>		<i>Extracting data from text</i>
	<i>Generating titles and headlines</i>		<i>Reformatting data</i>
Background research	<i>Summarization</i>	Mathematical derivations	<i>Classifying and scoring text</i>
	<i>Literature search</i>		<i>Extracting sentiment</i>
	<i>Formatting references</i>		<i>Simulating human subjects</i>
	<i>Translating text</i>		<i>Setting up models</i>
	<i>Explaining concepts</i>		<i>Deriving equations</i>
Promotion	<i>Various use cases</i>		<i>Explaining models</i>

Prompt responses here are generated using:

1. OpenAI's API ("o1-mini-2024-09-12", i.e. o1-mini)

- Requires paid subscription; deterministic and thus reproducible outputs

2. The Plus/Pro ChatGPT & Gemini web interfaces (*different OpenAI & Gemini models*)

- Require paid subscriptions; results will differ every time a prompt is executed
- Needed for GPTs, online search, uploading files, data analysis, etc.

→ Similar results can be expected with Claude, Mistral, etc.

→ The most trivial examples work relatively well with free versions and smaller models run locally

Ideation and feedback

Brainstorming

Research strategy

Feedback

Providing counterarguments

Ideation and feedback: *Useful in all research settings*

- LLMs can generate a wide range of ideas from simple inputs
- ***Iterative process:*** Generate, evaluate, and refine ideas
- ***Key benefit:*** Outside perspective, support for overcoming confirmation/selection bias
- ***More recent:*** Evaluate entire paper sections/drafts or compare several papers
- **Caveats:**
 - Risk of homogenizing ideas and reducing critical thinking
 - Feedback on complex topics may be overly generic or misleading

PROMPT: I am using microsimulation to evaluate the 2023/24 personal income tax reform in Rwanda on inequality. Can you brainstorm an outline for the paper I'm writing?

RESPONSE: tentative outline:

(o1 via the ChatGPT Plus web interface, formatting added; last bullet points omitted)

1. introduction

- brief background on the personal income tax system in Rwanda before the 2023/24 reform
- rationale for examining the tax changes and their intended goals
- significance of studying how these reforms affect inequality indicators

2. policy context and literature background

- overview of key features of the 2023/24 personal income tax reform
- summary of existing research on personal income taxation in Rwanda and similar contexts
- discussion of previous findings on how income tax reforms have shaped income distribution

3. data and microsimulation approach

- description of the data sources, including household income surveys and administrative tax records
- explanation of the microsimulation model and how it is calibrated to the Rwandan context
- assumptions about behavioral responses and how they are integrated into the model

4. baseline scenario and reform simulations

- description of the baseline scenario (pre-reform system) used as a reference point
- step-by-step outline of the policy changes implemented in the simulation for the 2023/24 reform
- discussion of how each policy shift is expected to alter the income distribution

5. results

6. robustness checks and sensitivity analysis

7. conclusion

PROMPT:

I asked you earlier to create an outline for the research I'm doing on the PIT reform in Rwanda.

Could you provide some **viewpoints I should consider when writing about the rationale for examining these tax changes and their intended goals.**

For context, the reform increased the tax-exempt threshold in 2023 and lowered marginal rates for the top quartile. Be brief.

RESPONSE:

(GPT-4o via the ChatGPT Plus web interface

- Assess whether the reform achieves its intended goals of reducing inequality and improving progressivity in Rwanda's tax system.
- Understand the trade-offs between revenue generation and social equity, especially in the context of Rwanda's development priorities.
- Analyze the implications of shifting tax burdens across income groups, with a focus on the top quartile and lower-income earners.
- Examine how the reform aligns with broader poverty reduction and inclusive growth strategies in Rwanda.
- Provide evidence on the effectiveness of PIT reforms in addressing inequality in a low-income country context, offering insights for future policy design.

PROMPT:

I'm working on an economic **model that predicts that eliminating VAT exemptions will increase inequality** in a developing country. What are the top 5 counterarguments? One line each.

RESPONSE:

*(Gemini
Advanced v2.0
Flash;
first and last
sentence
omitted)*

- 1. Informal sector dominance:** If a large share of consumption by the poor occurs in the informal sector, which already evades VAT, removing formal exemptions has little impact on their relative tax burden.
- 2. Targeted compensation:** Revenue gains from eliminating exemptions can finance targeted social programs (cash transfers, food subsidies) that specifically benefit low-income households, offsetting any regressive impact.
- 3. Reduced tax evasion:** A broader VAT base with fewer exemptions simplifies tax administration and reduces incentives for evasion, potentially leading to more equitable revenue collection overall.
- 4. Progressive expenditure incidence:** If the exempted goods and services are disproportionately consumed by wealthier households, removing exemptions could actually make the tax system more progressive.
- 5. Improved public services:** Increased VAT revenue can fund essential public services (education, healthcare) that disproportionately benefit the poor, indirectly reducing inequality in the long run.

PROMPT:

I'd like to estimate the causal effect of **school fee reductions** on **enrollment rates** in Vietnam, implemented in 36 out of 96 provinces in phases from 2014 to 2016. I have data on school enrollment rates in all provinces from 2010 to 2019.

Which research strategy would you recommend? Write it up as a Latex document, focusing on the equations while keeping other explanations brief.

RESPONSE:

(*Gemini Advanced v2.0 Flash; LaTeX output; discussion on identification and robustness omitted*)

Estimating the Effect of School Fee Reductions in Vietnam

Research Question: Estimate the causal effect of school fee reductions on enrollment rates, implemented in 36/96 provinces from 2014-2016.

Data: Provincial school enrollment rates (2010-2019).

Methodology: Difference-in-Differences (DID) with staggered adoption.

Basic DID Model: $Enrollment_{it} = \alpha + \beta \cdot Treatment_i \cdot Post_t + \gamma \cdot Treatment_i + \delta \cdot Post_t + X'_{it}\theta + \varepsilon_{it}$

Where:

- $Enrollment_{it}$: Enrollment rate in province i at time t .
- $Treatment_i$: 1 if treated, 0 otherwise.
- $Post_t$: 1 post-2014, 0 before.
- β : Causal effect.

Staggered DID (Event Study): $Enrollment_{it} = \alpha + \sum_{k=-4}^5 \beta_k \cdot Treatment_i \cdot D_{it}^k + \gamma \cdot Treatment_i + \delta_t + X'_{it}\theta + \varepsilon_{it}$

Where:

- D_{it}^k : Dummy for years relative to treatment.
- β_k : Effect k years relative to treatment (relative to $k = -1$).
- δ_t : Year fixed effects.

Writing and editing

Synthesizing text

Editing text

Evaluating text

Drafting academic text

Generating titles and headlines

Writing and editing: *Streamline the writing process*

- LLMs excel in generating coherent text from fragmented inputs and editing text
- ***Enhance existing text*** with additional context or examples
- ***Create*** engaging paper titles, headlines, or summaries in another style
- ***Evaluate and improve*** grammar, spelling, and clarity, or identify ambiguities
- **Caveats:**
 - Generated text may lack human nuances
 - Outputs require thorough evaluation and amendment

PROMPT: Write a paragraph in the style of an academic economist that integrates the following points. Include a topic sentence.

- progressive taxation plays a big role in mitigating income inequality in developing countries
- the effectiveness of such systems is often hampered by low compliance rates, large informal sector, etc. **[discuss additional challenges]**
- economists must explore strategies like leveraging tech for better tax enforcement **[offer additional solutions]**
- it is also important to consider international cooperation to combat tax evasion **[also offer specific ways in which this can be done]**
- conclude by stating a balance must be struck between ensuring tax fairness and avoiding overly burdensome taxation **[discussing why]**

RESPONSE: Progressive taxation is instrumental in mitigating income inequality in developing countries by ensuring that higher earners contribute proportionally more to public revenue, thereby enabling the provision of essential services and social safety nets. However, the effectiveness of progressive tax systems in these contexts is frequently undermined by challenges such as low compliance rates, the prevalence of large informal sectors, limited administrative capacity, and inadequate legal frameworks. To enhance the efficacy of progressive taxation, economists must explore strategies such as broadening the tax base to include previously untaxed activities, leveraging digital technologies for improved tax collection and enforcement, and implementing simplified tax regimes that reduce the compliance burden. Additionally, fostering international cooperation is crucial to combat tax evasion, which can be achieved through the establishment of global information-sharing agreements, harmonizing tax policies to prevent base erosion and profit shifting, and supporting capacity-building initiatives in developing nations. Ultimately, a delicate balance must be struck between ensuring tax fairness and avoiding overly burdensome taxation, as excessive tax pressures can disincentivize investment and economic growth, thereby potentially exacerbating the very inequalities the tax system seeks to alleviate.

(01-mini via ChatGPT API)

PROMPT:

Give a concise list of the top 5 stylistic shortcomings of the following abstract:

In the wake of the COVID-19 pandemic, several countries enacted tax and social protection measures to help mitigate the economic hardship faced by individuals and households. This experience underscores the need to better understand the impact of such programmes on incomes and poverty during crises, especially in low- and middle-income countries (LMICs) where they are most needed. This paper reviews existing empirical literature on the subject, conducting a scoping review on quantitative studies published between 2000 and 2022. Following a structured selection approach, we identify 39 studies about the role of social protection and taxation in LMICs during periods of crisis. The results reveal that LMICs often enact both vertical and horizontal expansions of existing unconditional cash transfer schemes during crises, although the applications vary by geography. Our research also suggests a sharp increase in recent studies because of the COVID-19 pandemic, and a large variety of methods used to study the subject. In terms of effects, social protection can help cushion households against crises, but the effect size depends on the policy adopted, country context, and type of crisis. The minimal empirical evidence about tax policy studies do not enable us to draw conclusions.

RESPONSE:
*(01-mini via
ChatGPT API)*

1. **Typographical Error:** The word "variety" is misspelled and should be corrected to "variety," which detracts from the professional presentation.
2. **Redundancy and Wordiness:** Phrases like "in the wake of" and "underscores the need to better understand" could be more concise, making the abstract clearer and more direct.
3. **Use of Unexplained Jargon:** Terms such as "vertical and horizontal expansions" are specialized and may not be immediately understandable to all readers without further clarification.
4. **Inconsistent Sentence Structure:** The abstract shifts between different sentence structures and perspectives, which can disrupt the flow and coherence of the text.
5. **Lack of Clarity in Conclusions:** The final sentence is vague regarding the implications of the minimal empirical evidence on tax policy, leaving the reader uncertain about the study's definitive outcomes.

Other use cases in ‘writing and editing’

Editing text

PROMPT: Can you rewrite the abstract to correct for shortcomings you identified?

PROMPT: Can you rewrite this paragraph to make it attractive and understandable to non-academic audiences?

Evaluating text (with advanced tools allowing for attachments)

PROMPT: Provide 5 obvious improvements that this paper needs before I can submit it to a journal. [attach the paper]

PROMPT: Compare the methods used in the two papers attached. [attach the papers]

Drafting academic text (with advanced tools allowing for web/article search) – use with care

PROMPT: I want to write a **literature review for my research on VAT compliance** among informal small firms in Kigali. It will cover several topics, but I would appreciate your help on the **subsection on VAT compliance in Rwanda, with possible comparisons to other countries**. Write the **first draft of this section** of the literature review based on relevant academic literature, where possible from high-impact journals. Make it comprehensive, with several references/sources and ample detail.

Background research

Summarizing text

Literature search

Formatting references

Translating text

Explaining concepts

Background research: *Micro tasks behind the core work*

- LLMs help streamline background research, like literature search, in various ways
- **Simple:** Format references across styles, translate texts, explain concepts
- **More complex:** Find and cite relevant papers, search dozens of sources to get an answer to a question (esp. with Gemini Pro 1.5 with Deep Research)
- **Caveats:**
 - Potential for fictional (or irrelevant) references
 - Verification is crucial
 - Outputs can be shallow

PROMPT:

Find 5 papers from reputable economics journals that analyze the effects of conditional cash transfer programs on poverty or inequality in African or Latin American countries. For now, I only need the title, authors, and journal as well as a two-sentence summary.

RESPONSE:

(GPT-4o via
ChatGPT Plus
web interface,
incl. web search
and links to
JSTOR;
formatting
edited)

Here are five studies [...]

1. “Conditional Cash Transfers, Resources, and Political Participation in Latin America”

Author: Gregory S. Schober, **Journal:** Latin American Research Review (2019)

Summary: This study investigates how CCT programs in Latin America influence political participation. The findings suggest that recipients of CCTs are more likely to engage in political activities, indicating that such programs can enhance democratic engagement among the poor.

2. “Cash Transfers in Latin America: Effects on Poverty and Redistribution”

Authors: Nora Lustig, Carola Pessino, John Scott, **Journal:** Economía (2014)

Summary: This paper evaluates the impact of cash transfer programs on poverty and income distribution in Latin America. The authors find that these programs have significantly reduced poverty and inequality, with varying degrees of effectiveness across different countries.

3. “Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua”

Author: Ana C. Dammert, **Journal:** Economic Development and Cultural Change (2009)

Summary: This research examines the varied effects of Nicaragua's CCT program on child labor and school attendance. It shows that the program effectively reduces child labor and increases school attendance, with more pronounced benefits for certain demographic groups.

4. “Conditional Cash Transfers as a Tool of Social Policy”

Author: Armando Barrientos, **Journal:** Economic and Political Weekly (2011)

Summary: This article discusses the role of CCT programs in social policy, focusing on their design and implementation in low- and middle-income countries, particularly in Latin America. It highlights the importance of program design in achieving desired social outcomes.

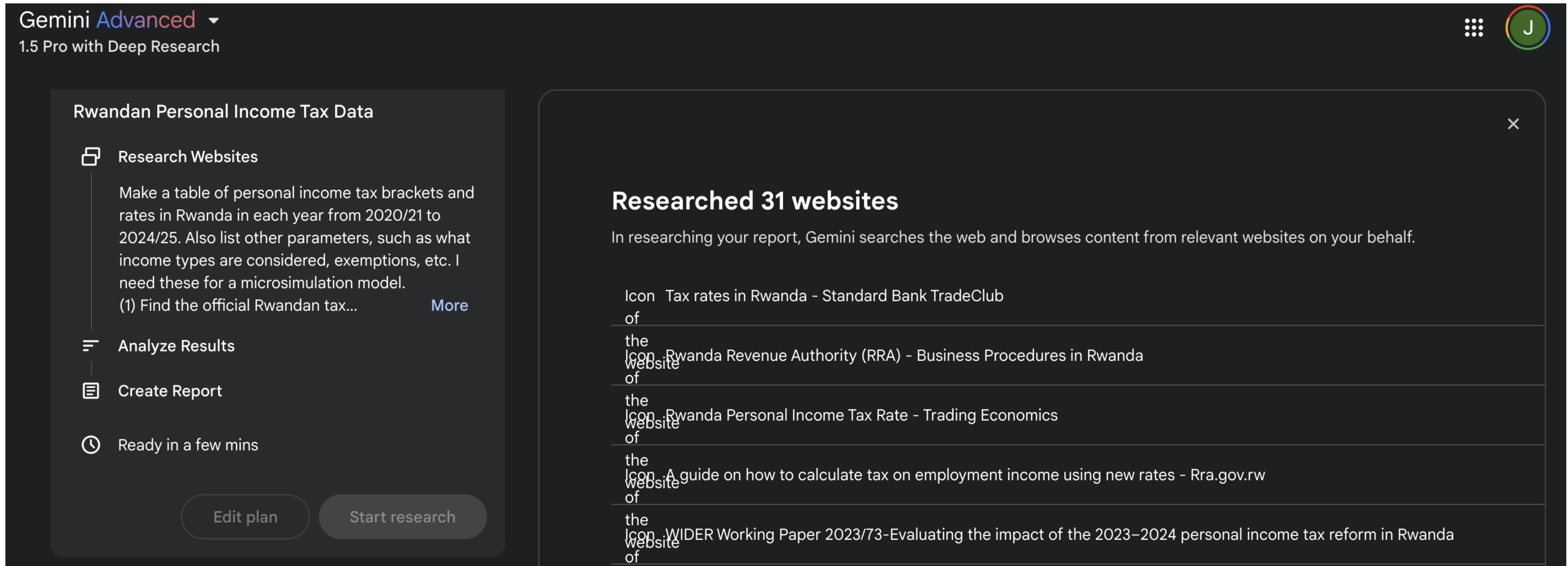
5. “Evaluating the Impact of Conditional Cash Transfer Programs: Lessons from Latin America”

Authors: Laura B. Rawlings, Gloria M. Rubio, **Journal:** World Bank Research Observer (2005)

Summary: This paper assesses the effectiveness of CCT programs in Latin America in reducing poverty and promoting human capital development. The authors conclude that these programs have been successful in improving health and education outcomes among the poor.

These studies provide valuable insights into the effectiveness of CCT programs in addressing poverty and inequality in African and Latin American contexts.

PROMPT: Make a table of personal **income tax brackets and rates in Rwanda** in each year from 2020/21 to 2024/25. Also list other parameters, such as what income types are considered, exemptions, etc. I need these for a microsimulation model.



The screenshot shows the Gemini Advanced 1.5 Pro with Deep Research interface. On the left, a sidebar titled "Rwandan Personal Income Tax Data" lists three main sections: "Research Websites", "Analyze Results", and "Create Report". Under "Research Websites", there is a detailed description of the prompt and two buttons: "Edit plan" and "Start research". The "Start research" button is highlighted with a light gray background. On the right, a large panel displays the results of the search. It starts with the heading "Researched 31 websites" and a brief explanatory text. Below this, five website entries are listed, each consisting of an icon, the website name, and a truncated URL. The entries are: "Tax rates in Rwanda - Standard Bank TradeClub", "Rwanda Revenue Authority (RRA) - Business Procedures in Rwanda", "Rwanda Personal Income Tax Rate - Trading Economics", "A guide on how to calculate tax on employment income using new rates - Rra.gov.rw", and "WIDER Working Paper 2023/73-Evaluating the impact of the 2023–2024 personal income tax reform in Rwanda".

Gemini Advanced ▾
1.5 Pro with Deep Research

Rwandan Personal Income Tax Data

Research Websites

Make a table of personal income tax brackets and rates in Rwanda in each year from 2020/21 to 2024/25. Also list other parameters, such as what income types are considered, exemptions, etc. I need these for a microsimulation model.

(1) Find the official Rwandan tax... More

Analyze Results

Create Report

Ready in a few mins

Edit plan Start research

Researched 31 websites

In researching your report, Gemini searches the web and browses content from relevant websites on your behalf.

- Icon Tax rates in Rwanda - Standard Bank TradeClub
of the Icon Rwanda Revenue Authority (RRA) - Business Procedures in Rwanda
- the Icon Rwanda Personal Income Tax Rate - Trading Economics
website of the
- Icon A guide on how to calculate tax on employment income using new rates - Rra.gov.rw
website of the
- Icon WIDER Working Paper 2023/73-Evaluating the impact of the 2023–2024 personal income tax reform in Rwanda
website of the

(Gemini Advanced 1.5 Pro with Deep Research; Gemini web interface)

PROMPT:

Make a table of personal **income tax brackets and rates** in Rwanda in each year from 2020/21 to 2024/25. Also list other parameters, such as what income types are considered, exemptions, etc. I need these for a microsimulation model.

RESPONSE:

(*Gemini Advanced Pro 1.5 with Deep Research; Gemini web interface; some parts of the Google doc created)*

Personal Income Tax in Rwanda (2020/21-2024/25)

This report provides a comprehensive overview of the personal income tax system in Rwanda, covering the tax years 2020 to 2024. It includes details on tax brackets, rates, taxable income, exemptions, deductions, and tax administration procedures. This information is crucial for constructing a microsimulation model and understanding the Rwandan tax landscape.

Personal Income Tax Brackets and Rates

The Rwandan personal income tax system is structured with progressive tax rates, meaning the tax rate increases as the taxable income increases. This system aims to distribute the tax burden more equitably across different income levels¹. The following tables summarize the personal income tax brackets and rates in Rwanda for each year from 2020 to 2024:

2020–2022

Annual Taxable Income (RWF)	Tax Rate
0 - 360,000	0%
360,001 - 1,200,000	20%
1,200,001 and above	30%

2023

Annual Taxable Income (RWF)	Tax Rate
0 - 720,000	0%
720,001 - 1,200,000	20%
1,200,001 and above	30%

2024

Taxable Income

In Rwanda, taxable income generally includes employment income, such as salaries, wages, bonuses, and benefits in kind. While the provided research material does not offer a definitive list of inclusions and exclusions, it is generally understood that income derived from employment activities is subject to personal income tax⁴.

It's important to consider the tax treatment of benefits in kind. These benefits, often provided by employers, are also considered part of taxable income. For example, housing benefits are taxed at 20% of the total employment income (excluding other benefits in kind), while motor vehicle benefits are taxed at 10% of employment income (excluding other benefits in kind). When an employer directly pays for an employee's house or motor vehicle rent, this payment is treated as a taxable allowance⁵.

Exemptions, Deductions, and Credits

The Rwandan tax system provides various exemptions, deductions, and credits that can potentially reduce an individual's tax liability. These provisions are designed to provide relief to taxpayers and encourage specific behaviors, such as saving for retirement or investing in the local economy⁶. These provisions can have a significant impact on different income groups, potentially reducing the tax burden for those who qualify.

Category	Description	Citation
Employment Expenses	Individuals can deduct expenses that are wholly and exclusively incurred in the production of their employment income. This includes expenses like travel and car expenses, as long as they are directly related to the individual's work and not reimbursed by the employer.	⁸
Retirement Contributions	Individuals can deduct contributions made to a qualified pension fund. This deduction is limited to the lower of 10% of the individual's	⁶

Works cited

1. A guide on how to calculate tax on employment income using new rates - Rra.gov.rw, accessed December 14, 2024, https://www.rra.gov.rw/en/details?tx_news_pi1%5Baction%5D=detail&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Bnews%5D=1669&cHash=8a281f98a1e1d9501765985f3a91fe8d
2. Rwanda's New Personal Income Tax Brackets/Rates Effective November 2023 — Orbitax Tax News & Alerts, accessed December 14, 2024, <https://orbitax.com/news/archive.php/Rwandas-New-Personal-Income-T-54125>
3. English: Personal Income Tax (PIT) - Rra.gov.rw, accessed December 14, 2024, <https://www.rra.gov.rw/en/taxes-fees/domestic-taxes/income-tax/personal-income-tax-pit-1>
4. English: Domestic Tax Services - Rra.gov.rw, accessed December 14, 2024, <https://www.rra.gov.rw/en/domestic-tax-services>
5. Rwanda - Individual - Income determination - PwC Tax Summaries, accessed December 14, 2024, <https://taxsummaries.pwc.com/rwanda/individual/income-determination>
6. Tax rates in Rwanda - Standard Bank TradeClub, accessed December 14, 2024, <https://www.tradeclub.standardbank.com/portal/en/market-potential/rwanda/taxes>
7. English: Income tax exemption - Rra.gov.rw, accessed December 14, 2024, https://www.rra.gov.rw/en/details?tx_news_pi1%5Baction%5D=detail&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Bnews%5D=2193&cHash=8b0160d1a4cd96180ad87086837e6b4
8. Rwanda - Individual - Deductions - Worldwide Tax Summaries Online, accessed December 14, 2024, <https://taxsummaries.pwc.com/rwanda/individual/deductions>
9. Rwanda gazettes new Income Tax Law - Tax News Update, accessed December 14, 2024, <https://taxnews.ey.com/news/2022-1782-rwanda-gazettes-new-income-tax-law>
10. corporate and personal income tax 2023 declaration and payment - Rra.gov.rw, accessed December 14, 2024, https://www.rra.gov.rw/fileadmin/user_upload/RRA_FAQs_INCOME_TAX_2024.pdf
11. Rwanda - Individual - Tax administration - Worldwide Tax Summaries Online, accessed December 14, 2024, <https://taxsummaries.pwc.com/rwanda/individual/tax-administration>
12. RRA urges taxpayers for early 2023 Annual Income Tax declarations for timely support, accessed December 14, 2024, https://www.rra.gov.rw/en/details?tx_news_pi1%5Baction%5D=detail&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Bnews%5D=1731&cHash=d42a800abccbda53ef5e9c5b5018f83d
13. English: Pay Income Tax - RRA, accessed December 14, 2024, <https://www.rra.gov.rw/en/taxes-fees/domestic-taxes/income-tax/pay-income-tax>

(*Gemini Advanced 1.5 Pro with Deep Research; Gemini web interface*)

Formatting references

Other use cases in ‘background research’

PROMPT: Convert this reference to bibtex format:

Gibson Masumbu and Kristi Mahrt: Assessing Progress in Welfare Improvements in Zambia: A Multidimensional Approach. In book: Growth and Poverty in Sub-Saharan Africa. Editors: Channing Arndt, Andy McKay, Finn Tarp. April 2016, Oxford University Press, Number of Pages 528

Translating text

PROMPT: Translate the following sentence into Spanish, German, and Finnish:

“This book comprehensively evaluates trends in living conditions in 16 major sub-Saharan African countries, corresponding to nearly 75% of the total population.”

Explaining concepts

PROMPT: How should commodity taxes be set in a developing country context as a part of the overall tax-benefit system? Discuss the normative aspects involved on the basis of public economics theory. Also comment on the relevant empirical information one would need to have to be able to make policy recommendations.

PROMPT: What factors determine the elasticity of intertemporal substitution, and why does it matter for understanding savings behavior and fiscal policy effectiveness?

Coding

Writing code

Explaining code

Translating code

Debugging code

Creating synthetic data

Creating scientific graphics

Coding: *An LLM as your everyday coding assistant*

- LLMs help write, edit, optimize, and debug code in Python, R, STATA, etc.
- Best used as an everyday coding assistant
- Helps explain your colleague's badly commented code in plain English
- Works well as a tutor for unfamiliar programming languages and libraries
- **Caveats:**
 - Code outputs may contain errors, necessitating iteration and thorough user review

09h30 – 10h15 **Integrating AI in Stata Programming: Perspectives on Enhancement and Constraints Across Skill Levels**
Ricardo Mora, Universidad Carlos III Madrid

PROMPT:

I have an output dataset from a Vietnamese microsimulation model, stored in "/Users/user/Desktop/vn_2023_std.txt". **Help me write efficient Stata code for four tasks:**

1. to insheet the txt file to Stata
2. to obtain a histogram of the distribution of variable 'ils_dispyx' (disposable income), among individuals between the ages of 18 and 64 (age='dag')
3. to calculate the share of individuals with zero disposable income
4. to calculate average disposable income among informal ('lfo'=0) and formal ('lfo'=1) workers.

Where required, use survey weights (variable 'dwt'). Provide the entire code that I can copy-paste to Stata to accomplish all tasks, with no other text.

Avoid using the svyset package, just show how to make the calculations.

RESPONSE (GPT-4 Turbo via Microsoft Copilot, Precise mode):

Here is the Stata code for your tasks:

```
// Task 1: Insheet the txt file to Stata
insheet using "/Users/user/Desktop/vn_2023_std.txt", clear

// Task 2: Obtain a histogram of the distribution of variable 'ils_dispyx' among
individuals between the ages of 18 and 64
histogram ils_dispyx if dag >= 18 & dag <= 64, frequency

// Task 3: Calculate the share of individuals with zero disposable income
gen zero_income = ils_dispyx == 0
egen total_weight = total(dwt)
egen zero_income_weight = total(dwt) if zero_income == 1
gen zero_income_share = zero_income_weight / total_weight
sum zero_income_share

// Task 4: Calculate average disposable income among informal and formal workers
egen informal_weight = total(dwt) if lfo == 0
egen formal_weight = total(dwt) if lfo == 1
egen informal_income = total(ils_dispyx * dwt) if lfo == 0
egen formal_income = total(ils_dispyx * dwt) if lfo == 1
gen avg_informal_income = informal_income / informal_weight
gen avg_formal_income = formal_income / formal_weight
sum avg_informal_income avg_formal_income
```

Please replace the file path with the actual path of your file if it's different. Also, make sure that the variable names match those in your dataset. If there are any issues, please let me know!
I'm here to help. 😊

PROMPT: Explain this Stata code in 10 short bullets:

```

* Folder with output datasets:
global output "/Users/user/Desktop"
* Years, definitions, and income concepts:
global years = "2022 2023"
global Nyear: word count ${years}
global definitions = "dispyx con"
global Ndef: word count ${definitions}
* Poverty rate matrix:
mat povR = J(${Ndef}, ${Nyear},.)
mat colnames povR = ${years}
mat rownames povR = "Disposable income" "Consumption"
matnames povR
mat list povR
* Loop to calculate poverty rates:
local col = 1
foreach year of global years {
    insheet using "${output}/vn_`year'_std.txt", clear
    * Equivalence scale (national definition)
    bysort idhh: gen ses_hh = sum(ses)
    * Household level equivalised income and consumption
    quietly foreach def of global definitions {
        bysort idhh: egen temp = total(ils_`def')
        gen `def'_equ = temp / ses_hh
        gen `def'_neg = `def'_equ
        replace `def'_neg = 0 if `def'_neg<0
        drop temp
    }
    * Matrix:
    local row = 1
    foreach def of global definitions {
        povdeco `def'_neg [aw=dwt], varpline(spl)
        mat povR[`row', `col'] = r(fgt0) * 100
        local row = `row'+1
    }
    local col = `col'+1
}

```

RESPONSE (*o1-mini via the ChatGPT Plus web interface*):

- Set Output Directory:** Defines a global macro output pointing to "/Users/user/Desktop" for storing input and output datasets.
- Specify Years:** Creates a global macro years with the values "2022 2023" and counts the number of years (Nyear = 2).
- Define Income Concepts:** Establishes a global macro definitions with "dispyx con" and counts the definitions (Ndef = 2).
- Initialize Poverty Matrix:** Creates a matrix povR with Ndef rows and Nyear columns, naming rows as "Disposable income" and "Consumption" and columns as the specified years.
- Loop Initialization:** Sets a local counter col to 1 to track the current column in the povR matrix.
- Import Yearly Data:** For each year in years, loads the corresponding dataset (e.g., vn_2022_std.txt) from the output directory using insheet.
- Calculate Equivalence Scale:** Groups data by household ID (idhh) and generates ses_hh, the sum of socio-economic status (ses) for each household.
- Compute Equivalised Measures:** For each income definition, calculates total income/consumption per household, divides by ses_hh to get equivalised values, sets negative values to zero, and removes temporary variables.
- Calculate Poverty Rates:** For each income definition, runs the povdeco command with the equivalised and non-negative income/consumption, applies sampling weights (dwt), retrieves the poverty rate (fgt0), converts it to a percentage, and stores it in the povR matrix.
- Update Matrix and Finalize:** Increments the column counter after each year, resulting in the povR matrix containing poverty rates (%) for "Disposable income" and "Consumption" across 2022 and 2023.

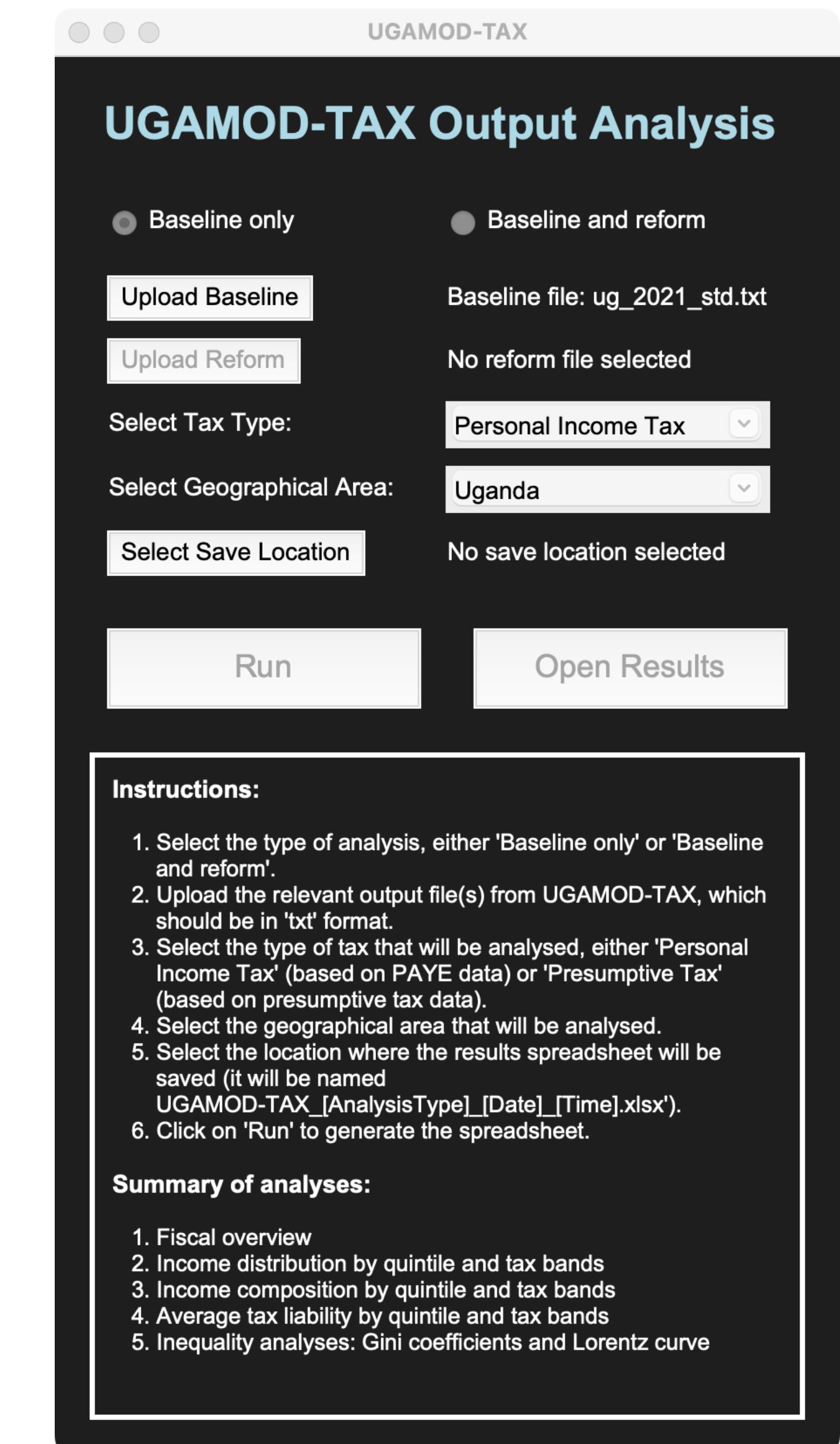
SET OF PROMPTS:

Write Python code for a simple program interface:

1. The user can upload a **Baseline** and/or **Reform** output file, produced with a tax-benefit microsimulation model (txt files)
2. User can select '**Baseline only**' or '**Baseline and reform**'
3. User can select the '**Tax Type**' analysed ('Personal Income Tax' or 'Presumptive Tax')
4. User can select the '**Geographical Area**' analysed (Uganda, North, South, West, East)
5. There's a '**Run**' button
6. There's an '**Open Results**' button
7. At the bottom, there is a **text field with instructions** and summary of analysis (I'll specify this later).

Let's do this in steps. Note that after we are done with finished the interface, I'll provide you with more information about the analysis that should occur when the user clicks on 'Run', and the files that will be produced.

Draft UI after an hour of iteration →
(01-Preview via ChatGPT Plus web interface)



Other use cases in ‘coding’

PROMPT: I create a variable ‘temp’ inside a loop like this in Stata: [...]. I received an error message: variable temp already defined r(110); How do I correct the error? Please **debug**.

PROMPT: **Translate** this Stata code to R: [...]

PROMPT: I will soon get access to data on school enrollment rates in all the 96 provinces from 2010 to 2019, along with information on household income and average parental education.
You suggested a staggered DID approach as a suitable research strategy.
Since I don't have the actual data yet, could you please provide me with Stata code to **generate a synthetic dataset** that I can use to test my methods? You can assume 10 schools per province

PROMPT: Provide **LaTeX code to create a graph** depicting the basic elements of a supply and demand model in an economy. The graph should include a supply curve and a demand curve (straight lines) as well as the equilibrium price and quantity, with dashed vertical and horizontal lines from the point where the curves intersect.

Mathematical derivations

Setting up models

Deriving/solving equations

Explaining models

Mathematical derivations

- LLMs are becoming increasingly better at mathematical reasoning and application
- Set up common models and solve basic equations
- Explain simple mathematical models and equations, aiding educational efforts
- **Caveats:**
 - Not yet consistently reliable for high-level derivations; careful oversight needed
 - May eventually impact manual problem-solving skills

PROMPT:

Set up the optimization problem of a consumer with CES preferences over two goods in LaTeX code.

RESPONSE (ChatGPT API):

[LaTeX code]

Copy-paste to LaTeX editor

Consumer Optimization Problem with CES Preferences

Consider a consumer with the following CES utility function over two goods, x_1 and x_2 :

$$U(x_1, x_2) = (\alpha x_1^\rho + (1 - \alpha)x_2^\rho)^{\frac{1}{\rho}} \quad (1)$$

where $0 < \alpha < 1$ is the share parameter, and $\rho \neq 1$ is the substitution parameter that determines the elasticity of substitution between the two goods.

The consumer's budget constraint is given by:

$$p_1x_1 + p_2x_2 = I \quad (2)$$

where p_1 and p_2 are the prices of goods 1 and 2, respectively, and I is the consumer's income.

The consumer's optimization problem is to choose x_1 and x_2 to maximize utility subject to the budget constraint:

$$\max_{x_1, x_2} U(x_1, x_2) = \max_{x_1, x_2} (\alpha x_1^\rho + (1 - \alpha)x_2^\rho)^{\frac{1}{\rho}} \quad (3)$$

subject to

$$p_1x_1 + p_2x_2 \leq I \quad (4)$$

This problem can be solved using the Lagrange multiplier method, where the Lagrangian is:

$$\mathcal{L}(x_1, x_2, \lambda) = (\alpha x_1^\rho + (1 - \alpha)x_2^\rho)^{\frac{1}{\rho}} - \lambda(p_1x_1 + p_2x_2 - I) \quad (5)$$

The first-order conditions for a maximum are:

$$\frac{\partial \mathcal{L}}{\partial x_1} = \alpha(\alpha x_1^\rho + (1 - \alpha)x_2^\rho)^{\frac{1}{\rho}-1} x_1^{\rho-1} - \lambda p_1 = 0 \quad (6)$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = (1 - \alpha)(\alpha x_1^\rho + (1 - \alpha)x_2^\rho)^{\frac{1}{\rho}-1} x_2^{\rho-1} - \lambda p_2 = 0 \quad (7)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = p_1x_1 + p_2x_2 - I = 0 \quad (8)$$

Solving these equations simultaneously will yield the optimal consumption bundle (x_1^*, x_2^*) .

PROMPT:

Can you explain what this model is, how to get from one step to the next, and what the intuition for each step is? Provide the full response in LaTeX format.

[Equations pasted in LaTeX format.]

$$\begin{aligned} & \max_{C(t), K(t)} \int_0^{\infty} e^{-\rho t} U(C(t)) dt \\ \text{s.t. } & \dot{K}(t) = F(K(t), L) - \delta K(t) - C(t) \\ & K(0) \text{ given} \\ & \lim_{t \rightarrow \infty} K(t)e^{-\rho t} \geq 0 \end{aligned}$$

$$\mathcal{H}(t, K(t), C(t), \lambda(t)) = e^{-\rho t} U(C(t)) + \lambda(t)[F(K(t), L) - \delta K(t) - C(t)]$$

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial C(t)} = 0 & \Rightarrow U'(C(t)) = \lambda(t) \\ \frac{\partial \mathcal{H}}{\partial K(t)} + \dot{\lambda}(t) = 0 & \Rightarrow \dot{\lambda}(t) = -\lambda(t)[F'(K(t), L) - \delta] \\ \lim_{t \rightarrow \infty} \lambda(t)K(t) &= 0 \end{aligned}$$

RESPONSE (GPT-4 via the ChatGPT Plus web interface):

[LaTeX code]

Copy-paste to LaTeX editor

Explanation of the Optimal Control Model

Consider the following optimal control problem in economics:

$$\begin{aligned} & \max_{C(t), K(t)} \int_0^{\infty} e^{-\rho t} U(C(t)) dt \\ \text{s.t. } & \dot{K}(t) = F(K(t), L) - \delta K(t) - C(t) \\ & K(0) \text{ given} \\ & \lim_{t \rightarrow \infty} K(t)e^{-\rho t} \geq 0 \end{aligned}$$

Model Explanation: This model represents a typical Ramsey-Cass-Koopmans model in economics, used for studying optimal consumption and saving over time. The objective is to maximize the discounted utility of consumption over an infinite horizon, subject to the capital accumulation constraint and initial capital.

Lagrangian:

$$\mathcal{H}(t, K(t), C(t), \lambda(t)) = e^{-\rho t} U(C(t)) + \lambda(t)[F(K(t), L) - \delta K(t) - C(t)]$$

Intuition: The Lagrangian, \mathcal{H} , combines the objective function and the constraint, with $\lambda(t)$ acting as the shadow price or multiplier. It represents the trade-off between current consumption and future capital accumulation.

First-Order Conditions:

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial C(t)} = 0 & \Rightarrow U'(C(t)) = \lambda(t) \\ \frac{\partial \mathcal{H}}{\partial K(t)} + \dot{\lambda}(t) = 0 & \Rightarrow \dot{\lambda}(t) = -\lambda(t)[F'(K(t), L) - \delta] \\ \lim_{t \rightarrow \infty} \lambda(t)K(t) &= 0 \end{aligned}$$

Understanding the Conditions: 1. The first condition equates the marginal utility of consumption with the shadow price of capital, balancing immediate utility and the cost of reduced capital. 2. The second condition represents the evolution of the shadow price over time, considering the marginal product of capital. 3. The transversality condition ensures no unbounded accumulation of capital in the long run, reflecting the finite value of future resources.

Data analysis

Creating figures from data

Extracting data

Reformatting data

Classifying and scoring text

Extracting sentiment

Simulating human subjects

Data analysis: *Streamline data-related tasks*

- LLMs help format, extract, and classify data, and discern text sentiment
- Generate and run code for analysis, with results presented in various formats
- Produce figures or tables from datasets of your own, or data on websites
- Replicate economic models
- **Caveats:**
 - Complex analyses may lack accuracy
 - Extensive iteration recommended
 - Review outputs carefully for mistakes

15h00 – 15h45 **Uses of Generative AI for Economics Research**
Kevin Bryan, University of Toronto

16h15 – 17h00 **Unlocking Economic Data with LLM**
Sérgio Correia, Board of Governors of the Federal Reserve System

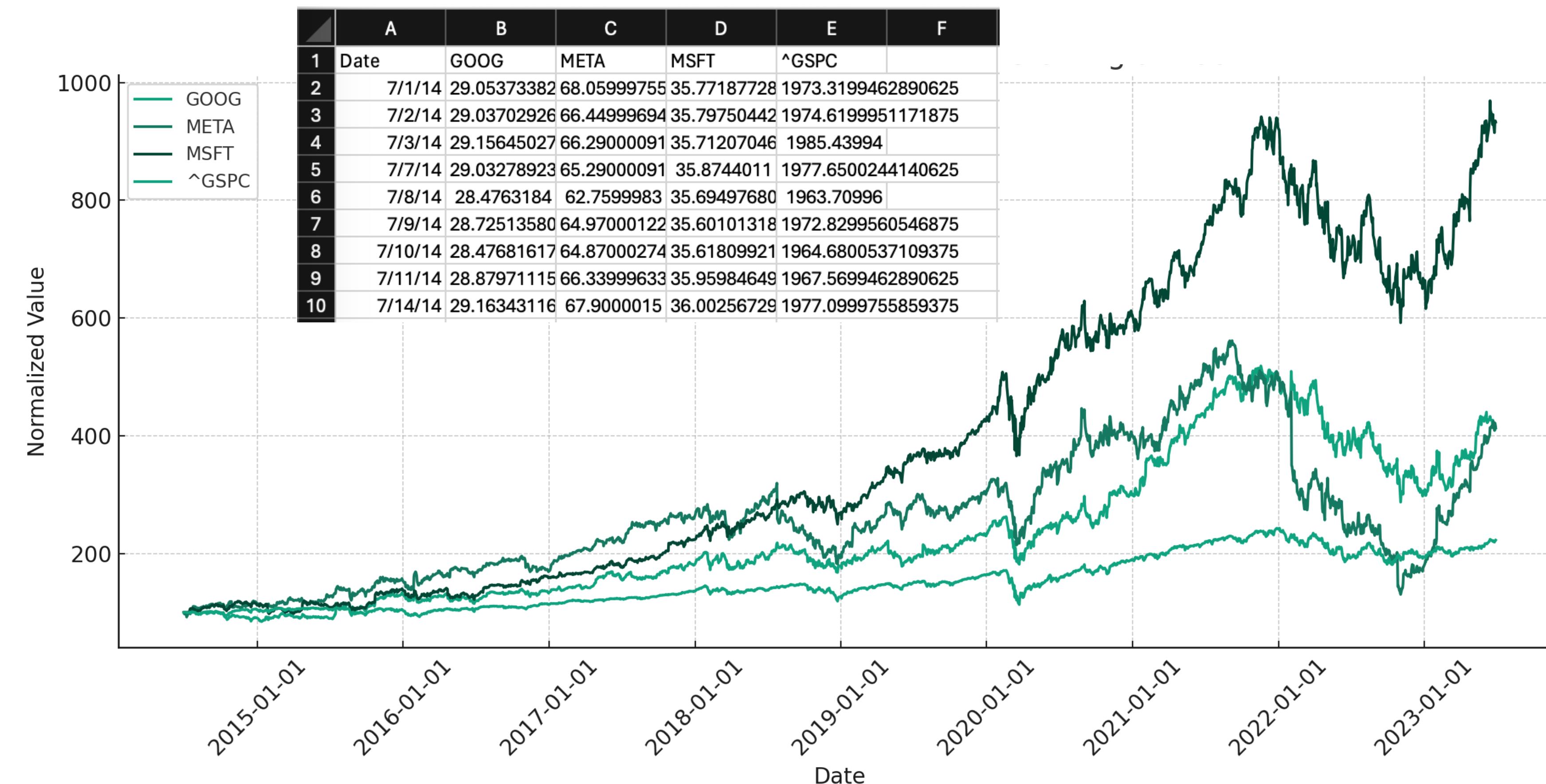
10h15 – 11h00 **How Large Language Models Support Statistical Analysis**
Peter Gruber, USI Università della Svizzera Italiana

PROMPT:

Please display a graph that shows the performance of these securities normalized to start at 100? (Attachment: stock_data.csv, which is itself generated via an LLM)

RESPONSE:

(GPT-4 via the ChatGPT Plus web interface, Advanced Data Analysis)



PROMPT:

Go online to extract details of all 'At a glance' indicators (both indicator and current estimate).

It is available at:

<https://fred.stlouisfed.org>

List in short bullet points.

AT A GLANCE	POPULAR SERIES	LATEST RELEASES	FEATURED
Consumer Price Index for All Urban Consumers: All Items +2.7 % Chg. from Yr. Ago on Nov 2024	Real Gross Domestic Product 2.8 % Chg. from Preceding Period on Q3 2024	U.S. / Euro Foreign Exchange Rate 1.0563 U.S. \$ to 1 Euro on 2024-12-06	Unemployment Rate 4.2 % on Nov 2024
Industrial Production Index -0.3 % Chg. on Oct 2024	10-Year Treasury Constant Maturity Rate 4.32 % on 2024-12-12	All Employees: Total Nonfarm Payrolls +227 Chg., Thous. of Persons on Nov 2024	Initial Claims 242,000 on 2024-12-07

RESPONSE:

(GPT-4o via the ChatGPT Plus web interface, incl. online search)

Here are the “At a Glance” economic indicators from the Federal Reserve Economic Data (FRED) website, along with their current estimates as of December 15, 2024:

- Consumer Price Index for All Urban Consumers: All Items: Increased by 2.7% from a year ago in November 2024.
- Real Gross Domestic Product: Grew by 2.8% in the third quarter of 2024 compared to the preceding period.
- Industrial Production Index: Decreased by 0.3% in October 2024.
- 10-Year Treasury Constant Maturity Rate: Stood at 4.32% on December 12, 2024.
- U.S. / Euro Foreign Exchange Rate: 1.0563 U.S. dollars per euro on December 6, 2024.
- Unemployment Rate: At 4.2% in November 2024.
- All Employees: Total Nonfarm Payrolls: Increased by 227,000 persons in November 2024.
- Initial Claims: 242,000 claims on December 7, 2024.

These indicators provide a snapshot of the current economic conditions in the United States.

PROMPT: Below are excerpts from the U.N. World Economic Situation Prospects reports from 2020 (top) and 2021 (bottom) related to international tourism in 2019 and 2020. Summarize the sentiment of both excerpts in one bullet point, and label them on a scale from 1 (very negative) to 5 (very positive), explaining your reasoning.

"Confidence in global tourism remains positive yet cautious for the remainder of 2019. Weakening economic indicators, trade tensions and Brexit-related uncertainties have started to take a toll on business and consumer confidence. The UNWTO Confidence Index points to more moderate growth in arrivals during the period September-December 2019, particularly in Europe and the Americas. ... UNWTO estimates 3 to 4 per cent growth in international arrivals globally for 2019, reflecting rising tourism demand overall, though at a slower pace. At the regional level, prospects are strongest for Asia and the Pacific, where arrivals are expected to have grown 5 to 6 per cent"

"The tourism industry has been the hardest-hit services sector, as COVID-19 brought international travel to a virtual standstill from March 2020 onward. Before the pandemic, travel services accounted for almost one third of developing countries' services exports. International tourist arrivals (overnight visitors) are estimated to have plunged by about 70 per cent globally in 2020 (UNWTO, 2020e), which would represent 1 billion fewer international arrivals than in 2019 and a potential loss of US\$ 1.1 trillion in international tourism receipts, the largest decline ever. The shock has put 100 million to 120 million direct tourism jobs at risk, with large spill-overs into other sectors."

Simulating human subjects

PROMPT: For each of the following registered voters, provide a one-word rating to indicate whether they are more likely to approve or disapprove of a policy proposal to loosen US immigration restrictions on highly-educated workers:

1. A 32-year old female lawyer in Houston, TX who is Hispanic
2. A 73-year old male retired doctor in Rome, GA who is Caucasian
3. A 25-year old software engineer in San Francisco, CA who is Asian and identifies as non-binary
4. A 45-year old female grocery check-out clerk in Brooklyn, NY who is second-generation Italian
5. A 55-year old male insurance agent in Lake City, FL who is Caucasian

From: Korinek, A. (2023): “Generative AI for Economic Research: Use Cases and Implications for Economists”. Journal of Economic Literature 61 (4): 1281–1317.

"AI for Economists" resources at:

futurai.fi

UNU-WIDER's webpage at:

wider.unu.edu

Follow us on LinkedIn, Facebook, and Twitter @UNUWIDER