

Generativ AI

Joacim Tåg

IFN and Hanken (www.jtag.se)

September 2023

Agenda

1. Vad är generativ AI?
2. Hur kommer generativ AI att påverka ekonomin?
3. Att forska med ChatGPT4

Vad är generativ AI

Vad är generativ AI?

- Programstyren process som **producerar, manipulerar och syntetiserar** data
- Generativ AI *framställer nytt material*, särskiljande AI ("discriminative AI") *kategoriserar*
- Generativ AI har funnits i decennier (dialogbotten Eliza, MIT 1966)
- Moderna versioner som **ChatGPT** och **Dall-E** är *förutsegelsemaskiner*

Vad är ChatGPT

- ChatGPT (Generative Pre-trained Transformer) är en språkmodell (**LLM**) utvecklad av OpenAI för att förstå och generera text
- Den *förutsäger* efterföljande ord i ett givet sammanhang för att skapa mänsklig text
- ChatGPT använder Transformer-arkitekturen, en djupinlärningsmetod som gör det möjligt för modellen att förstå mönster, sammanhang och språkstruktur från stora mängder text
- ChatGPT finjusteras med Reinforcement Learning från mänsklig feedback (RLHF)

Översikt av LLMs

MODEL	RUNS CODE	SEE IMAGES	READS FILES	INTERNET CONNECTION	PERSONALITY	WHAT IS IT	WHEN TO USE IT
ChatGPT/ GPT-3.5				No	Neutral, bland	This is the free version of ChatGPT that came out in November. It is very fast and pretty solid at writing and coding tasks.	It is fast and cheap and capable, but other models are now better. Also, It is not connected to the internet, so don't use it like a search engine.
ChatGPT/ GPT-4		Soon		No	Helpful, a little preachy	Currently available to paying customers. The most powerful LLM for most purposes. Does a lot.	Better at everything (writing, coding, summarizing) than GPT-3.5 Still not connected to the internet.
ChatGPT/ Code Interpreter/ Plugins	Yes		Yes	Limited	Helpful sometimes witty	Adds new capabilities to ChatGPT. Plugins and web browsing still have rough edges. Code Interpreter lets the AI run Python code & work with files.	Code Interpreter is great for working with data of all sorts, and is the most generally impressive AI implementation so far. Plugins are generally not that useful, yet. The web browsing is being updated.
Bing (Uses GPT-4 in creative & precise modes)		Yes	Yes	Yes	Friendly but also somewhat insane	Connected to the internet. It is GPT-4 with a grab-bag of powerful features. The weirdest model. Creative mode lets you use GPT-4 and is free.	Does a remarkably wide range of stuff, from working as a browser companion to creating art. Is connected to the internet. Has a strong, weird personality.
Bard	Yes	Yes		Yes	None	Bard is a name for a variety of models used by Google. Current offerings are weak, but improving over time	This should be your last choice of AIs for now, but recent updates show promise. High levels of hallucination.
Claude 2			Yes	No	Pleasant when not scolding	Almost as powerful as GPT-4, but built to be "safer" and more pleasant to use. Has 100k context window - enough "memory" to hold a book.	Terrific on working with large documents, due to its extended context window and ability to upload content. It is very new so it is unclear how it stacks up.

Exempel på vad LLMs är bra på

- **Skrivassistent:** Förbättrar skrift, genererar innehåll, och översätter
- **Kodning:** Assisterar i kodutveckling och problemlösning
- **Pedagogik:** Svarar på frågor och skapar övningsmaterial
- **Dataanalys:** Genererar insikter från stora datamängder
- **Kundtjänst och teknisk support:** Hanterar exempelvis automatiserade kundinteraktioner och kan ge teknisk support

Risker med generativ AI

- Hallucinationer och vinkling
 - Uppkommer på grund av att AI "gissar" hellre än att säga att den inte vet
 - Kan justeras i algoritmen, men är en balans mellan korrekthet och kreativitet
- Säkerhetsrisker
 - Terrorister och Nordkorea
 - "Fake news" och annan effektiv propaganda är lätt att generera
- Kan leda till att AGI utvecklas ("*Artificial General Intelligence*")
 - Butlin et al (2023): "*our analysis suggests that no current AI systems are conscious, but also suggests that there are no obvious technical barriers to building AI systems which satisfy these indicators*"
 - Existentiell risk med självförbättrande AGI (vi är inte längre smartast på planeten, 10% tror vi dör ut)
 - Svårt att justera belöningsfunktionen för den är så komplex ("AI Alignment")
 - Tre scenarion utanför "*Terminator IRL*"
 - "Maktkoncentration och snabb automatisering"
 - "Hitler, Stalin och Gengis Khan"
 - "Gem-maskinen"

Påverkan på ekonomin

Påverkan på ekonomin

- Effekter på arbetsmarknaden (Eloundou et al 2023)
 - **80%** av jobben kommer att se en **tiodel** av sina arbetsuppgifter påverkade
 - **20%** av jobben kommer att se **hälften** av sina arbetsuppgifter påverkade
 - **15-56%** av alla arbetsuppgifter kan utföras snabbare till samma kvalitetsnivå
 - Påverkar i större utsträckning **välbetalda** jobb som kräver **hög** utbildning
 - **Minskar** 90-10 löneklyftan (Webb 2020)
- Speciellt exponerade sektorer
 1. Tech sektorn (kodning)
 2. Mediasektorn (innehållskapande)
 3. Rättssektorn (advokater som skall behandla mycket lagtext)
 4. Utbildningssektor (lärare och utbildning)
 5. Grafisk design och marknadsföring (kan genereras av AI)
 6. Kundservice och teknisk support (kan ersättas nästan helt av chatbottar)

Påverkan på ekonomin

- Effekter på produktivitet
 - Noy and Zhang (2023, *Science*): RTC, "mid-level professional writing tasks", "college students"
 - "Time taken **decreases by 0.8 SDs** and **output quality rises by 0.4 SDs**"
 - "**Inequality between workers decreases**, as ChatGPT compresses the productivity distribution by benefiting low-ability workers more"
 - "ChatGPT mostly substitutes for worker effort rather than complementing worker skills"
 - "Restructures tasks towards idea-generation and editing and away from rough-drafting"
 - "Exposure to ChatGPT increases **job satisfaction** and self-efficacy"
 - Brynjolfsson et al (2023, WP): "staggered adoption", "customer service agents"
 - "Access to the tool **increases productivity**, as measured by issues resolved per hour, by **14 percent** on average, with the **greatest impact on novice and low- skilled workers**, and **minimal impact on experienced and highly skilled workers**"
 - "The AI model **disseminates the potentially tacit knowledge of more able workers** and helps newer workers move down the experience curve"

Att forska med ChatGPT4

Vad kan ChatGPT4 används till?

- **Generera nya forsknings idéer**
 - Brainstorming
 - Evaluera idéer
 - Ta fram motargument och skriva korta ”referee rapporter”
- **Hjälpa till att skriva text**
 - Skriva texten och sammanfatta
 - Editera och rätta text
 - Evaluera stilen och tydligheten i texten
 - Skapa minnesvärda titlar
 - Generera tweets

Vad kan ChatGPT4 används till?

- **Sammanfatta litteratur**

- Sammanfatta text och skriva översikter (jenni.ai)
- Hitta relevant forskning
- Formatera referenser
- Översätta text från andra språk
- Förklara text

- **Kodning och analys av data**

- Skriva kod
- Förklara kod
- Översätta kod från ett språk till ett annat
- Hitta fel i koden

Hur interagera med ChatGPT4

- **”Promptcrafting”**

- Förklara kontexten, uppgiften, stilen och längden (var motsatsen till vag)
- Var så specifik och detaljerad som möjligt
- Ge exempel

- **Tips**

- Är din personliga assistent (*”I want you to act as X”*)
- Viktigast är bara att komma igång, flera interaktioner ger ofta samma resultat
- Be om flera alternativ
- Konstant förbättring i att förstå enkla ”prompts”

Exempel

- Skapa första utkast på text

1. Sätt en timer på i 30 minuter och skriv text utan att sluta (alternativt diktera)
2. Kopiera in i ChatGPT4 med följande prompt:

"Please remove redundant words from the following passage and make it coherent and cohesive."

- Förbättra argumentation

- *"Please critique the following passage and let me know if my argument is clear enough or not."*

Exempel

- Förstå koncept
 - *"Please explain the following passage in simple words. I am having difficulty understanding X."*
- Strukturera en artikel
 - *" Prepare an outline for a journal article/dissertation on X. You can specify the number of chapters."*
- Förslag på titel
 - *"Here is my abstract, please provide ten remarkable and memorable title suggestions"*

Exempel

- Skriva e-post

”Context: Arranging an economic conference on labor and finance in Stockholm

Task: Please draft an invitation email asking if prominent researcher X would like to attend and present a paper.”

- Strukturera en artikel

- *” Prepare an outline for a journal article/dissertation on X. You can specify the number of chapters.”*

- Förslag på titel

- *”Here is my abstract, please provide ten remarkable and memorable title suggestions”*

Exempel

- Översätt

 - *”Translate this text from english to Swedish:”*

- Policyslutsatser

 - *”Provide ten policy conclusions from this introduction to my academic article:”*

- Fixa referenser

 - *”Here is my reference list, translate into the Chicago Manual of Style format”*

- **ChatGPT4**
 - Custom instructions
 - Advanced data analysis
- **Övriga att hålla koll på (överkurs, the "wait equation")**
 - bing (internet access)
 - claude.ai (men bara tillgänglig i US/UK ännu)
 - scite.ai (svar från "peer-reviewed" forskning)
 - jenni.ai och litmaps.com (översikter)
 - paperpal.com (grammarly+)
 - stablediffusionai.ai / midjourney.com (bilder)
 - aragon.ai (professionella pressbilder)

Custom instructions

Custom instructions ⓘ

What would you like ChatGPT to know about you to provide better responses?

I am a professor in economics doing empirical research in labor and corporate finance. I am based in Stockholm. I code using Stata.

131/1500

Hide tips ⓘ

How would you like ChatGPT to respond?

When replying, I want you to use the guidelines suggested by Deirdre McCloskey in her book "Economic Writing". But don't say that you are.

140/1500

Enable for new chats

Cancel

Save

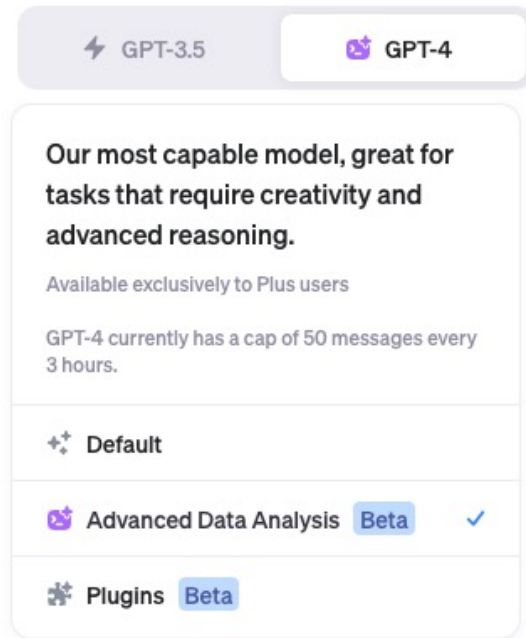
Thought starters

- Where are you based?
- What do you do for work?
- What are your hobbies and interests?
- What subjects can you talk about for hours?
- What are some goals you have?

Thought starters

- How formal or casual should ChatGPT be?
- How long or short should responses generally be?
- How do you want to be addressed?
- Should ChatGPT have opinions on topics or remain neutral?

Advanced data analysis / Code interpreter



The screenshot shows the OpenAI chat interface with two model selection buttons at the top: "GPT-3.5" (disabled) and "GPT-4" (active). Below the buttons, a text box describes GPT-4 as the most capable model, suitable for creative and advanced reasoning tasks, and notes it is available exclusively to Plus users with a 50-message cap every 3 hours. A settings menu is open, showing "Default" as the current mode, "Advanced Data Analysis" (Beta) as the selected mode with a checkmark, and "Plugins" (Beta) as an available option.

⚡ GPT-3.5 🧠 GPT-4

Our most capable model, great for tasks that require creativity and advanced reasoning.

Available exclusively to Plus users

GPT-4 currently has a cap of 50 messages every 3 hours.

⚙️ Default

🧠 Advanced Data Analysis **Beta** ✓

⚙️ Plugins **Beta**




Increases in house prices can affect entrepreneurship among young people in several ways:






1. **Barrier to Entry:** High house prices may limit the ability to access capital, as homes are often used as collateral for loans. This constraint may inhibit young entrepreneurs from starting new businesses.
2. **Opportunity Cost:** Rising house prices may push young individuals to prioritize homeownership over entrepreneurial pursuits, as the fear of being priced out of the housing market increases.
3. **Wealth Effect:** For those who already own property, increased house prices can create a sense of additional wealth, possibly enhancing risk-taking and entrepreneurial activities.
4. **Regional Disparities:** High house prices in urban centers might push young entrepreneurs to less expensive areas, affecting the geographical distribution of startups.
5. **Rent Pressure:** In areas with soaring house prices, rental costs may also rise. This increased expenditure may reduce available capital for entrepreneurial activities, particularly for those who are renting living or working spaces.

In summary, while rising house prices might encourage some young property owners to venture into entrepreneurship, they can generally pose barriers to entry, alter priorities, and change geographical business dynamics for young potential entrepreneurs.

Yes, I know about these methods. Can you suggest concrete examples of policy changes or shocks from Sweden that could be used? 



Certainly! Here are some concrete examples of policy changes, shocks, or regional variations that could be used as instrumental variables or focal points for causal analysis in Sweden:   

1. **Municipal Zoning Laws:** Changes in zoning regulations within specific municipalities might create exogenous variation in housing prices, enabling you to examine their impact on entrepreneurship.
2. **Property Tax Reforms:** Any substantial reforms or adjustments to property taxation could serve as an instrumental variable, affecting housing prices but arguably unrelated to entrepreneurship directly.
3. **Rent Control Policies:** Sweden's history of rent control might provide opportunities to explore how adjustments in such policies have influenced house prices in different regions.
4. **Infrastructure Investments:** Major investments in public transportation or other infrastructure might affect property values in specific areas. For example, the expansion of the Stockholm metro system.
5. **Urban Renewal Projects:** Specific urban renewal or redevelopment projects can create localized changes in property values, such as the redevelopment of certain districts in Malmö or Gothenburg.
6. **Environmental Regulations:** Introduction or changes in environmental regulations, such as those affecting building codes, may have localized effects on housing prices.
7. **Housing Subsidy Programs:** Programs aimed at assisting homeownership among certain demographics, such as first-time buyers, may also lead to localized changes in housing prices.
8. **Financial Crisis Impact:** The effects of the global financial crisis on Sweden's housing market could provide an interesting shock to study, especially as they varied regionally.
9. **Interest Rate Changes by Riksbank:** The central bank's decisions on interest rates could provide a broader macroeconomic tool to explore how they have indirectly impacted house prices.
10. **Migration Patterns:** Sweden's varying migration patterns (both internal and international) might offer opportunities to study how population shifts affect housing demand and prices in different regions.

X1

 BO0101A5_20230823-062511.xlsx
Spreadsheet



It looks like you've uploaded a file named `BO0101A5_20230823-062511.xlsx`. How can I assist you with it?



X1

Its a file of newly produced houses across swedish municipalities across years. Do you understand the structure of the data?



I will need to inspect the file to understand its structure. Let me take a quick look at the contents.



Finished working Show work ▾

The data appears to represent the number of newly built apartments in different Swedish municipalities across various years. The structure includes:

- Municipality names in the first column, starting from the third row.
- The type of housing (e.g., "flerbostadshus") in the second column.
- The number of newly built apartments for each year from 1975 to 2022 in the subsequent columns.

Please confirm if this understanding is correct, and let me know how I can assist you with this data.

x1

Let's run a diff-in-diff analysis looking at how the removal of property taxes in 2008 differentially affected new apartments in Stockholm relative to Malmö, Göteborg and Uppsala. Give me the regression output and plot the year times treated coefficients with confidence intervals over time in a graph



Certainly! We'll conduct a difference-in-differences (DiD) analysis to estimate the effect of the removal of property taxes in 2008 on new apartments in Stockholm, relative to the other cities: Malmö, Göteborg, and Uppsala.



For the DiD analysis, we'll consider the following variables:

- **Outcome Variable (Y):** Number of new apartments
- **Treatment Variable (D):** Binary variable indicating whether the municipality is Stockholm (1) or one of the control cities (0)
- **Post-Treatment Variable (P):** Binary variable indicating whether the year is after the removal of property taxes in 2008 (1) or before (0)
- **Interaction Term (DP):** Interaction between the treatment and post-treatment variables

The regression model we'll estimate is:

$$Y_{it} = \alpha + \beta_1 D_i + \beta_2 P_t + \beta_3 (D_i \times P_t) + \gamma X_{it} + \epsilon_{it}$$

where Y_{it} is the outcome for municipality i in year t , D_i is the treatment variable, P_t is the post-treatment variable, and X_{it} are other controls (e.g., year fixed effects).

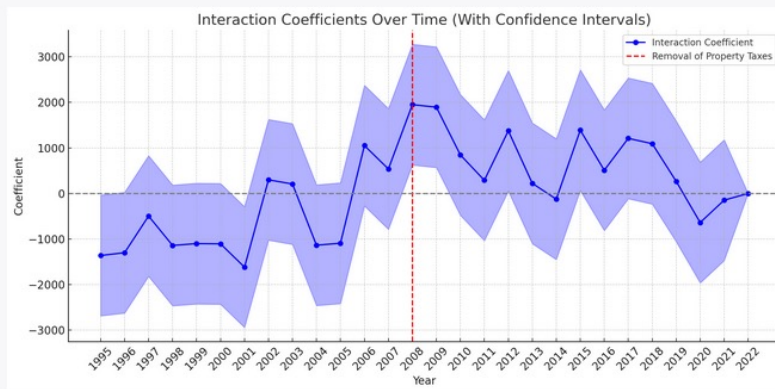
The coefficient β_3 on the interaction term will give us the differential effect of the treatment on the treated group (Stockholm) relative to the control group (Malmö, Göteborg, Uppsala) after the removal of property taxes.



We'll proceed with the difference-in-differences (diff-in-diff) regression by explaining the "New Apartments" variable with year dummies, a treated dummy (for Stockholm), and their interactions. Then we'll plot the interaction coefficients.



Finished working Show work ▾






Here's the plot showing the interaction coefficients over time from the difference-in-differences (diff-in-diff) regression. The regression explains new apartments with year dummies, a treatment dummy (for Stockholm), and their interactions.

- The blue line represents the interaction coefficients, capturing the differential effect on new apartments in Stockholm compared to Gothenburg, Malmö, and Uppsala for each year.
- The shaded area represents the confidence intervals for these coefficients.
- The red dashed line marks the year 2008 when the removal of property taxes occurred.
- The grey dashed line at zero serves as a reference line.

This plot visualizes how the policy change in 2008 affected the number of new apartments in Stockholm relative to the other cities.

Please let me know if there's anything else you'd like to know or do!



Certainly! Below is the Stata code to clean the dataset and produce the final figure. Note that this code assumes that you have the raw dataset in Excel format and that you've imported it into Stata.   

```
stata Copy code  
  
* Import the dataset (replace "your_dataset.xlsx" with your actual file name)  
import excel "your_dataset.xlsx", firstrow  
  
* Keep only the relevant columns: Year, Municipality, New_Apartments  
keep Year Municipality New_Apartments  
  
* Filter for years starting from 1995  
gen year_filter = (Year >= 1995)  
keep if year_filter == 1  
drop year_filter  
  
* Generate the treatment variable for Stockholm, Gothenburg, Malmo, and Uppsala  
gen Treatment = .  
replace Treatment = 1 if Municipality == "0180 Stockholm"  
replace Treatment = 0 if Municipality == "1480 Göteborg" | Municipality == "1480 Malmö"  
  
* Convert New_Apartments and Year to integers (if they are not already)  
destring New_Apartments, replace  
destring Year, replace  
  
* Run the diff-in-diff regression  
gen year_dummies = "Year_" + string(Year)  
gen interaction_term = year_dummies + "*" + string(Treatment)  
  
regress New_Apartments i.year_dummies##i.Treatment  
  
* Generate the plot using coefplot (assuming you have installed the coefplot package)  
coefplot, vertical drop(_cons) title(Interaction Coefficients Over Time (With Treatment))
```

- **Tips till forskare**

- Ändra ditt arbetsflöde för att använda **ChatGPT4** (sätt som hemsida i webbläsaren)
- Är en *investering*, du blir mer produktiv på sikt
- Genom experimentering lär du dig vad som fungerar (bli inte sur på dåliga svar)
- Använd *Custom Instructions* och lär dig *Code Interpreter*
- Prenumenera på Ethan Mollicks nyhetsbrev: www.oneusefulthing.org

- **Tips till ledningen**

- Stå för alla kostnader med experimentering
- Se till att forskningsassistenterna är experter på ChatGPT4
- Svenska sammanfattningar och inlägg till sociala media borde genereras centralt
- Utbilda minst 1ggr/år