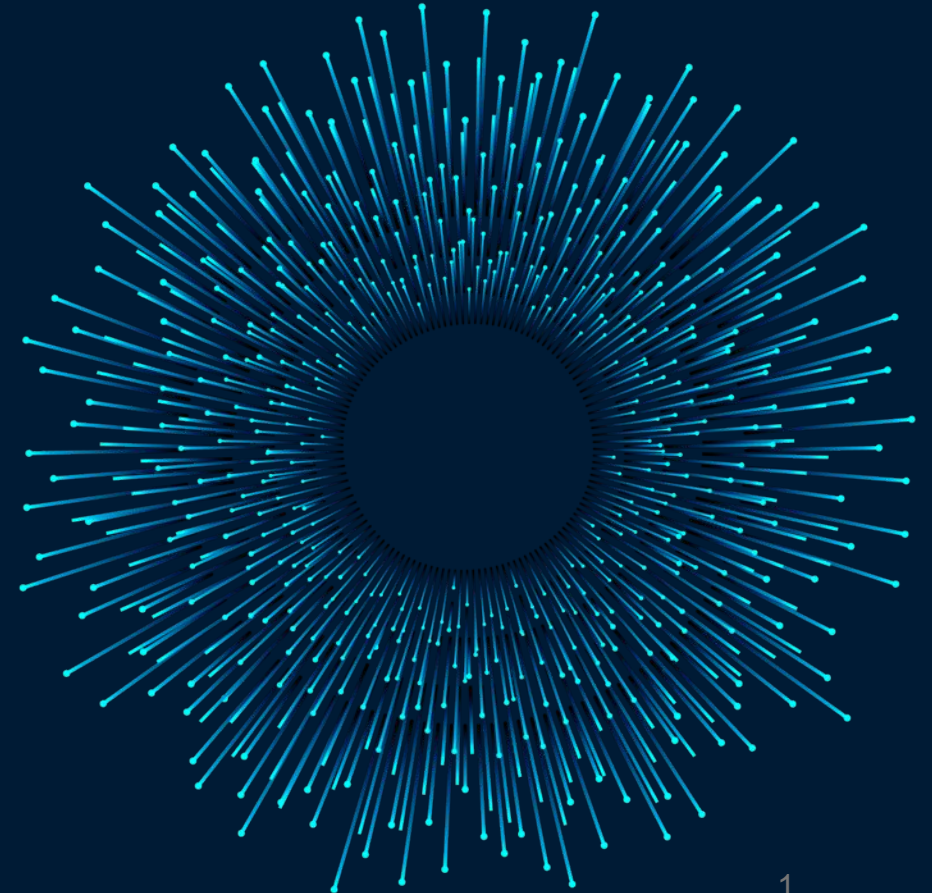


# GDP-B: A New Measure of Economic Welfare

**Erik Brynjolfsson**

ESCOE Annual Conference on Economic Measurement

May 22, 2025



# Outline

1. GDP vs well-being
2. The Digital Welfare of Nations
3. A Broader Basket of Goods: Digital vs Conventional
4. LLMs Agents for economic research and LLM Time Machines
5. Externalities, Limitations, and Extensions
6. The Future of Economic Measurement — A Call to Action

# Team behind Stanford GDP-B Project

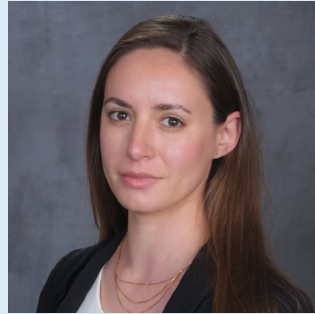
## Core research team



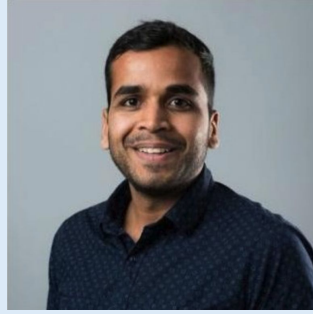
Erik  
Brynjolfsson



David  
Nguyen



Sophia  
Kazinnik



Avi  
Collis

## Key collaborators



Kevin  
Fox



Erwin  
Diewert



Felix  
Eggers

## GDP-B Advisory Committee



Diane Coyle



Jason Liu



Paul Schreyer



Leonard Nakamura



Rachel Soloveichik

# *1. GDP v Welfare*

# How Are We Doing?

The [United States economy](#) grew faster than any other large advanced economy last year — by a wide margin — and is on track to do so again in 2024.

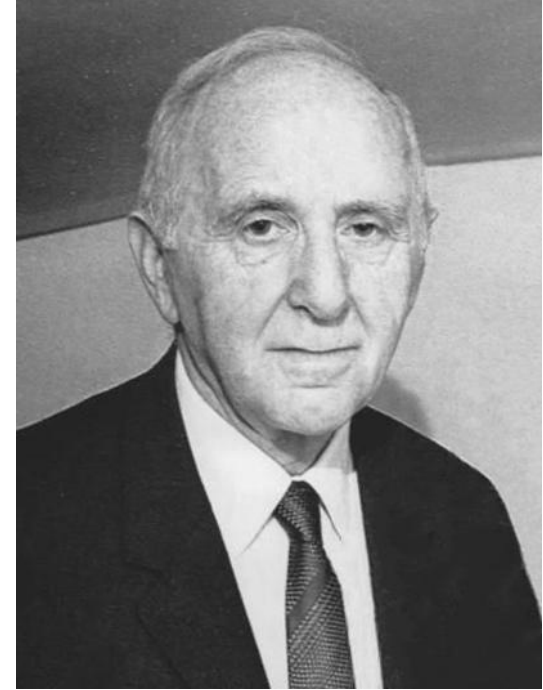
India Q3 GDP growth surges to 8.4%, exceeding expectations

## ***Are We in a Productivity Boom? For Clues, Look to 1994.***

Thirty years ago, the U.S. entered an era of productivity gains that enabled healthy growth. Experts are asking if it could happen again.

“...a measure for standard of living: average real gross domestic product (GDP) per capita” – Boston Fed

“Productivity is the most important determinant of the standard of living” – Forbes



**“The welfare of a nation can scarcely be inferred from a measurement of national income as defined [by GDP.]”**  
**- *Simon Kuznets, 1934***

1. GDP is a measure of *production*, not well-being
2. Productivity is simply GDP/hours worked

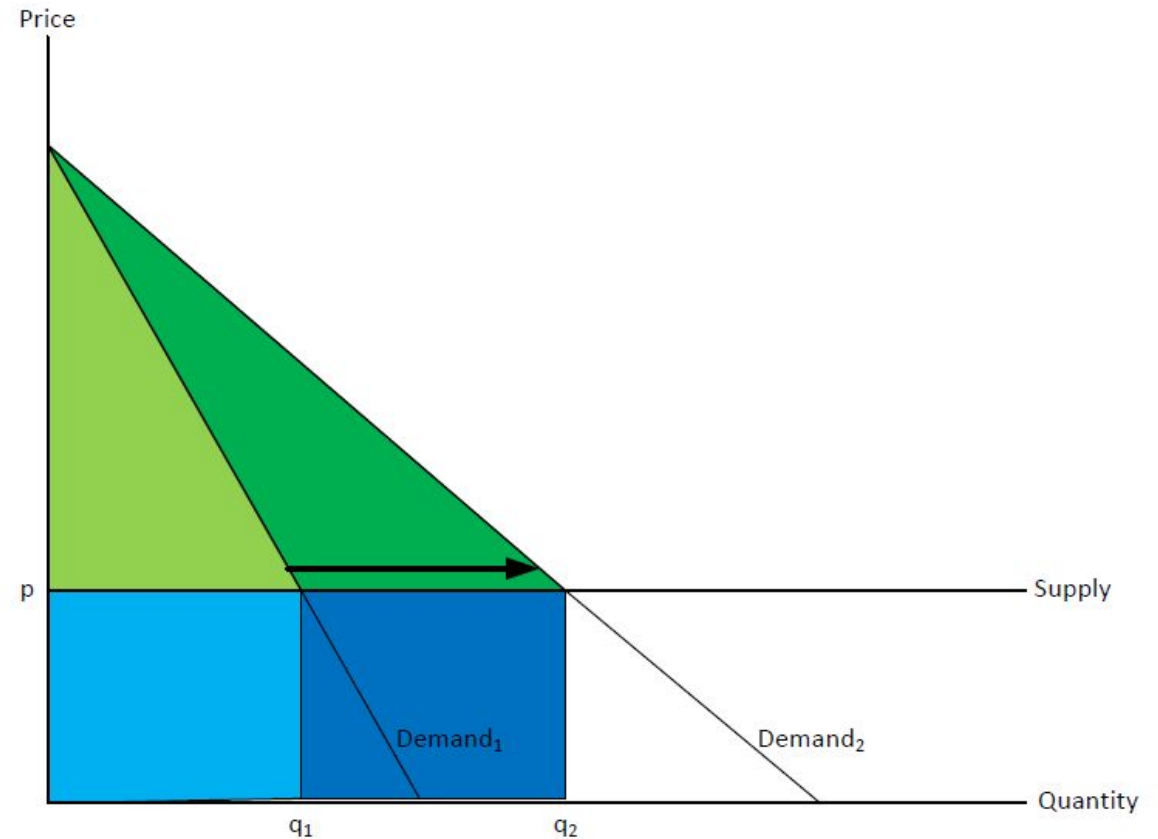
# **GDP vs. Consumer Welfare**

# $\Delta$ Production vs. $\Delta$ Consumer Surplus

## Case 1: Classic Goods

E.g. Automobiles, haircuts, food

GDP  $\uparrow$ , Consumer Surplus  $\uparrow$

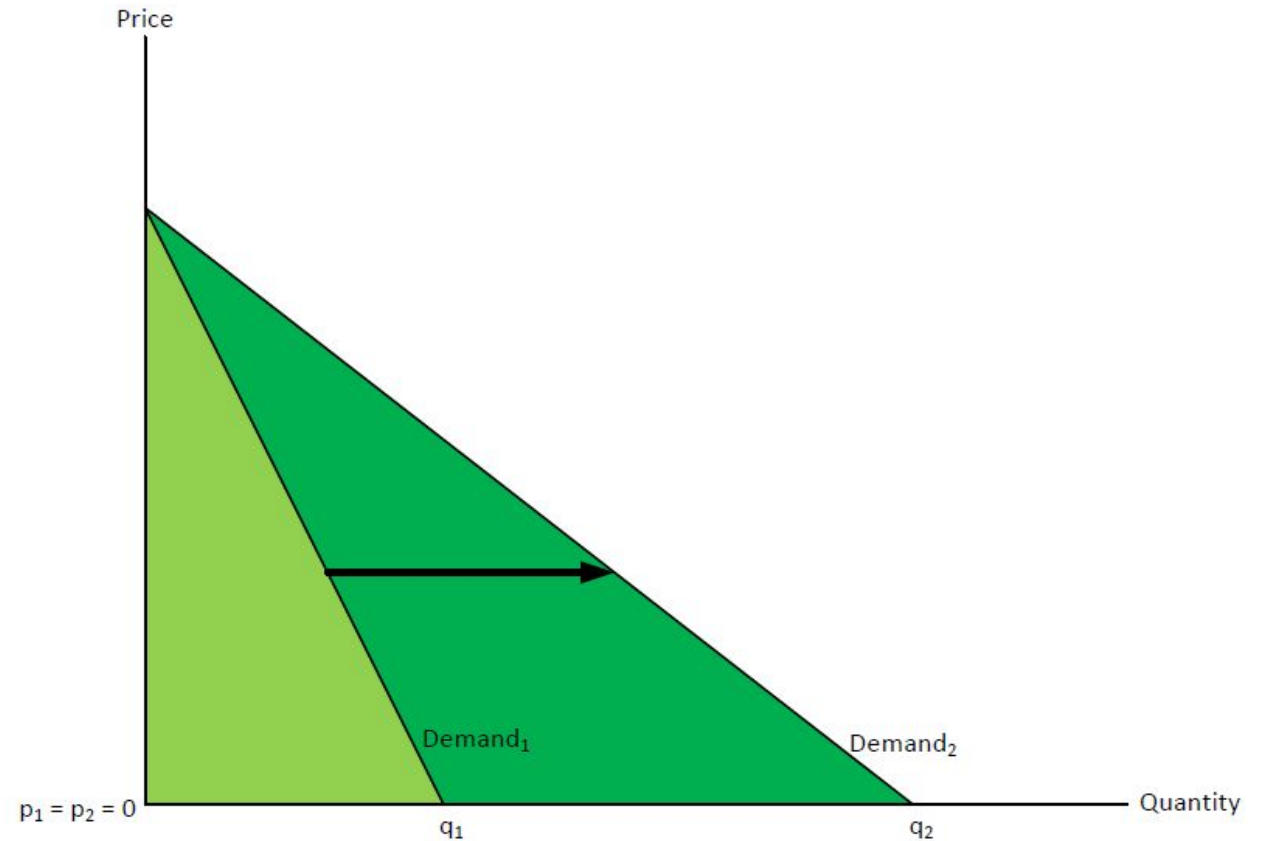


# $\Delta$ Production vs. $\Delta$ Consumer Surplus

## Case 2: Digital Goods

E.g. Increased use of free maps on smart phones or more digital photos;  
Special case: Free digital apps that never existed before

GDP no change,  
Consumer Surplus  $\uparrow$

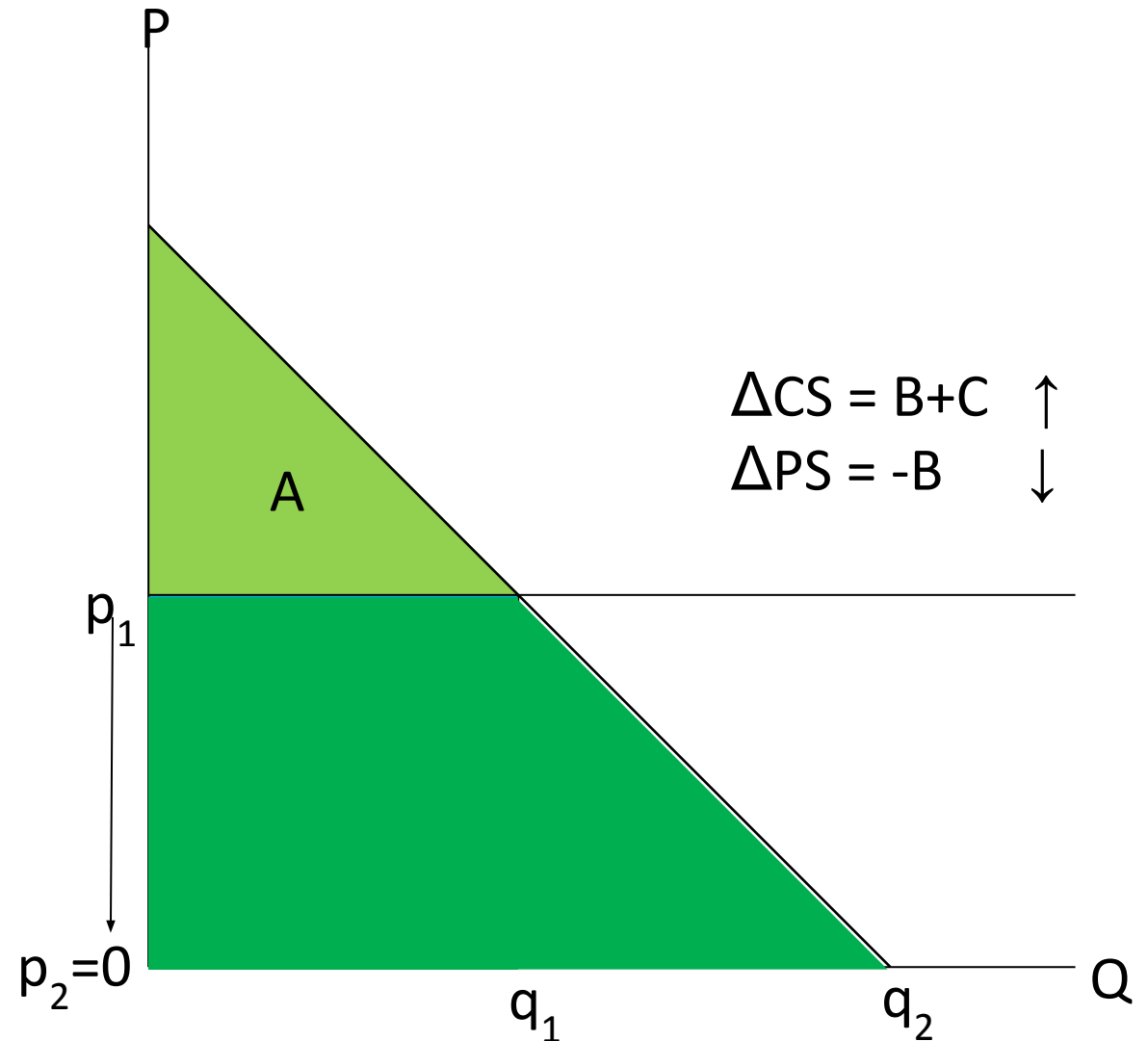


# $\Delta$ Production vs. $\Delta$ Consumer Surplus

## Case 3: Transition Goods

E.g. Encyclopedia  
(Wikipedia vs. Britannica)  
Chemical photography to digital  
photography

GDP  $\downarrow$ , Consumer Surplus  $\uparrow$

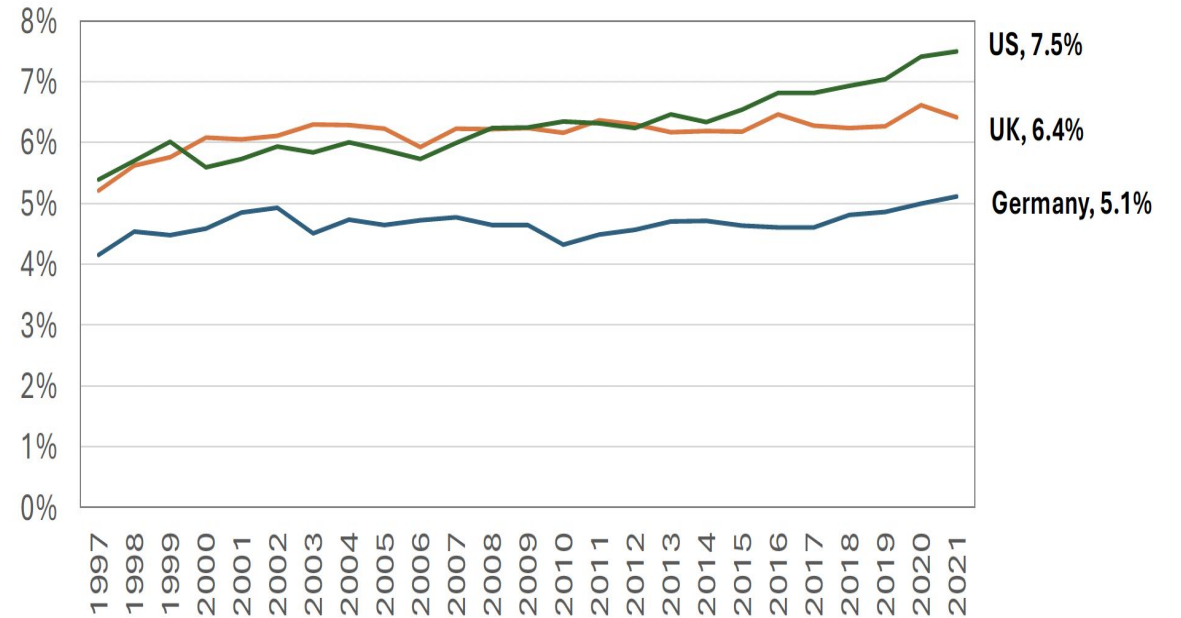


# Free Goods: Many Digital Goods and Services

“There’s a lack of appreciation for what’s happening in Silicon Valley, because we don’t have a good way to measure it.”

- Hal Varian

An explosion of free digital goods



# New Goods: Smartphones and Cameras

1. Photos taken worldwide
  - 2000: 80 billion photos
  - 2021: 1.6 trillion photos
  - Price per photo has gone from 50 cents to 0 cents.
2. Increase doesn't show up in GDP measures since...
  - Price index for photography includes price of film, developing, and cameras, all of which are vanishing
  - Photos are mostly shared, not sold (non-monetary transaction)
3. GDP went down further when cameras were absorbed into smartphones
  - 121 million digital cameras sold in 2010, but just 8 million in 2023.

# Smartphones aren't just absorbing cameras

Smartphones substituted

- Camera
- Alarm Clock
- Music Player
- Calculator
- Computer
- Land Line
- Game Machine
- Movie Player
- Recording Device
- Video Camera

Plus:

- Data plan
- GPS Map and directions
- Web Browser
- E-book reader
- Fitness monitor
- Instant messaging



**Radio Shack**  
AMERICA'S TECHNOLOGY STORE

**PRESIDENTS' BIRTHDAY SALE!**

**0% INTEREST!**  
NO PAYMENTS UNTIL MAY!  
NO DOWN PAYMENT!

**HURRY! OFFER ENDS TUESDAY FEBRUARY 19**

**COME IN AND TAKE ADVANTAGE OF THESE OTHER FANTASTIC VALUES!**

**INTRODUCTORY SPECIAL!**  
TANDY 1000 TL/3 Complete System  
Save \$670  
**\$1599**  
Low As \$199 Per Month - Reg. \$2269.95  
• 286-Based PC Compatible  
• Color Monitor  
• 20MB SmartDrive Hard Drive  
• Easy-to-Use 10-in-1 DeskMate Software  
• BONUS PACKAGE  
• Lotus Spreadsheets For DeskMate  
• DeskMate Q&A Write  
• Quicken  
• 2-Button Mouse

**VHS Camcorder**  
Save \$100  
**\$799**  
Low As \$199 Per Month - Reg. \$899.00  
Realistic Model 182 includes video light for indoor shooting, 2-hr battery, with accessories. #14-823

**3-Way Speaker With Massive 15" Woofer**  
Save \$110  
**14995**  
Low As \$19 Per Month - Reg. \$259.95  
Optima Match "Bee" 100-watt horned out bass you can feel. 4-1/2" tweeter, 5" midrange. #14-823

**All-Weather Stereo**  
Cut 34%  
**1188**  
Realistic STEREO-MATE™ AM/FM personal receiver shrugs off heat, water, #12-142. #1188.00. See Us In-Store.

**AM/FM Clock Radio**  
30% Off  
**1388**  
Chromatonic™-261 clock radio's compact size cuts nighttime clutter. #12-142. See Us In-Store.

**In-Ear Stereo Phones**  
HALF PRICE!  
**788**  
Realistic™ in-ear phones weigh just 0.6 ounce! With carry pouch. #12-817. See Us In-Store.

**Mobile Cellular Telephone**  
Save \$100  
**\$199**  
Low As \$19 Per Month - Reg. \$299.00  
#17-1076

**Deluxe Portable CD Player**  
Save \$40  
**15995**  
Low As \$19 Per Month - Reg. \$199.95  
Realistic CD-3250 has 10-selection memory. Headphones extra. #41-071.

**Tiny Dual-Superhet Radar Detector**  
Save \$60  
**7995**  
Reg. \$85.95  
Road Patrol XX™ detector lets you drive with confidence. Separates X and K-band tones. #12-161.

**Compact 10-Channel Desktop Scanner**  
Save \$30  
**9995**  
Low As \$10.95  
Realistic PRO-17 lets you catch the news as it happens! Hear police, fire, rail, military. 105 more. #12-161.

**Mobile CB With Channel Controls on Mike**  
HALF PRICE!  
**4995**  
Reg. \$99.95  
Realistic TRC-430 lets you get highway info or help—you'll never have to drive alone. #14-104.

**Our Easiest-to-Use Phone Answerer**  
Cut 17%  
**4995**  
Reg. \$59.95  
DUALPHONE™ TAD-241 answerer is ready to use—just plug it in. Has built-in announcement. #43-263.

**Handheld Voice-Actuated Cassette Tape Recorder**  
40% Off  
**2995**  
Reg. \$49.95  
Realistic CTR-85 makes an excellent "notetaker" for students, secretaries or executives. #14-104.

**20-Memory Speed-Dial Phone**  
Cut 33%  
**2995**  
Reg. \$44.95  
Radio Shack CR-292 Speed Dialer is ideal for home or office. Touch-tone. #12-142. #2995.00. See Us In-Store.

Check Your Phone Book for the Radio Shack Store or Dealer Nearest You

Most Major Credit Cards Welcome



# We Need A New Approach: *GDP-B*

**GDP-B can be assessed for all types of goods**

**It is especially important for two types of goods:**

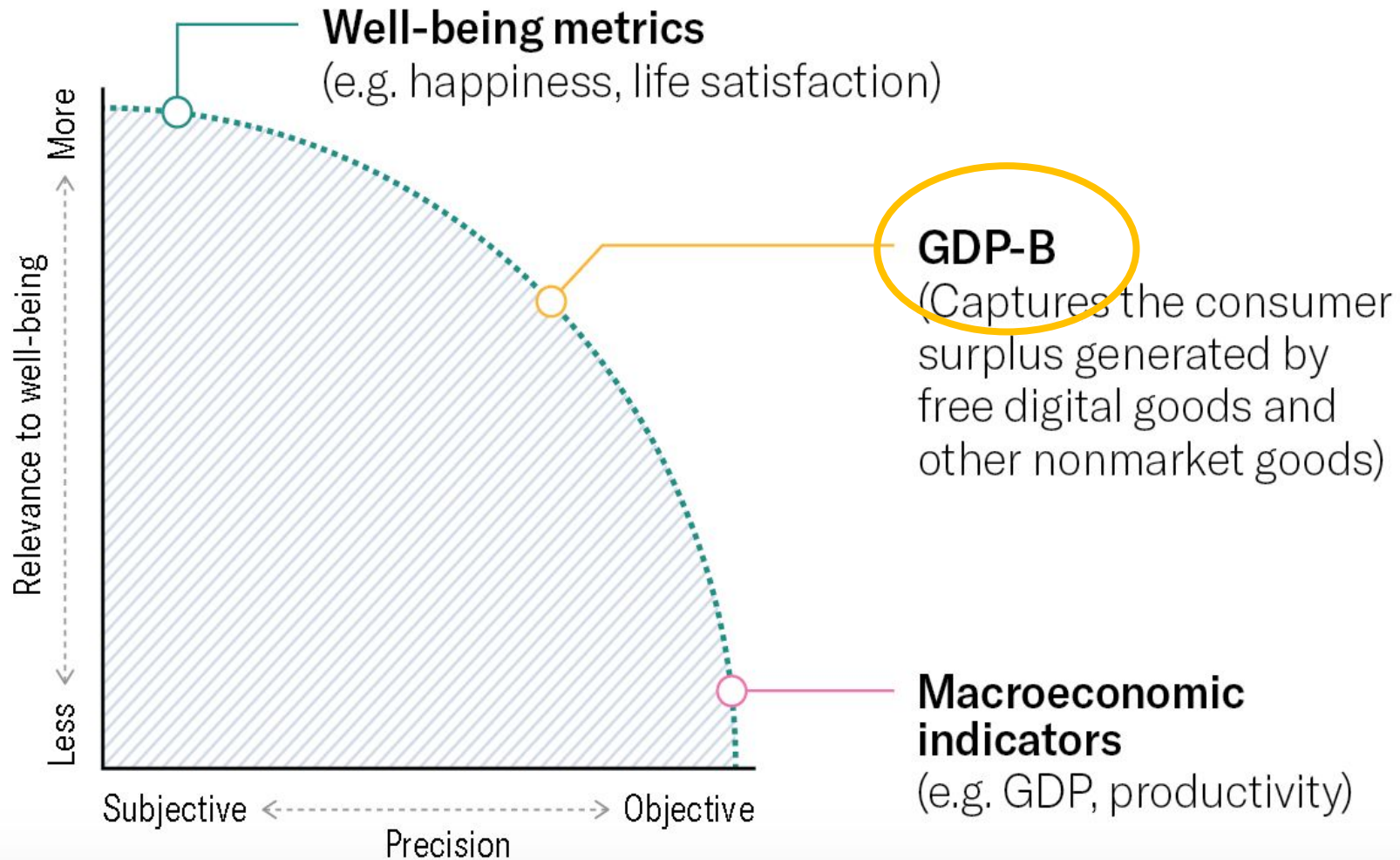
**1. Free goods**

- E.g. Facebook, Wikipedia

**2. New goods**

- E.g. Smartphones

# A Dashboard of Metrics for the Economy



# Our Approach

**Develop a new framework for measuring welfare change.**

- 1. Use cost and expenditure functions to re-work theory allowing for there to be free goods (with an imputable price).**
  - Using reservation prices and the total income method
- 2. Derive an explicit term that is the value of a new good's contribution to welfare change and GDP growth.**
  - Welfare change mismeasurement if it is omitted from statistical agency collections.
  - Derive a lower bound on the addition to real GDP growth from the introduction of a new good.

Details in Brynjolfsson, E., Collis, A., Diewert, W. E., Eggers, F., & Fox, K. J. (2025) GDP-B: Accounting for the value of new and free goods in the digital economy. *AEJ: Macro*.

# Empirical Approach

**Brynjolfsson, Collis & Eggers (*PNAS* 2019) propose an approach to directly estimate consumer welfare by running **massive online choice experiments**.**

- 1. We run incentive compatible discrete choice experiments**
  - “Incentive compatible” => participants risk losing access to the good
  - Recruit a representative sample of the US internet population via online survey panel
  - Use data to estimate the consumer valuation of Facebook
- 2. Quantify the adjustment term to real GDP growth (GDP-B) for the contribution of Facebook from 2004 to 2017**
- 3. Run additional incentive compatible discrete choice experiments to estimate the consumer valuation of several popular digital goods**
  - Instagram, Snapchat, Skype, WhatsApp, digital Maps, LinkedIn, Twitter, and Facebook
  - Conducted in a lab in the Netherlands
- 4. Explore the welfare gains from new goods: case study of smartphone cameras**

# Scale up using Google Consumer Surveys

The Daily Globe

Top Stories

World

US

Business

## Fair Use Digital Circulation Strategy Information Overload

*The Work of Art in the Age of Mechanical Reproduction*

Matthew Dodd |

Jurgen Habermas R&D Android cops  
beat The Weekender mathewi Tim  
Carmody attracting young readers  
tweets, collaboration tags the medium is  
the message blog plagiarism horse-race  
coverage advertising the other longer  
Book Review....



Privacy put the paper to bed Fuego news.me photo source: proimos/flickr

Please complete the following survey to access this premium content.

Would you prefer to keep access to Facebook or go without access to Facebook for one month and get paid \$5?

- Give up Facebook and get paid \$5
- Keep access to Facebook

# Key Findings

## 1. Choice experiments generate plausible demand curves

- Valuations are consistent across BDM lotteries, best-worst scaling and SBDC experiments
- Incentive compatible experiments often imply *higher* valuations

## 2. Median valuations

Search > email > maps > video > e-commerce > social media > messaging > music

## 3. Consumer surplus from Facebook in USA:

\$38/month for median consumer

## 4. This approach can be scaled up to numerous goods and services

# Facebook study with Incentive Compatibility

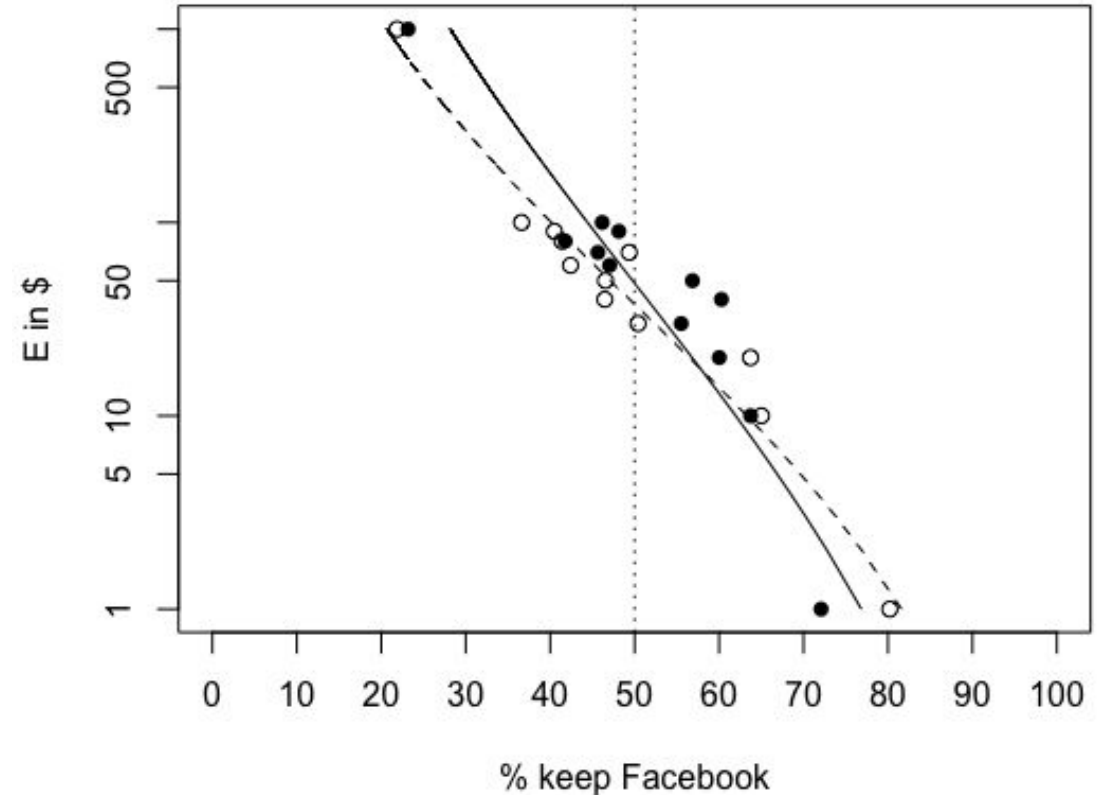
$n_{2016} = 1497, n_{2017} = 1388$

Median WTA:                   \$37.76 / month  
  [\$27.19, \$51.97]

## Heterogeneity in valuation

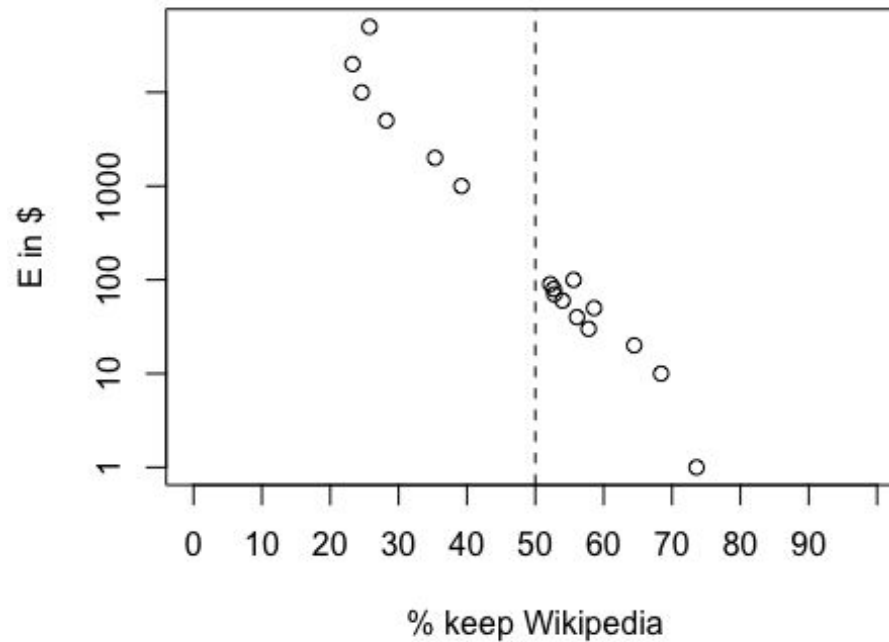
Higher valuations for people with

1. More time spent on Facebook
2. More friends they have
3. More frequent posting
4. More videos watched
5. Female
6. Older
7. Less use of Instagram or Youtube

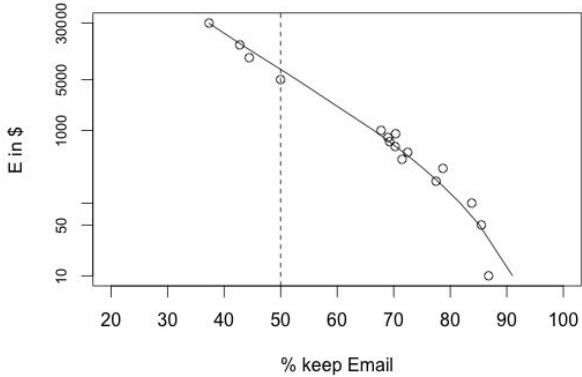


# Some Implied Demand Curves and WTA

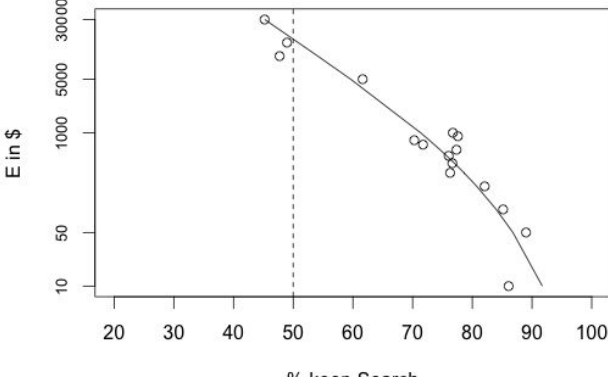
**Wikipedia:  $WTA_{\text{median}} = \$150/\text{year}$**



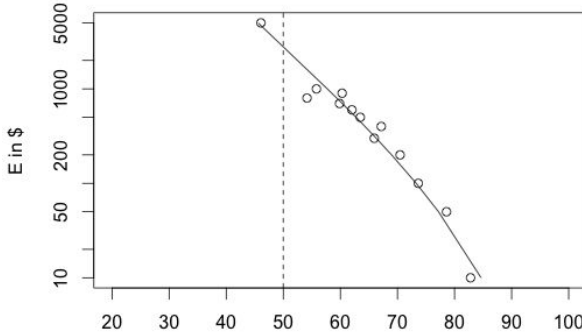
# Most widely used categories of digital goods



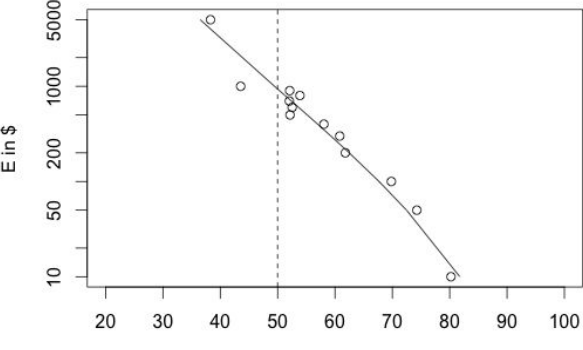
**Email**



**Search**

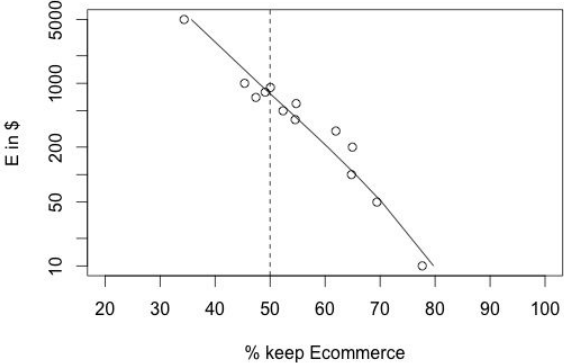


**Maps**

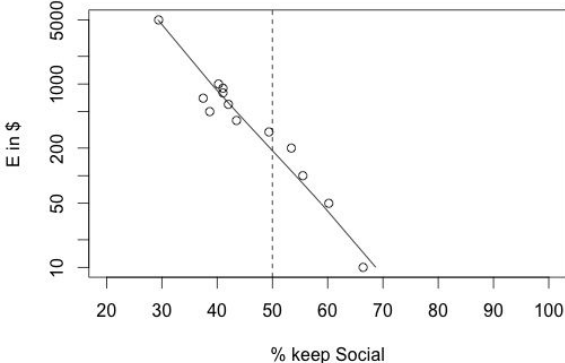


**Video**

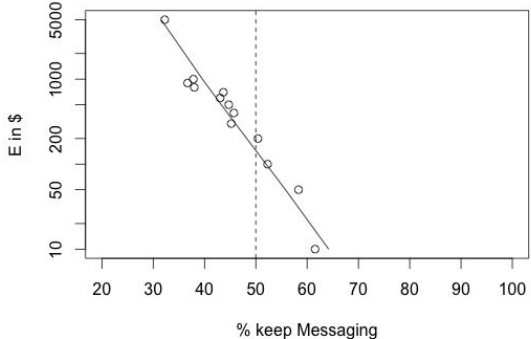
# Most widely used categories of digital goods



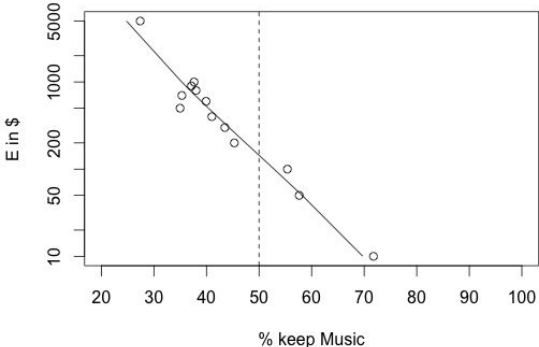
**E-Commerce**



**Social Media**



**Messaging**



**MUSIC**

# Welfare Change and the Free Goods Problem

Consumer holding  $Z^{**} > 0$  free goods has utility  $u^{**} \equiv f(x^{**}, z^{**})$ .

“Global” willingness to accept (WTA) function for the disposal of  $z^{**}$  as follows:

$$W_A(u^{**}, p, z^{**}) \equiv c(u^{**}, p, 0_M) - c(u^{**}, p, z^{**})$$

That is, the amount of expenditure needed to achieve the same utility without access to the free good.

# GDP-B growth from $t=0$ (before “free”) to $t=1$

- Consumer has a total income ( $T$ ) that is used to achieve the level of utility at an observed equilibrium,  $t=0,1$ :
- $T^t = p^t \cdot q^t + w^t \cdot z^t$  (market income plus imputed income), where  $z^0 = 0$
- GDP-B growth =  $T^1/T^0$ 
  - Deflate this by the GDP deflator

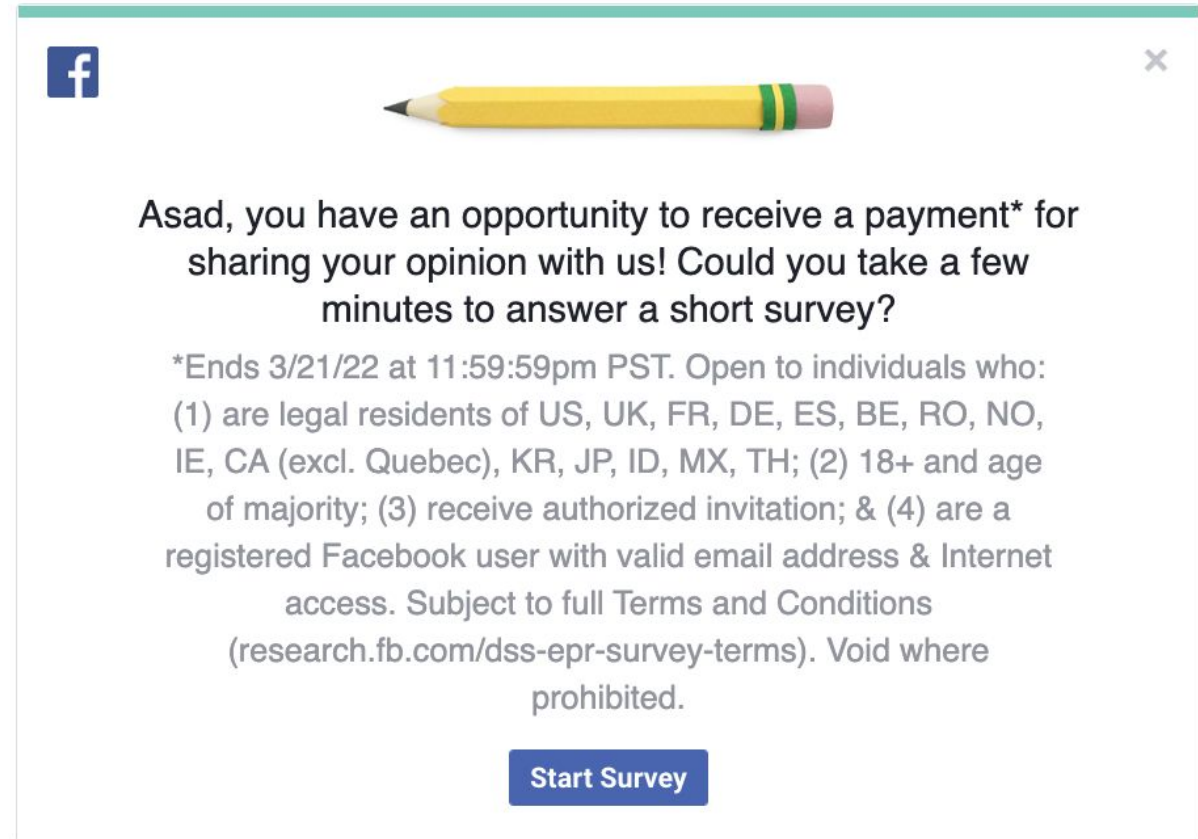
# Facebook's contribution to GDP-B

	GDP-B growth
Percentage Points, 2003-2017	0.68
Per year	0.05
GDP-B Growth per year without Facebook (i.e. GDP growth)	1.83
GDP-B Growth per year with Facebook	1.87

## ***2. The Digital Welfare of Nations***

# A platform for choice experiments

- Partnership with Meta
  - Large-scale (N=40,000) representative studies
  - strong incentive compatibility
- Weights using detailed information Meta has about its users to reduce bias from non-response
  - Differential Facebook use among different subpopulations
- International scope (13 countries)
  - FB penetration 60%



A screenshot of a Facebook notification for a survey. The notification features a blue Facebook 'f' logo in the top left corner and a yellow pencil icon in the top right corner. The main text reads: "Asad, you have an opportunity to receive a payment\* for sharing your opinion with us! Could you take a few minutes to answer a short survey?". Below this, there is a detailed disclaimer: "\*Ends 3/21/22 at 11:59:59pm PST. Open to individuals who: (1) are legal residents of US, UK, FR, DE, ES, BE, RO, NO, IE, CA (excl. Quebec), KR, JP, ID, MX, TH; (2) 18+ and age of majority; (3) receive authorized invitation; & (4) are a registered Facebook user with valid email address & Internet access. Subject to full Terms and Conditions (research.fb.com/dss-epr-survey-terms). Void where prohibited." At the bottom of the notification is a blue button with the text "Start Survey".

Asad, you have an opportunity to receive a payment\* for sharing your opinion with us! Could you take a few minutes to answer a short survey?

\*Ends 3/21/22 at 11:59:59pm PST. Open to individuals who: (1) are legal residents of US, UK, FR, DE, ES, BE, RO, NO, IE, CA (excl. Quebec), KR, JP, ID, MX, TH; (2) 18+ and age of majority; (3) receive authorized invitation; & (4) are a registered Facebook user with valid email address & Internet access. Subject to full Terms and Conditions (research.fb.com/dss-epr-survey-terms). Void where prohibited.

Start Survey

# Valuing Facebook using SBDC experiments

- We ensure compliance with actual usage
  - Important to avoid hypothetical bias
  - For FB, incentive compatible valuations are larger than hypothetical valuations
- Randomly choose price levels from \$5 to \$100
  - Equivalent amounts in other countries
- 1 in 50 chance of getting selected
  - ~800 selected, 400 chose deactivation and offered \$ (total \$20k paid out)
  - Almost all of them come back (2 years later 92.7% still active vs. 92.9% for the rest)



**Would you be willing to stop using Facebook for one month in exchange for \$40? You may be randomly chosen and offered \$40 to stop using Facebook based on your answer to this question.**

- No, I am not willing to stop using Facebook for one month in exchange for \$40
- Yes, I am willing to stop using Facebook for one month in exchange for \$40

Continue

# Best-Worst Scaling: Relative valuations

- Choose the most valuable and least valuable option from a list of items (Louviere, 1987)
- Relative valuations of different goods
  - Can include all types of goods including those for which incentive compatible design is difficult
  - Doesn't suffer from hypothetical bias
- Balanced incomplete block design to ensure all pairs of goods are evaluated together sufficiently
  - Each user answers 3 randomly selected questions



vs

The image shows the Facebook logo in its characteristic blue, lowercase font.

# Best-Worst Scaling: Relative valuations

- List of items: Most popular digital goods
  - Search Engine: Google
  - Social Media: Facebook, Twitter, Instagram, Snapchat, TikTok
  - Instant Messaging: WhatsApp
  - Maps: Google Maps
  - Video: YouTube
  - E-commerce: Amazon
- Also a non-digital item for comparison:
  - Meeting friends in person

 Facebook Survey

Which of these three situations are you **MOST WILLING** to experience and which are you **LEAST WILLING** to experience?

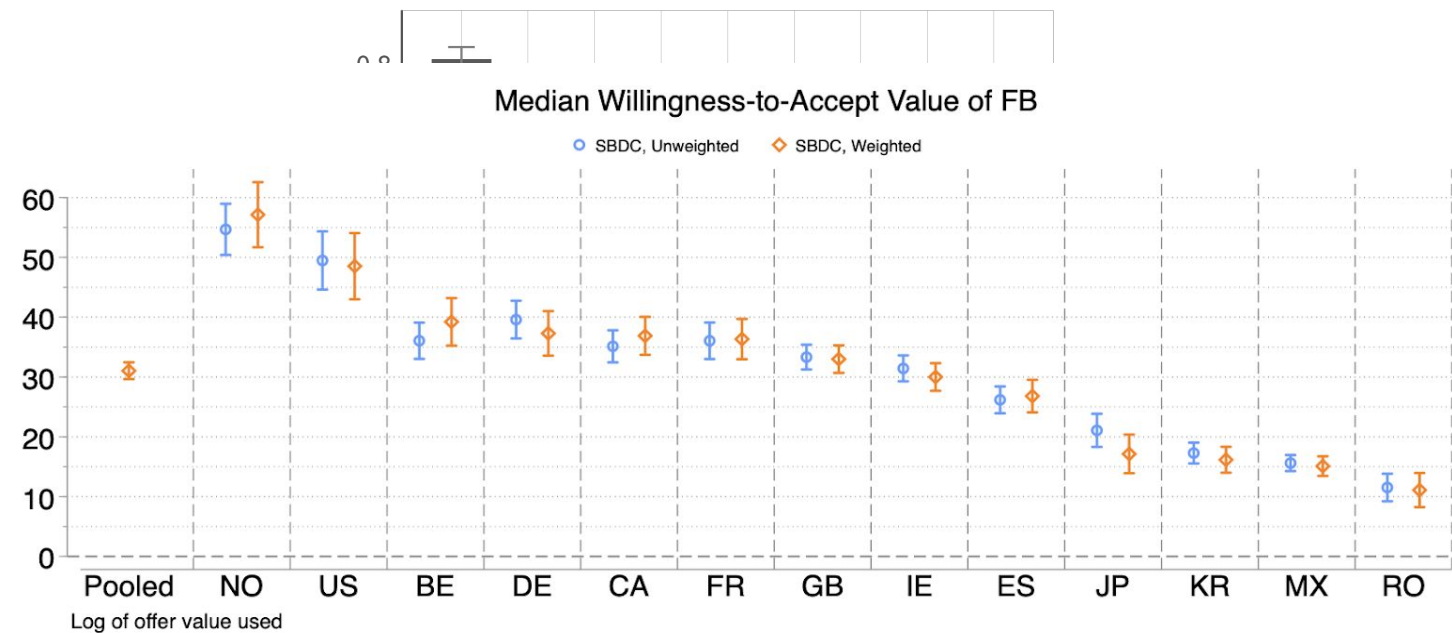
MOST WILLING		LEAST WILLING
<input type="radio"/>	You don't use YouTube for 1 month	<input type="radio"/>
<input type="radio"/>	You don't use WhatsApp for 1 month	<input type="radio"/>
<input type="radio"/>	You don't use Snapchat for 1 month	<input type="radio"/>

Submit

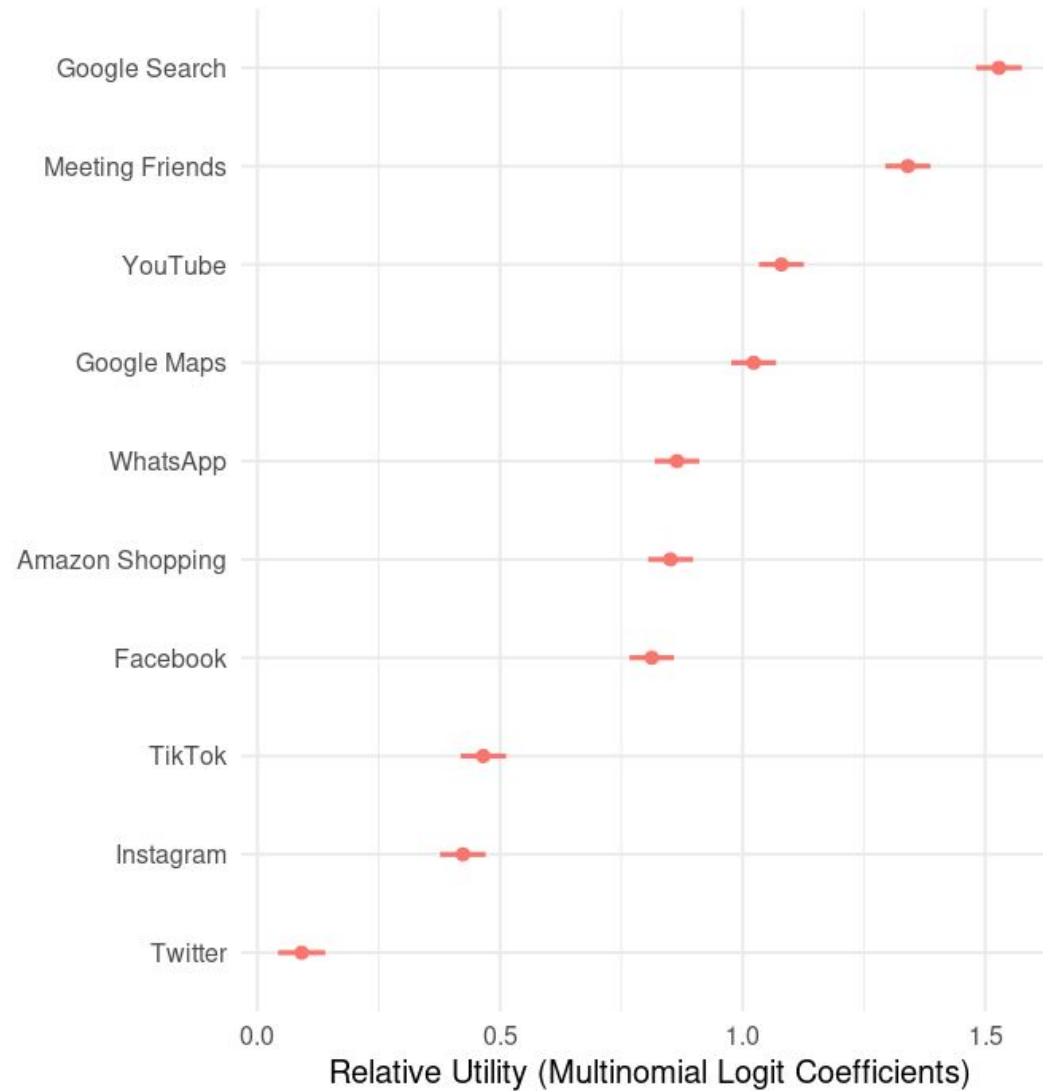
Use FB valuations from previous experiment to convert relative valuations to absolute dollar figures

# Facebook valuation using SBDC experiments

- 19% have a valuation < \$5 to give up Facebook for a month, 27% have a valuation > \$100
- Demand curve estimated using weighted logit model
- Overall median monthly valuation across 13 countries: \$31
  - Min. \$11 in Romania, Max. \$57 in Norway



# Relative valuations using Best-Worst Scaling

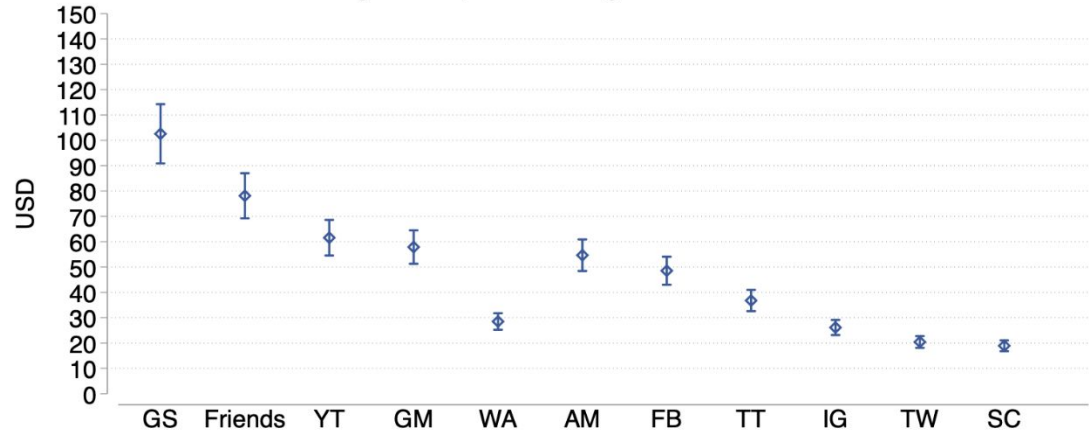


# Valuations of digital goods across countries

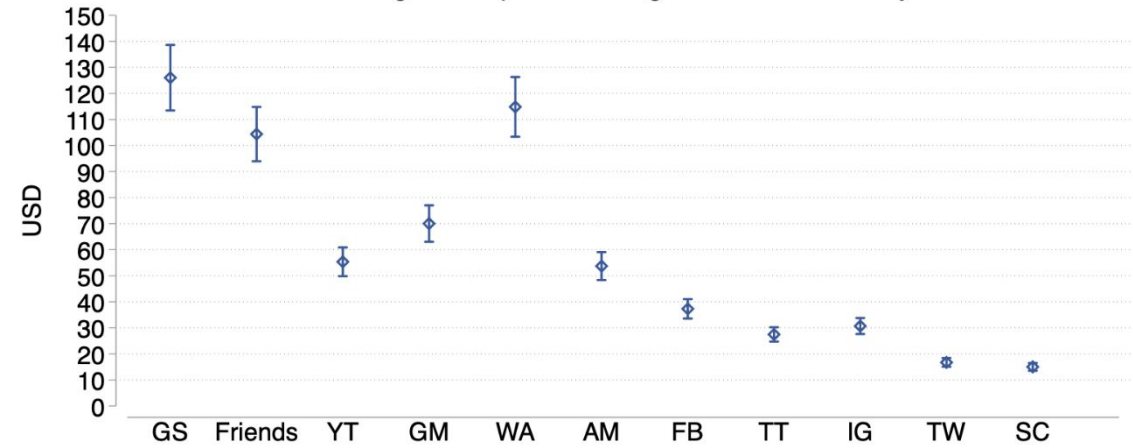
- Calibrate valuations of other digital goods using incentive-compatible SBDC Facebook valuations and relative utilities from best-worst scaling

# Valuations of digital goods across countries

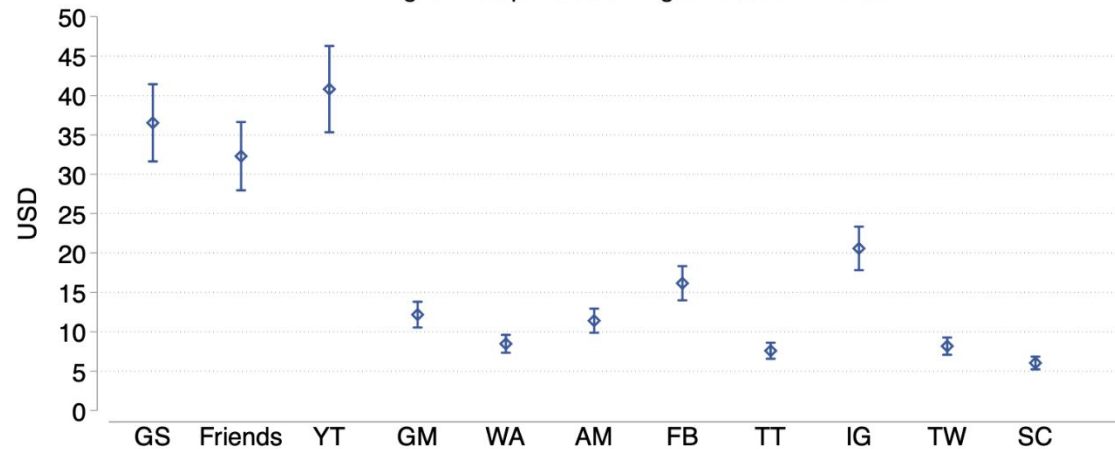
Willing-to-Accept Value of Digital Goods in United States



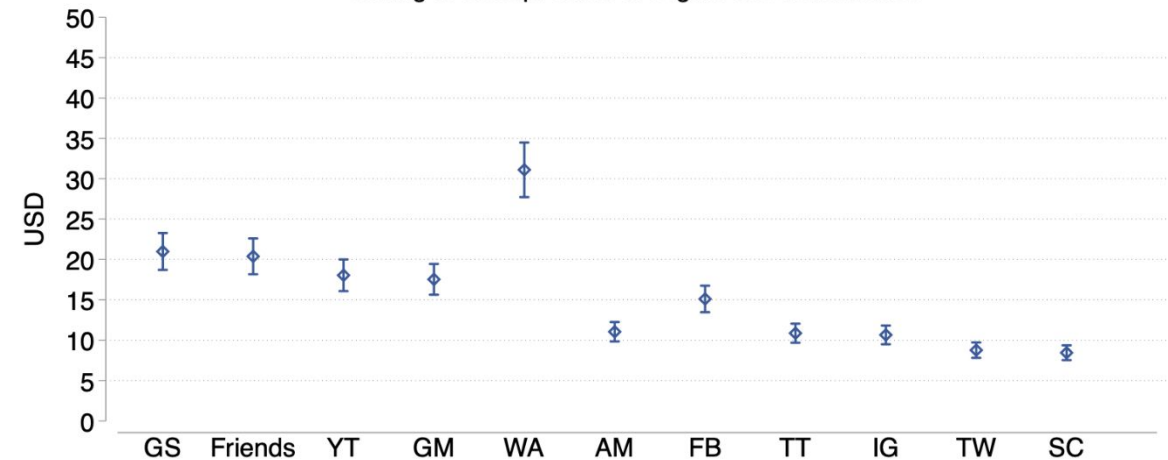
Willing-to-Accept Value of Digital Goods in Germany



Willing-to-Accept Value of Digital Goods in Korea



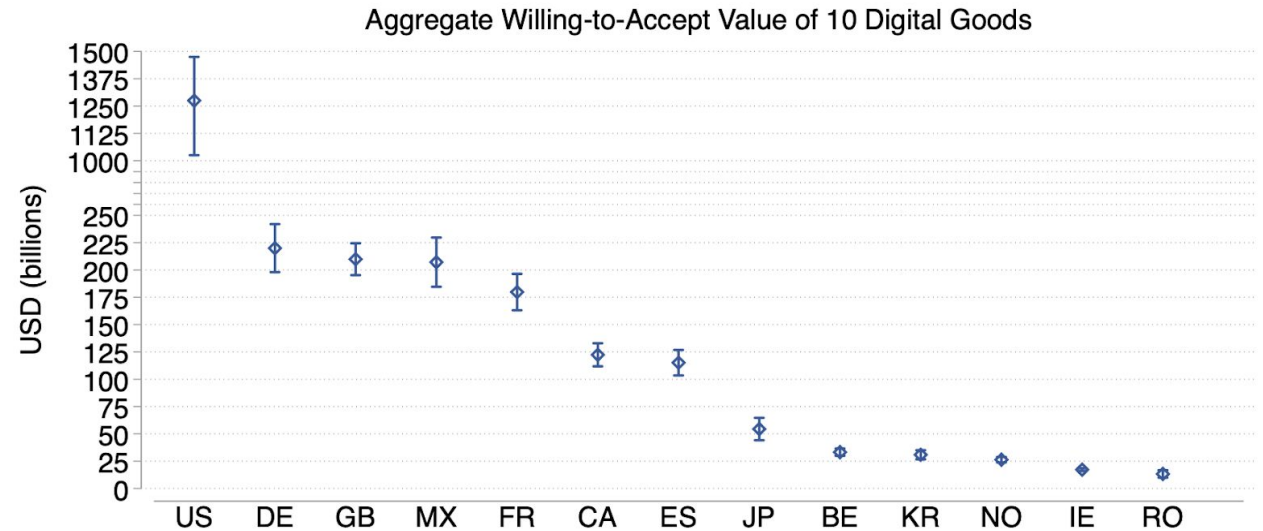
Willing-to-Accept Value of Digital Goods in Mexico



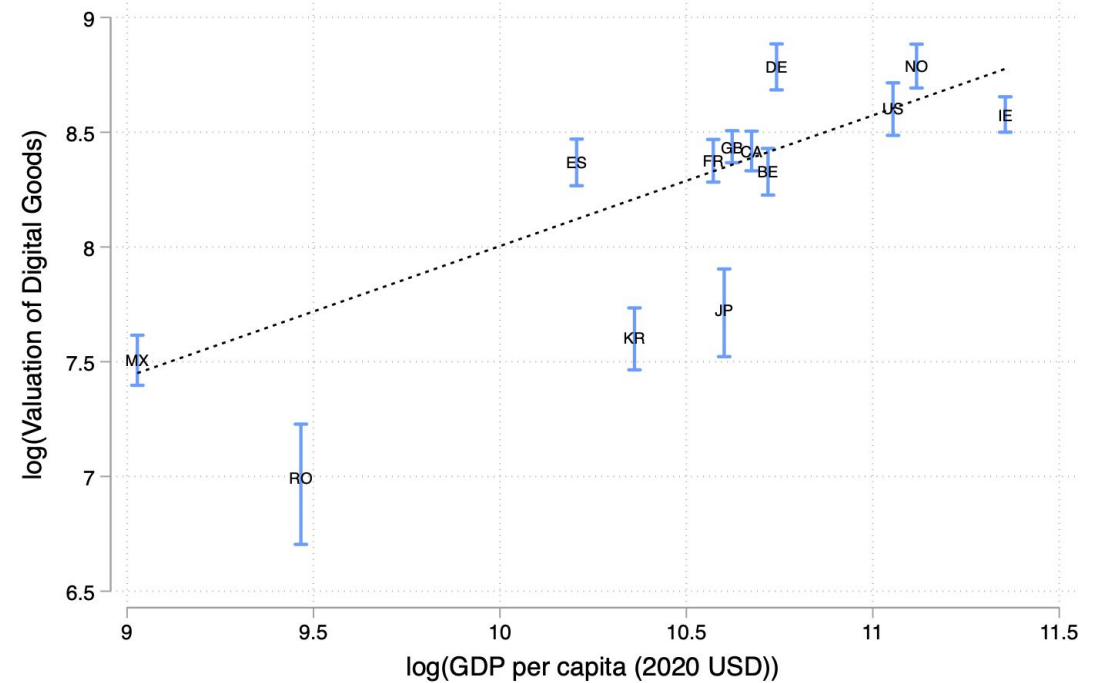
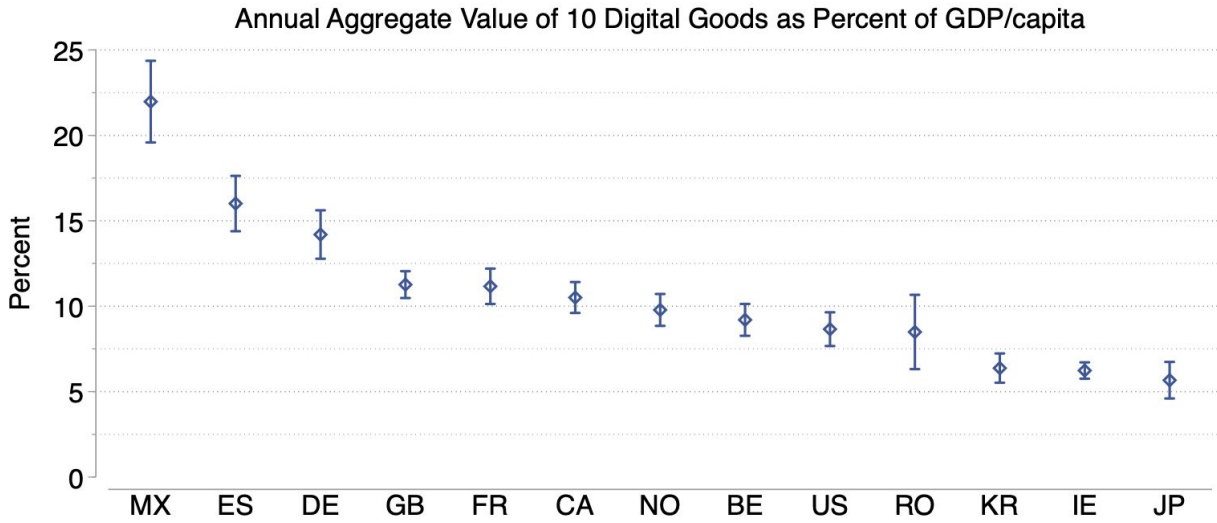
# Welfare gains from digital goods across countries

These 10 digital goods generate a total of \$2.5 trillion in welfare across these countries

- Ranging from \$13 billion in Romania to \$1.29 trillion in US



# Welfare gains as a percent of GDP

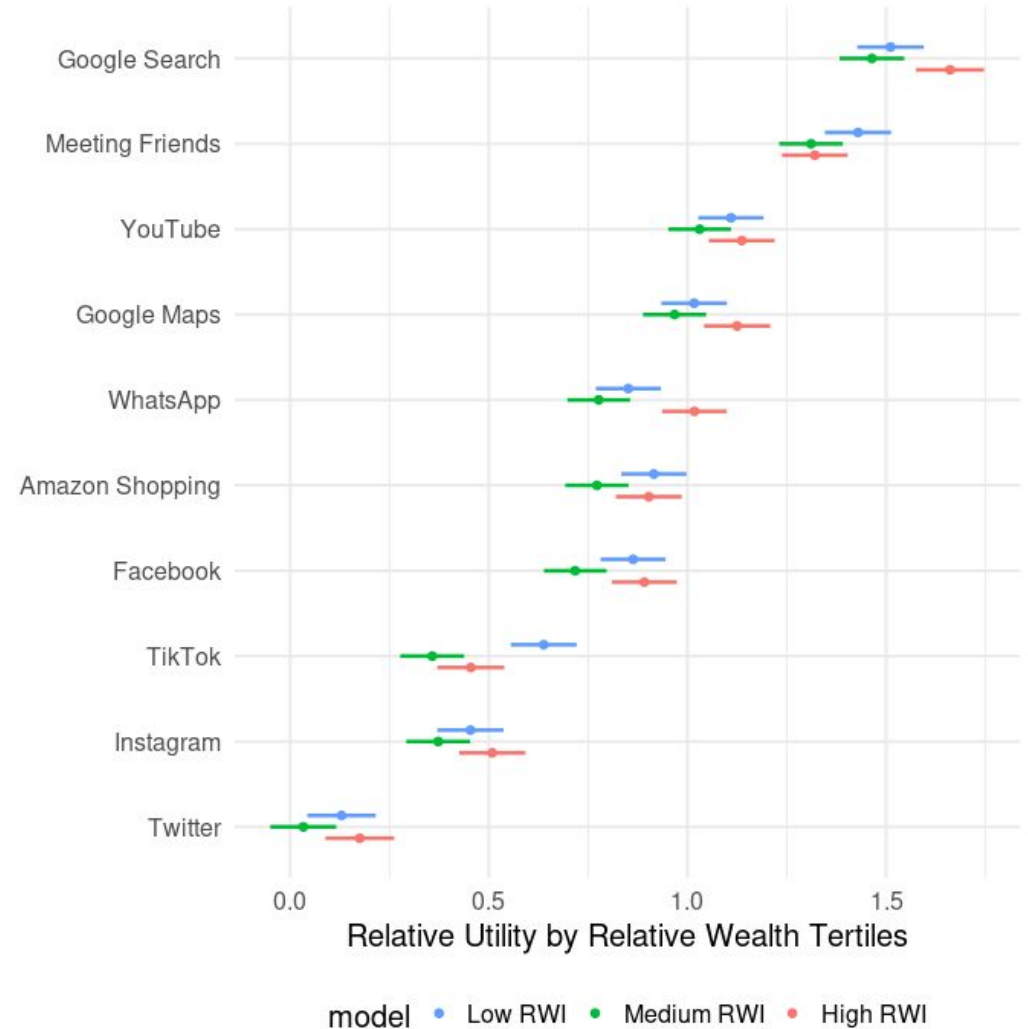


**Takeaway: Digitization reduces welfare inequality across countries.**

A 1% increase in a country's GDP per capita is associated with a 0.68% increase in users' valuation of the 10 digital goods relative to GDP per capita.  
( $\beta = 0.68$ , p-value = 0.002).

# Welfare gains within countries

- No systematic association between wealth and valuations
  - Differences rarely significant across terciles -> lower RWI users benefit more
  - Poorest and wealthiest users have greater value for most digital goods
  - In some cases (Google Search), users in top tercile derive highest welfare
  - In some other cases (TikTok), users in lowest tercile derive highest welfare



# Digitization and welfare inequality

## **Conclusion: Welfare gains from digital goods disproportionately benefit lower-income countries and lower-income individuals within the US**

- Free digital goods are available to both higher and lower-income individuals and at the same quality
  - A billionaire has access to the same quality search engine or messaging app as the average person
  - Lower-income individuals are expected to consume higher quantities of free digital goods because they have less money to spend on other goods, and also because their opportunity cost of time is lower



r/Showerthoughts · 6 yr. ago  
university

The super-rich have fast cars, private jets, many large homes, exclusive designer fashion, live-in chefs, original artworks, and the same phone as you.

### ***3. Surplus from digital vs. conventional goods***

# Scale Up GDP-B to a Representative Basket of Goods

## Team

- Erik Brynjolfsson, Avi Collis, David Nguyen, and Sophia Kazinnik

## Assess valuations for basket of hundreds of goods and services

- Both traditional and digital
- Also experimenting with several public goods and quality of life metrics
- Included several attention and robustness checks in online experiment

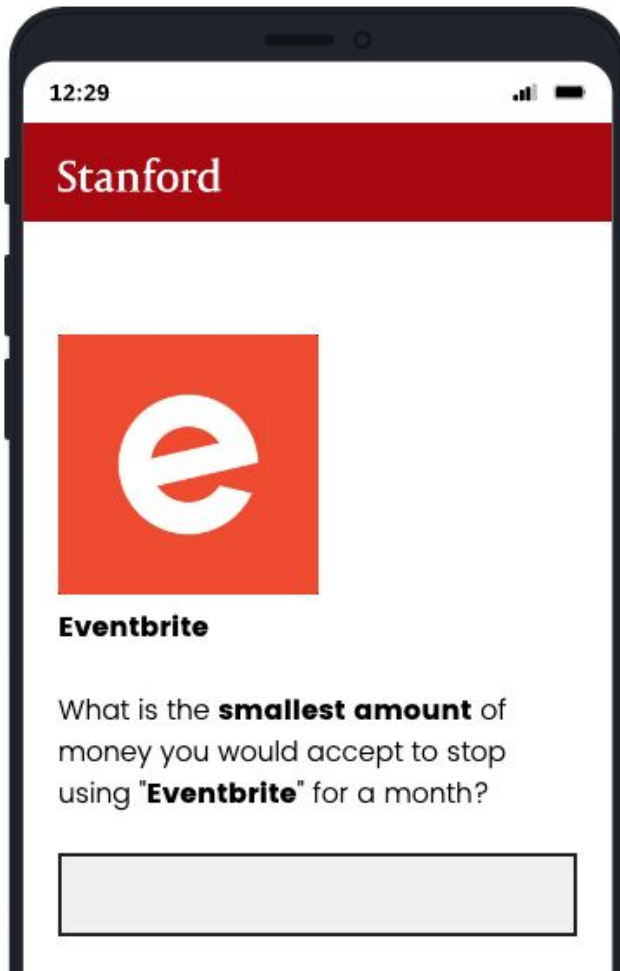
## Run on Prolific platform

- About 30,000 respondents
- Selected to be representative of demographic groups
- Ongoing, first waves completed
- Analysis underway – very preliminary results

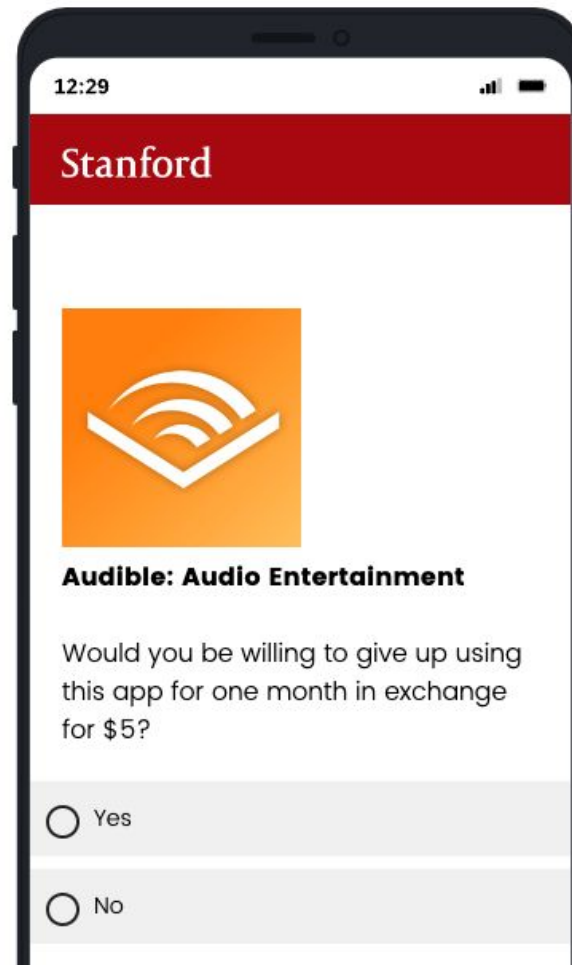
# Item types – Examples

- **Traditional:** Milk, Coffee, Rib Eye Steak, Frying pans, Sneakers, Vehicles, Contact lenses, TVs, Smartphone, Laptop, Financial / tax services, Life insurance
- **Digital:** Google search, ChatGPT, Wikipedia, WhatsApp, PayPal, TikTok, investment apps
- **Public goods:** Electricity at home, healthcare, police, roads, national defense
- **Quality of life:** Not suffering from cancer / poor air / chronic pain / identity theft, freedom of speech
- **Reference items:** Earning \$100 / \$500 / \$1,000 / \$2,000 less for next month or year

## Open ended text



## Cash offer



- Intuitive, no training needed.
- Works well on mobile.
- Flexible design, works for monetary and non-monetary offers.
- Easy to add logos and pictures.

# Data collection process



Run massive Single Binary Discrete Choice surveys via Prolific, representative of US population (N=33,000) + open box survey (N=4,400).



Data collected in December 2024 and March 2025.



Average response time of 2:30 minutes for 15 questions.



Explicit and implicit attention checks (i.e. “fake apps”) passed by 97.3% of respondents.

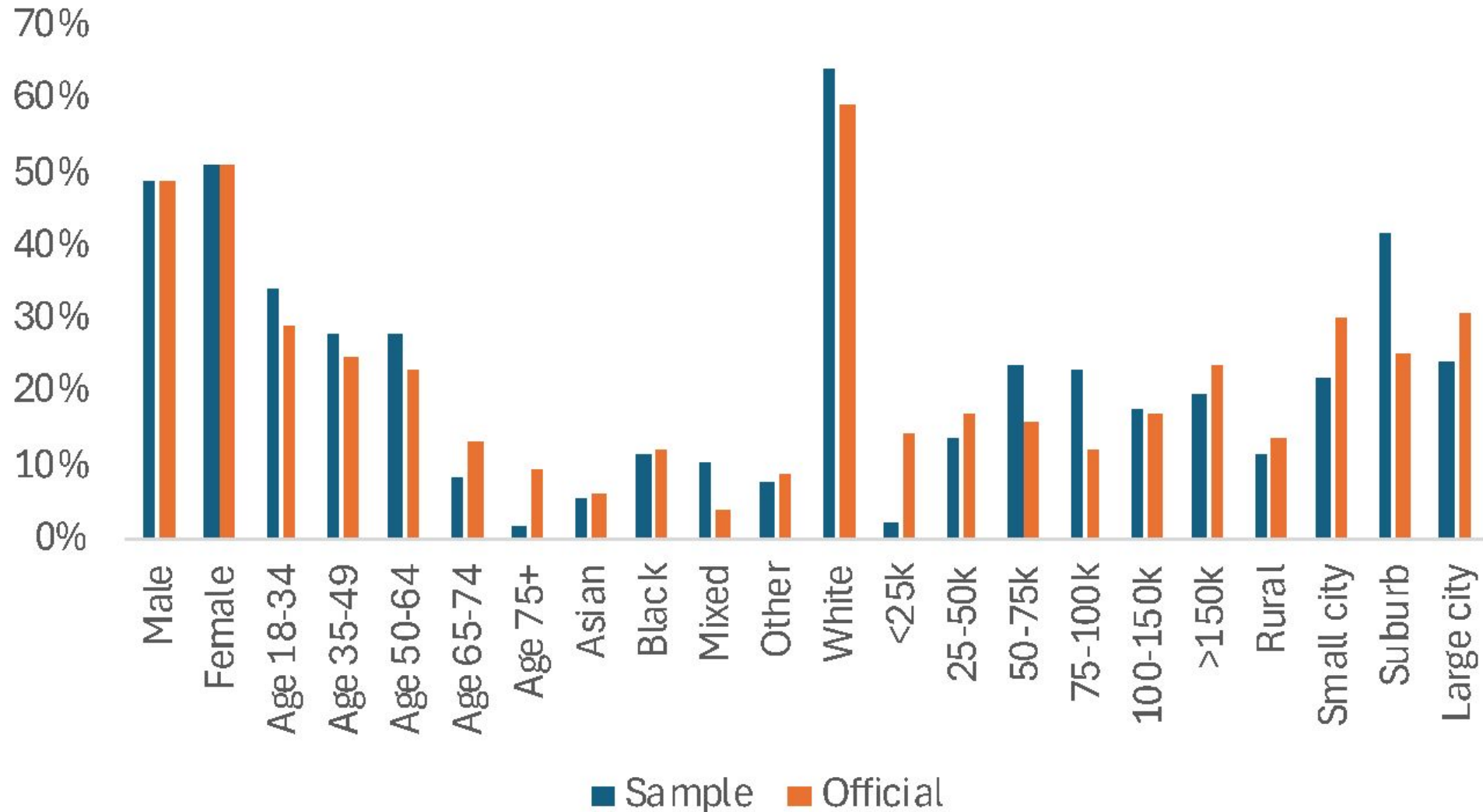
# GDP-B: Traditional consumption items

- The goal is to build a **representative basket for the traditional and digital economy**.
- For traditional items we rely on Consumer Expenditure Surveys (CE) from the Bureau of Labor Statistics (BLS).
- We choose 264 items covering most of the consumer spending apart from housing.
- For example: eggs, apples, cookies, health insurance, sports equipment,

Level	All items		In GDP-B	
	Freq.	Percent	Freq.	Percent
1	14	1.98	4	1.52
2	84	11.90	34	12.88
3	188	26.63	66	25.00
4	215	30.45	78	29.55
5	150	21.25	71	26.89
6	38	5.38	11	4.17
7	17	2.41	-	-
Total	706	100.00	264	100.00

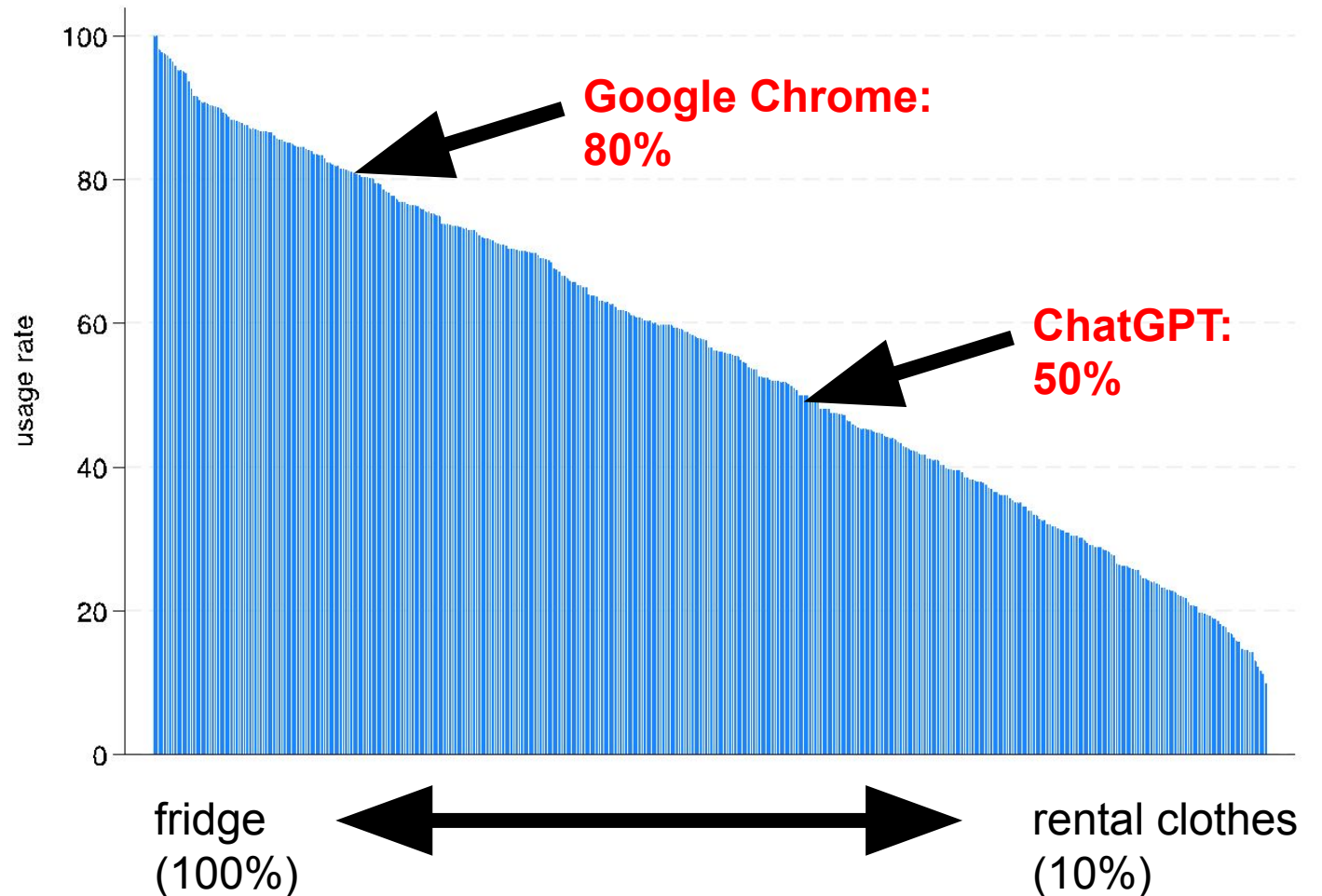
Table 2: Number of items in GDP-B by level

# Survey representativeness



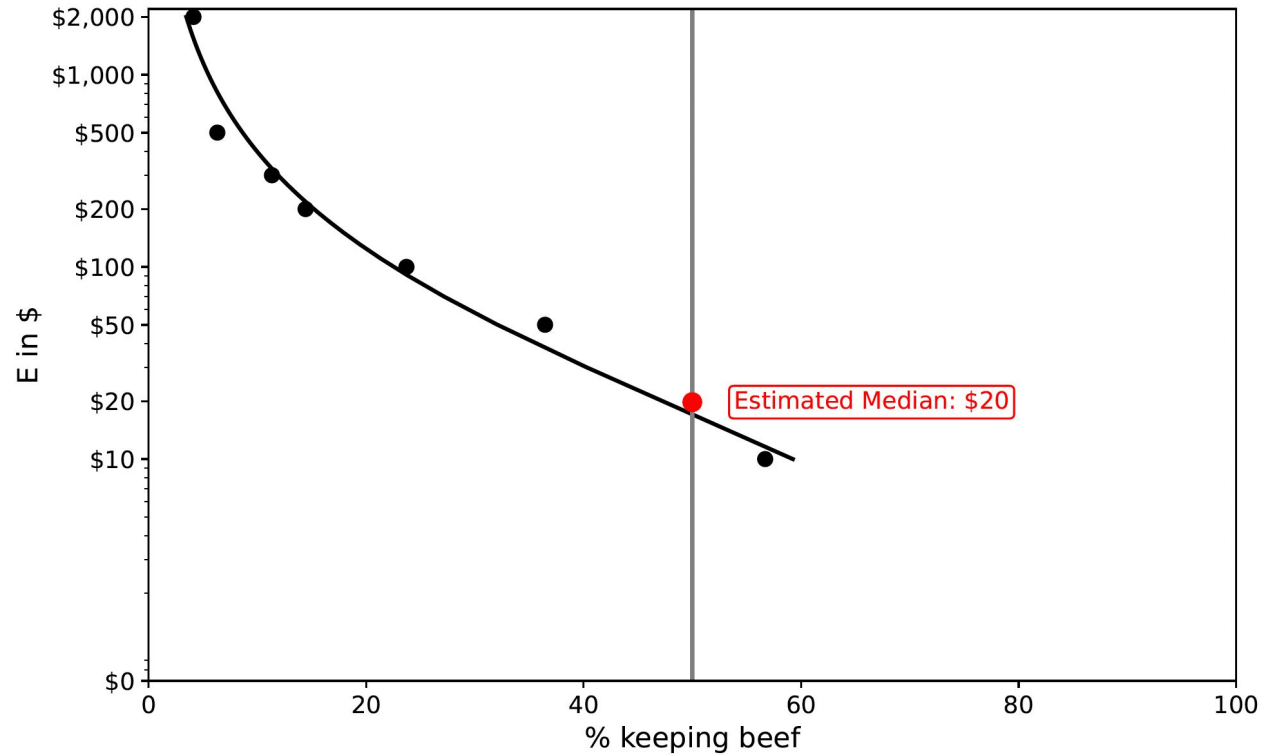
# Usage rates range from 10% to 100%

- Average usage rate is 44%, ranging from 10% (rental clothes) to 100% (fridge).
- High usage rates (80% or more) observed for one third of items.
- Low usage rates (30% or less) observed for 17% of items.

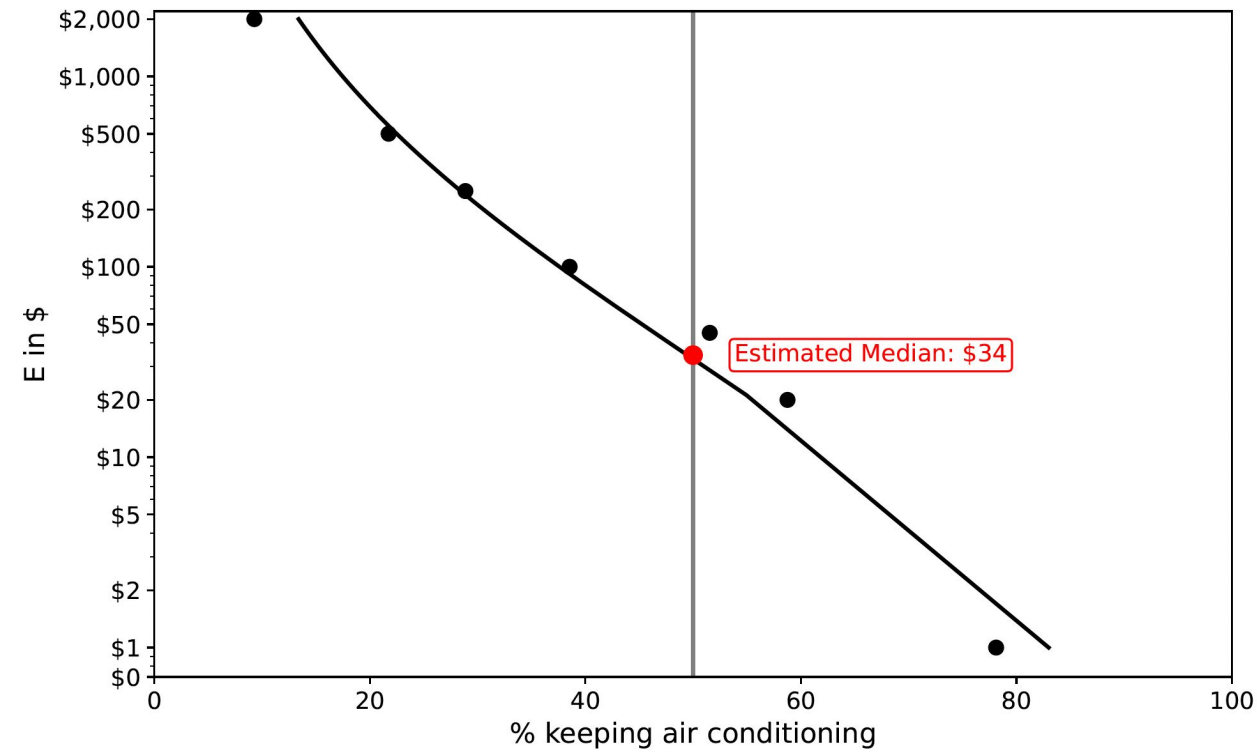


# Traditional item examples: Beef and AC

## Beef (any type)

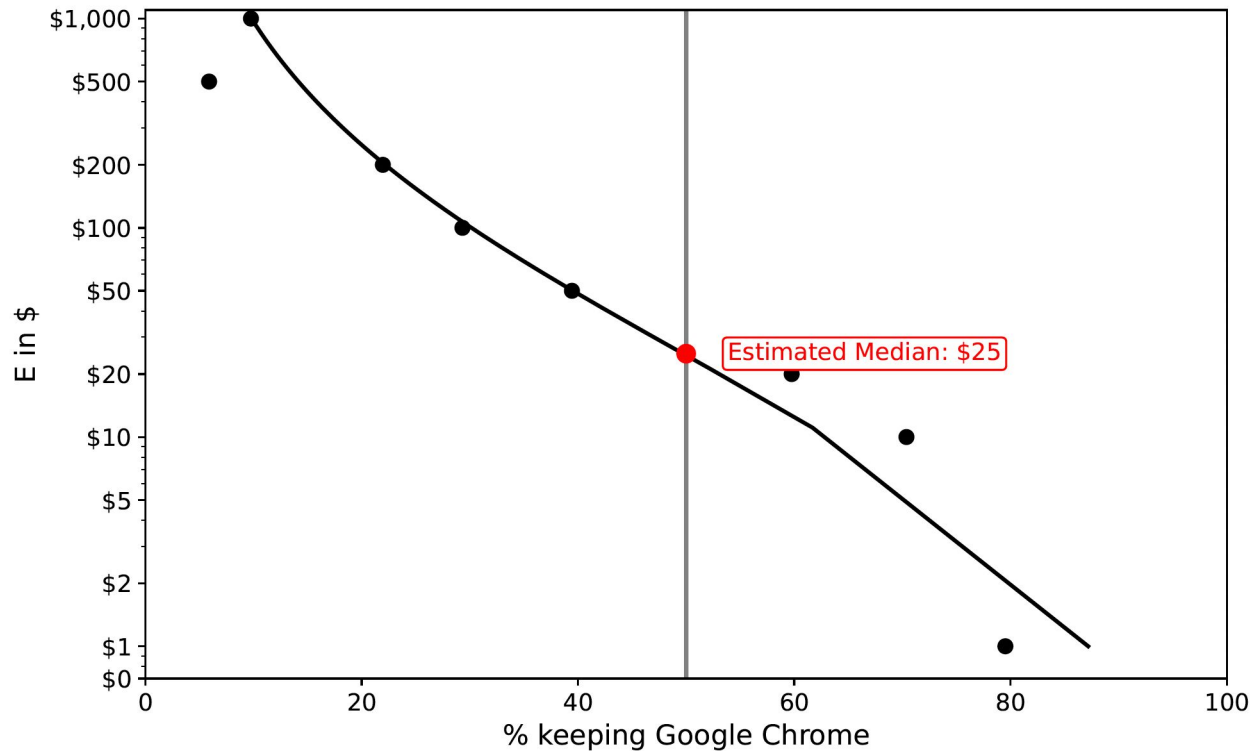


## Air conditioning

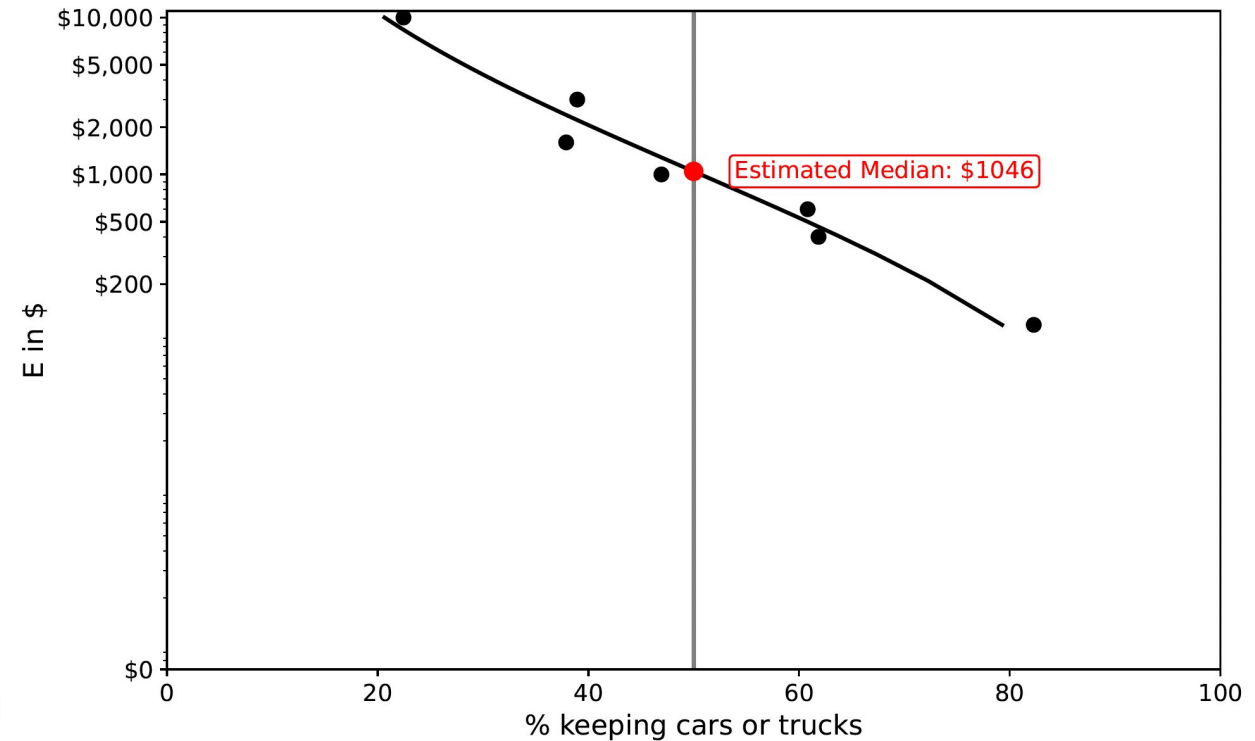


# Google Chrome and Private Cars

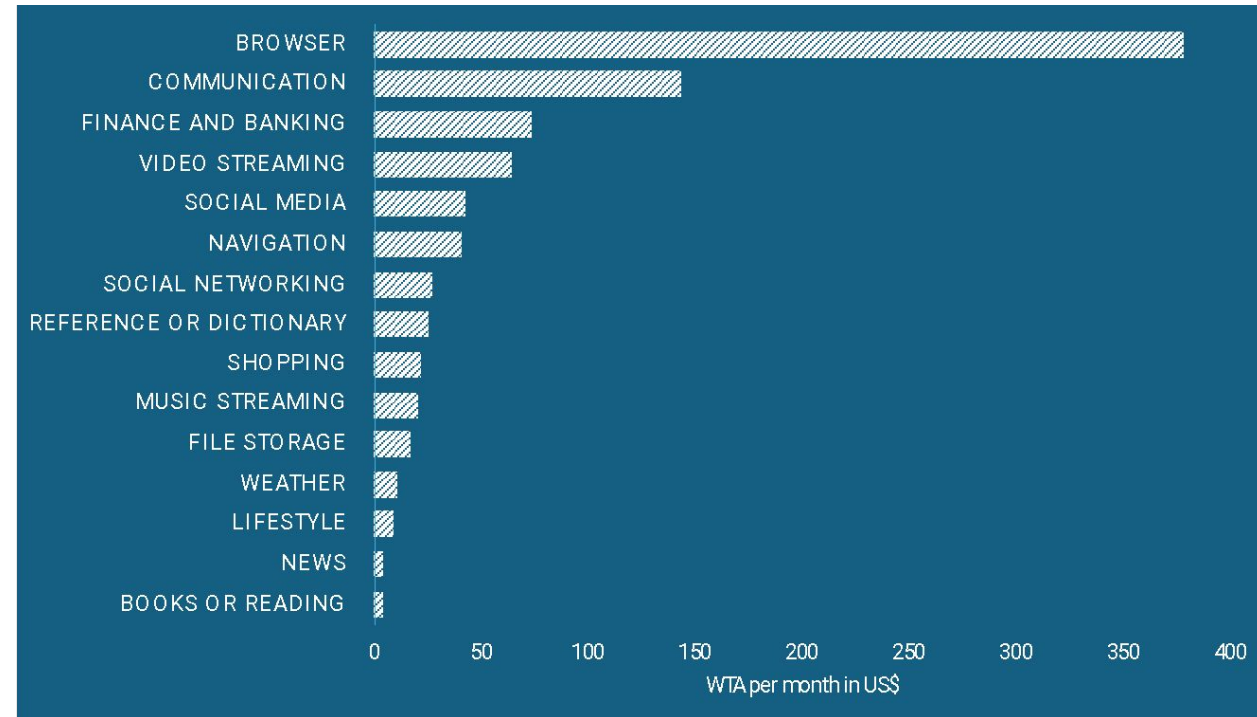
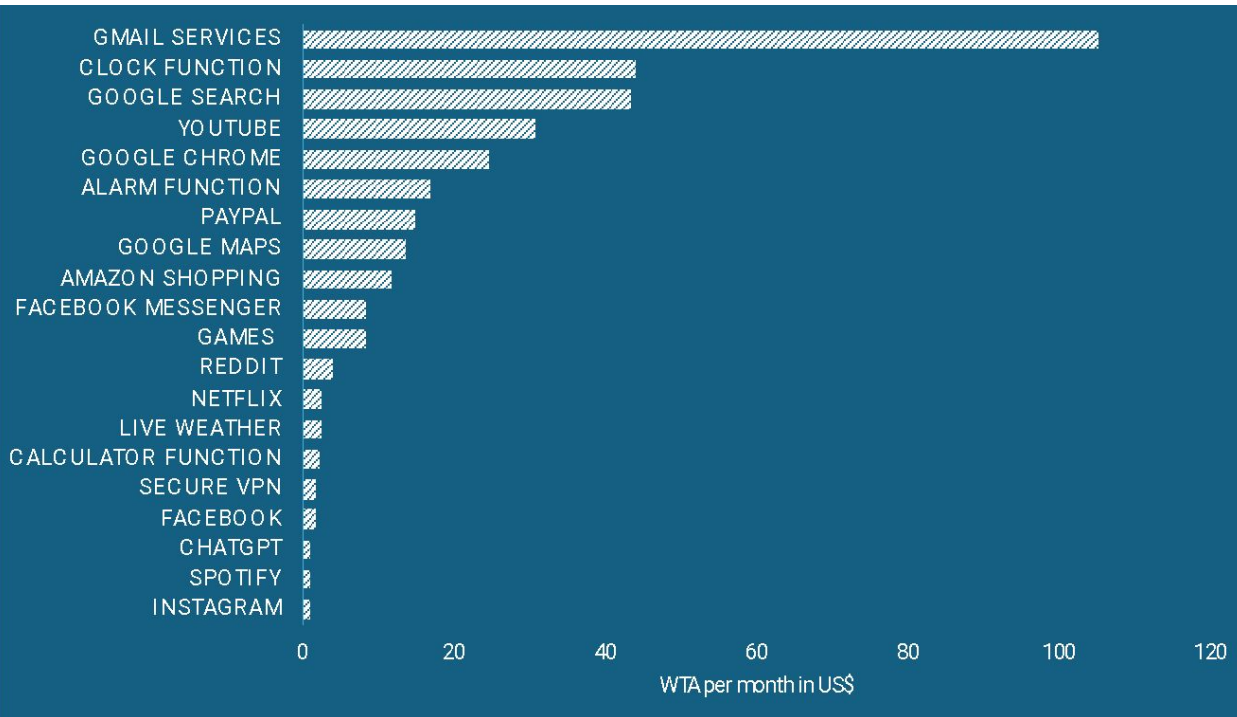
## Google Chrome



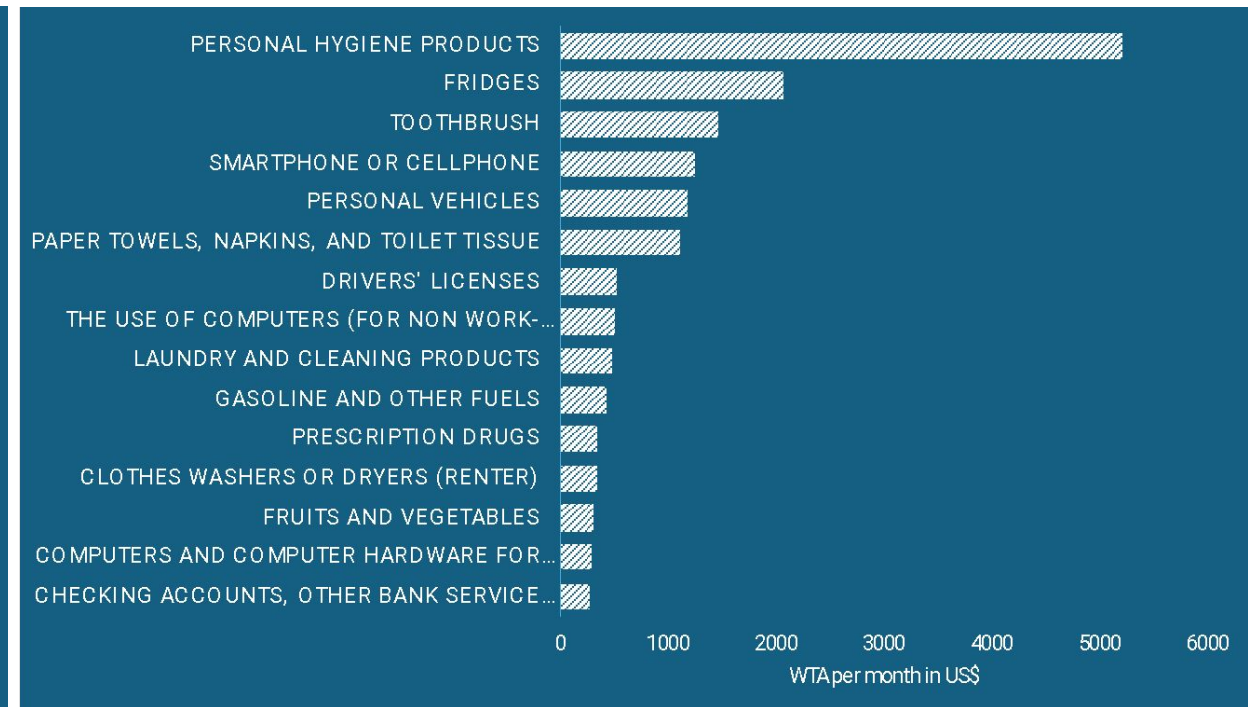
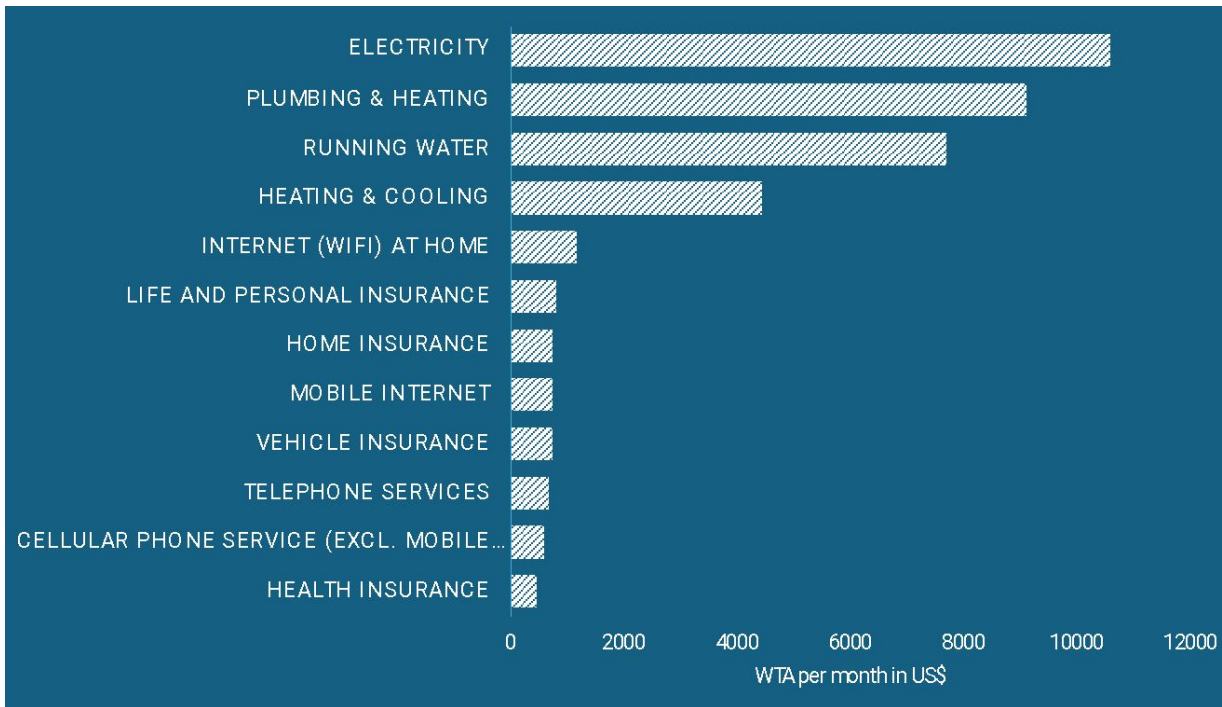
## Cars (private use)



# GDP-B: Digital items and categories



# GDP-B: Basic utilities and traditional items



# GDP-B for the digital economy

- The surplus from all 24 digital categories adds up to **\$900 per month**, around **18% of average monthly income** in the US.
  - Accounts for substitutability within categories (e.g. Google v Apple Maps) but not across categories (e.g. Social media v news).
- Not all digital items are free (Netflix, Spotify, ChatGPT), so **GDP-B measures surplus on top of savings**.
  - For example, surplus from video and audio streaming is \$166, needs to be added to the \$31 the average person spends on these per month.
  - Google Maps is free so it's 100% surplus.

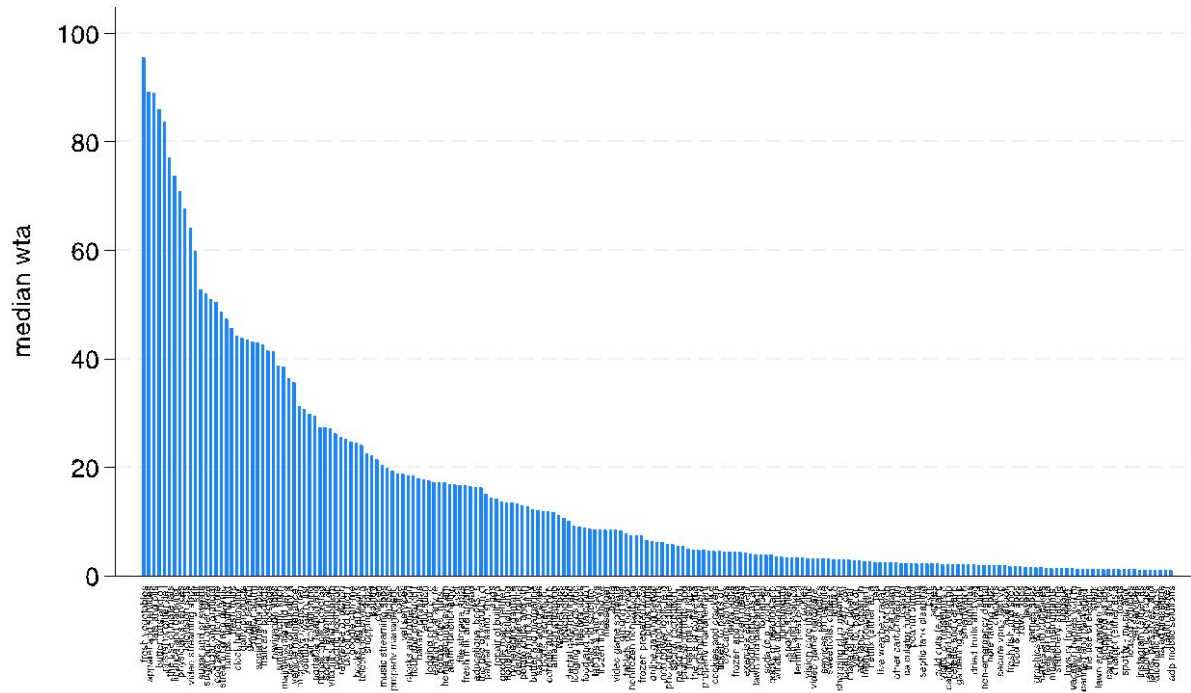
# Substitution effects for traditional items

Meats, poultry, fish, and eggs						\$179.9 \$38.8 4.6
\$16.8 Beef 1.6 \$10.3	\$2.2 Pork 0.3 \$7.2	\$22.1 Poultry 3.1 \$7.1	\$3.3 Fish and seafood 0.6 \$6.1	\$14.4 Eggs 4.4 \$3.3	\$4.3 Other meats 0.9 \$5.0	\$63.1 \$38.8 1.6
\$8.3 Ground beef 2.1 \$4.7	\$0.9 Bacon, breakfast sausage, and... 0.3 \$2.6	\$21.4 Chicken 4.0 \$5.3	\$1.7 Fresh fish and shellfish 0.5 \$3.5		\$0.1 Frankfurters 0.07 \$1.4	\$40.6 \$30.1 1.3
\$1.2 Roasts 0.9 \$1.3	\$0.3 Pork chops 0.3 \$1.2	Other poultry	\$0.9 Processed fish and seafood 0.3 \$2.6		\$3.5 Lunchmeats 1.2 \$2.9	
\$1.2 Steak 0.4 \$3.0	\$0.8 Ham 0.7 \$1.2				\$0.3 Lamb, organ meat, and game 0.75 \$0.4	
Other beef	Other pork					

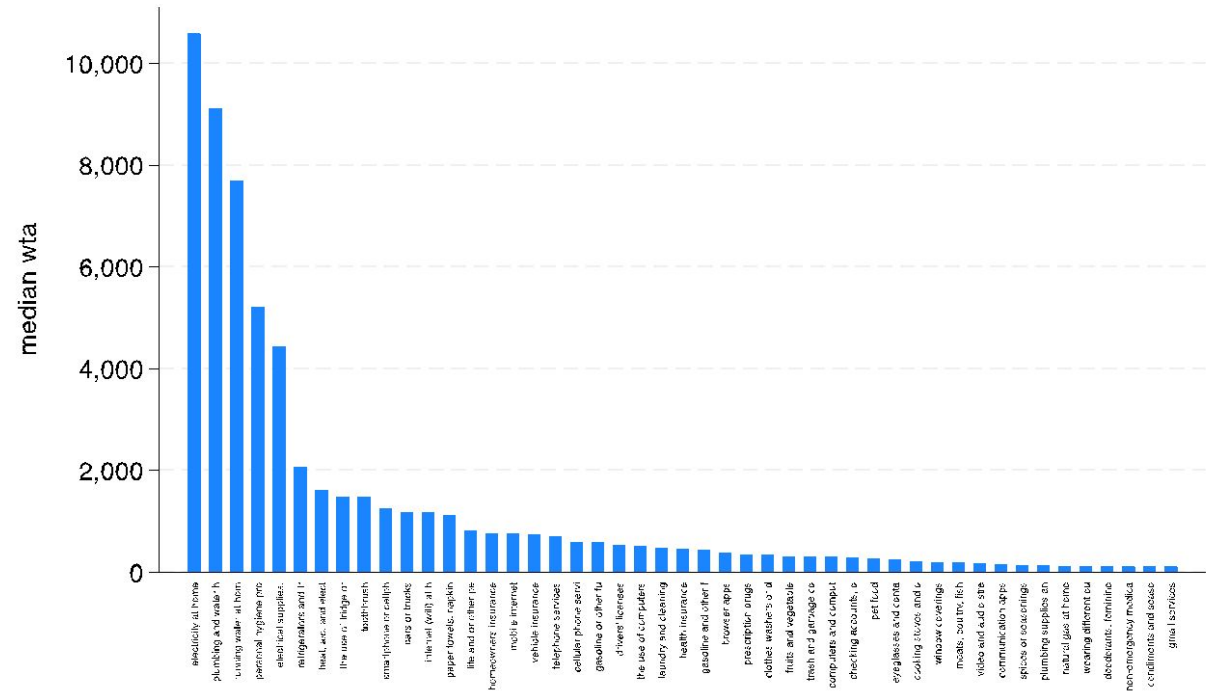
Green is consumer surplus. Red is expenditure. Yellow box is the ratio.

# GDP-B results: summary

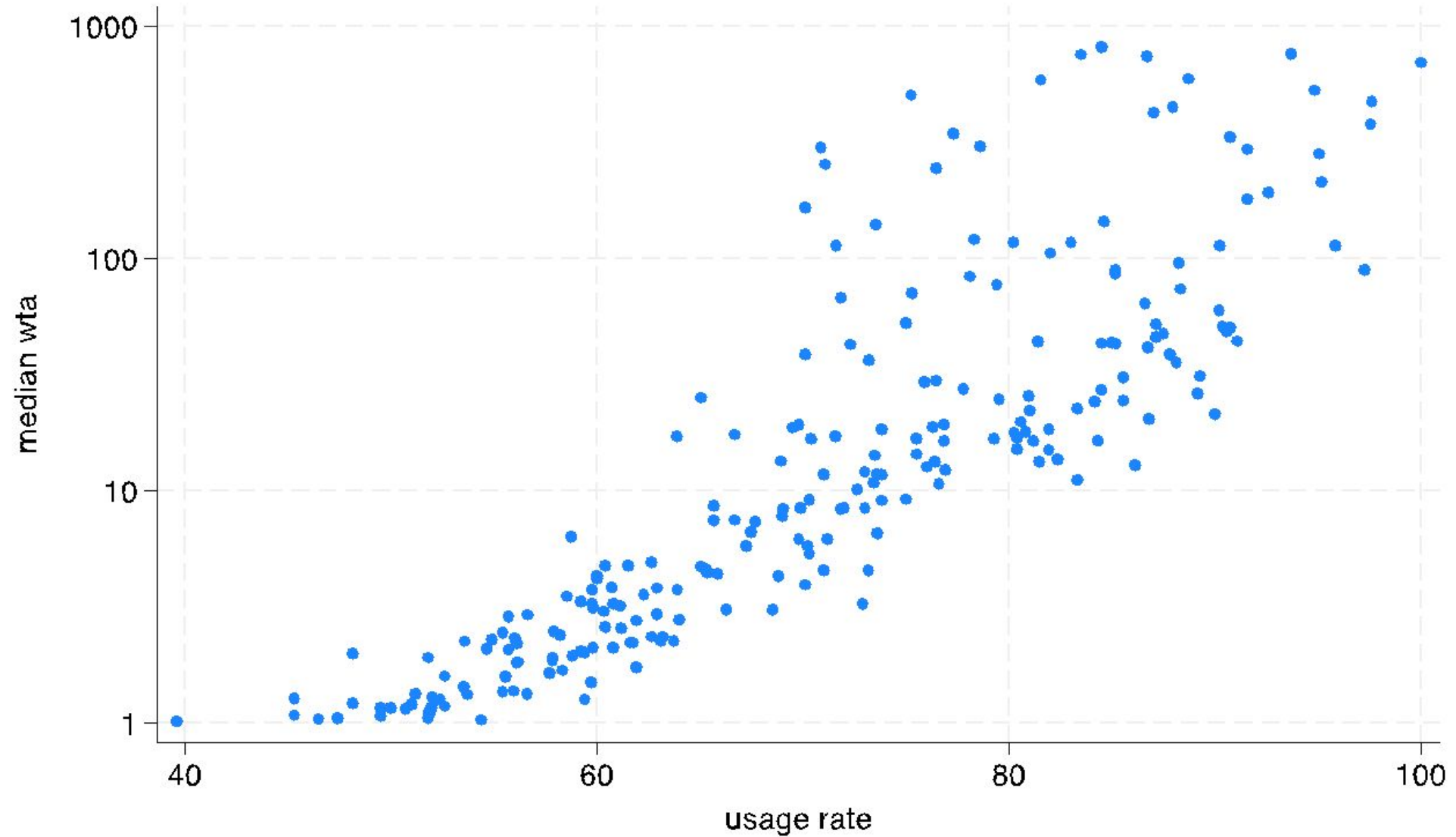
**Median WTA =  
\$1-100  
(203 items)**



**Median WTA =  
\$100+  
(47 items)**



# Usage correlated with WTA



# Findings – very preliminary

- As expected, highest median valuations for utilities at home, healthcare, insurance, and personal hygiene, but also internet and electronic devices.
  - Electricity at home,
  - Running water at home
  - Personal hygiene products
  - Cars, health insurance, mobile internet, wifi, computer, smartphone, fridge, toothbrush
- Lowest median valuations for tinder app, riding motorcycles, tobacco, cigarettes, preschool services.
  - But: tobacco and cigarettes have high mean, indicating high valuation of those that *do* smoke.
- Reassuring: Very low valuations for fake apps.
- Access to ChatGPT has median valuation of \$25/month
  - Comparable to eating bananas or going to the museum.

# ***4. LLM Agents for Economic Research***

# Issues with Survey Data

Survey research is facing a dual challenge:

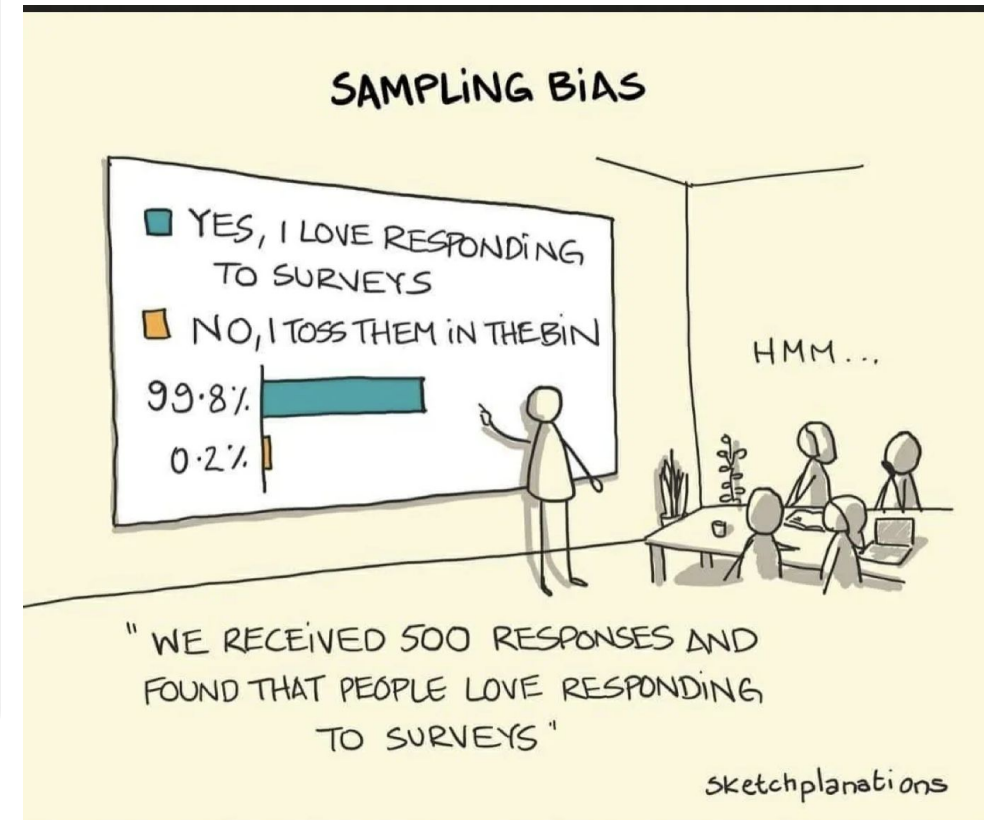
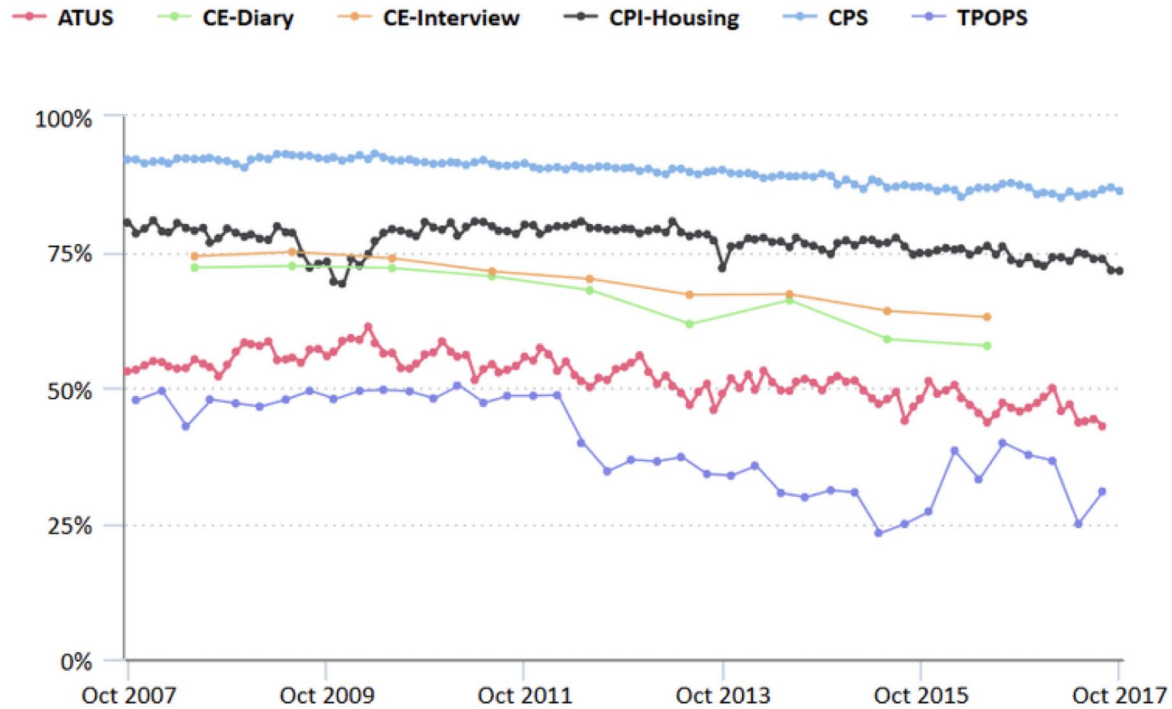
- 1) Escalating costs
- 2) Deteriorating response rates

Survey data is getting worse:

- Demographic patterns of non-response
- Dependence on online platforms (e.g., MTurk, Prolific) – “professional respondents”, identity misrepresentation, etc.
- Respondents are using LLMs...(Zhang et al., 2025; Veselovsky et al.,2023)

# Issues with Survey Data

Household survey response rates, October 2007–October 2017



# LLMs as Synthetic Respondents

Growing body of literature shows that LLMs produce responses consistent with documented patterns of human behavior:

- Computer science (*Park et al., 2024*)
- Economics (*Horton, 2023*)
- Psychology (*Dillion et al., 2023; Kosinski, 2023*)
- Sociology (*Bail, 2024; Ashokkumar et al., 2024*)

→ LLMs as survey data sample extenders

# LLM Agents as Survey Respondents

LLMs calibrated on human input would be a first step to:

- Richer high-frequency data at a lower cost through hybrid surveys
- Scaling up current work (e.g., to estimate GDP-B)

# LLM Agents as Survey Respondents

LLMs calibrated on human input would be a first step to:

- Richer high-frequency data at a lower cost through hybrid surveys
- Scaling up current work (e.g., to estimate GDP-B)

Could LLMs replicate survey responses? What are best practices?

# LLM Agents as Survey Respondents

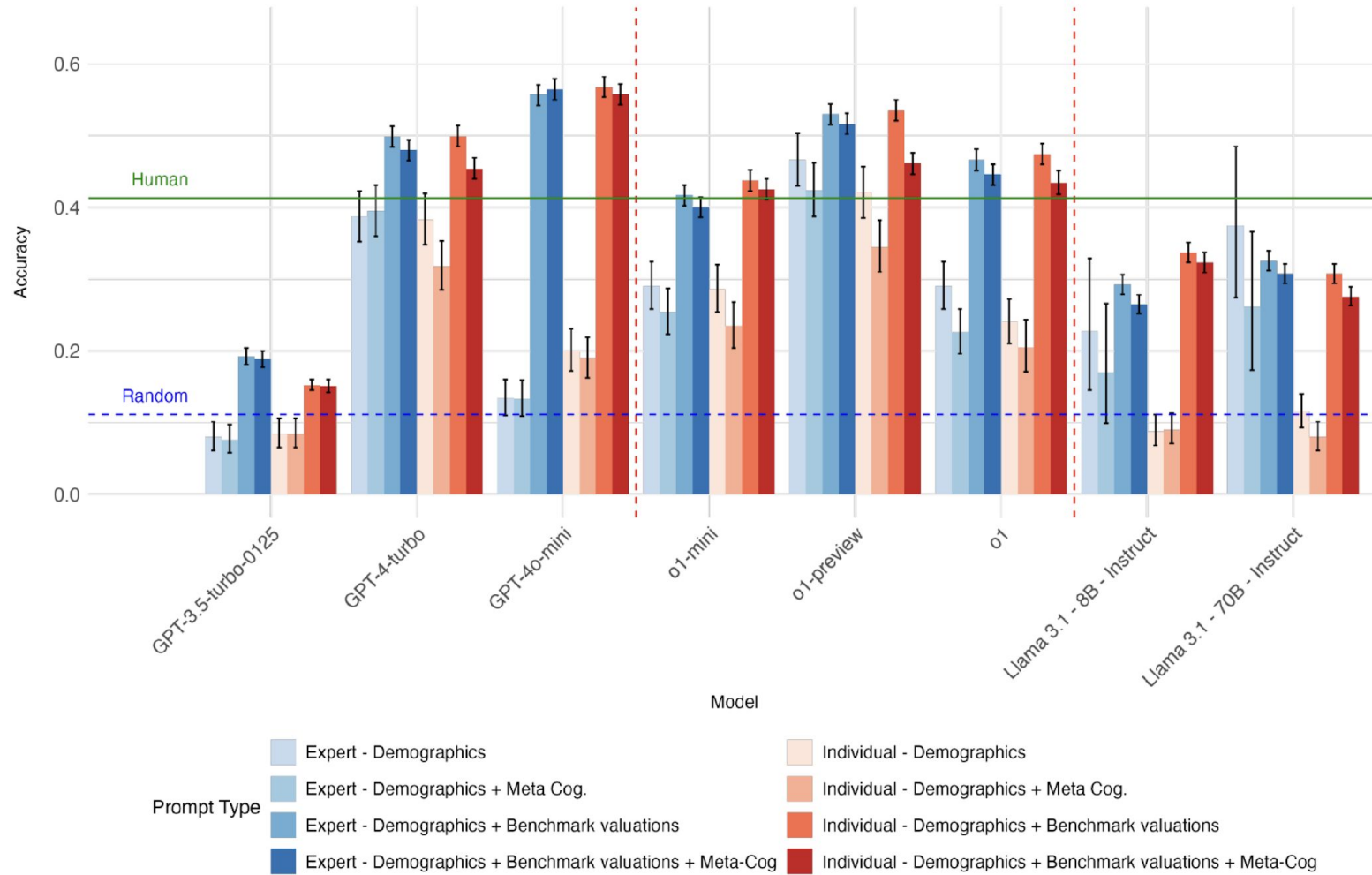
LLMs calibrated on human input would be a first step to:

- Richer high-frequency data at a lower cost through hybrid surveys
- Scaling up current work (e.g., to estimate GDP-B)

→ *“Augmenting Human Survey Responses with GenAI: An Application to Economic Research”*

*by Erik Brynjolfsson, Jose Ramon Enriquez, Sophia Kazinnik, David Nguyen*

# Model Accuracy





**Kris Gulati**

@krisgulati

Why has no one done the llm approach to calculate GDP(-B)? We probably no longer need large scale incentivized discrete choice experiments..

4:42 AM · Aug 15, 2024 · 189 Views

- We have demonstrated that **LLM Agents can consistently achieve good accuracy** in replicating WTA experiments (UK data).
  - Working Paper presented later today and coming out soon (Brynjolfsson, Enriquez, Kazinnik, Nguyen).
- Our **very large dataset of 32,000 respondents** from GDP-B surveys allows us to further train agents for the US.

# An Interesting Application

## The LLM Time Machine

*What if we asked the LLM to give valuations not only for the current year, but for past years?*

# Data

## Humans

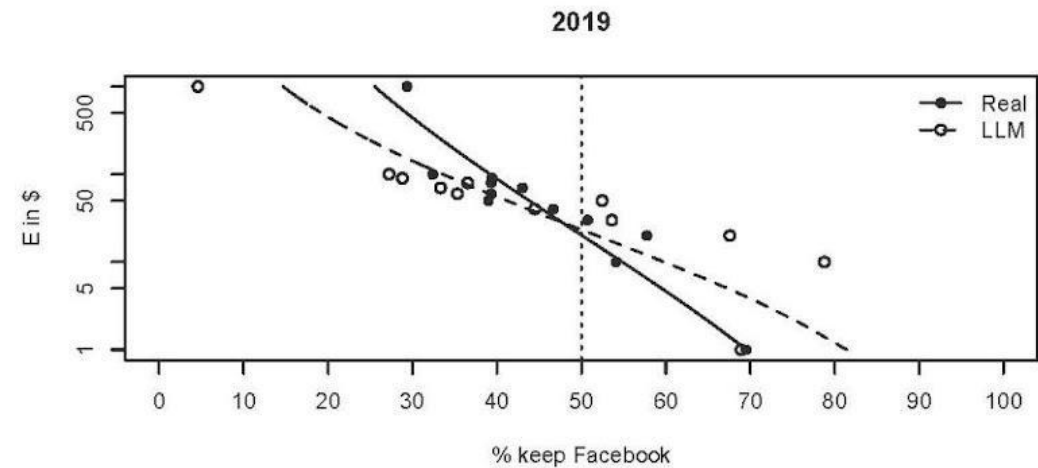
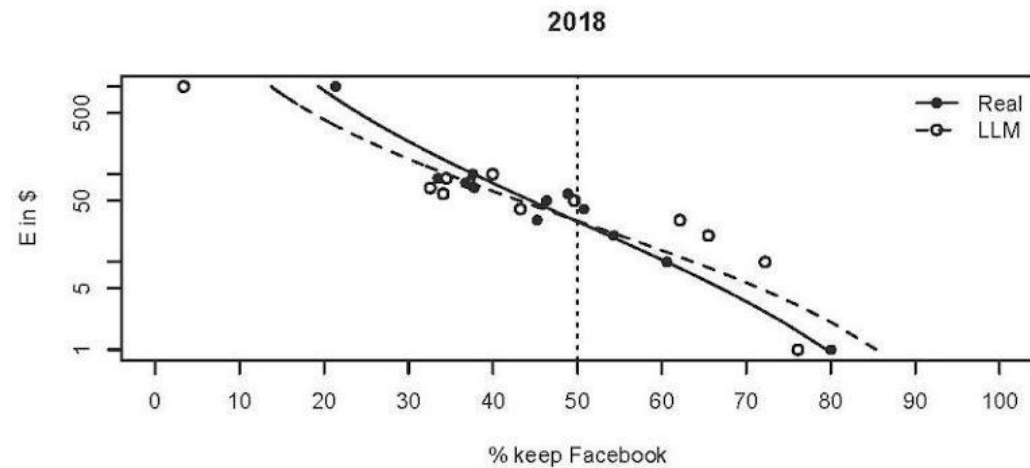
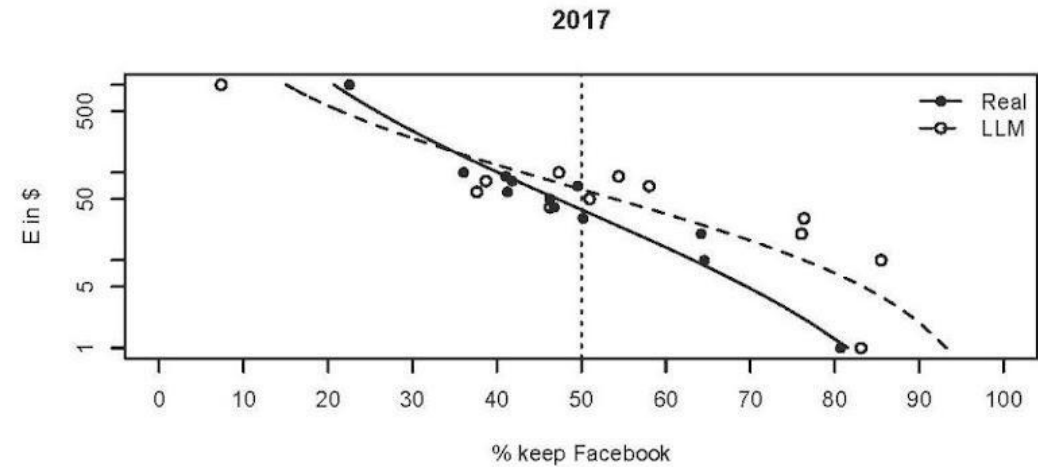
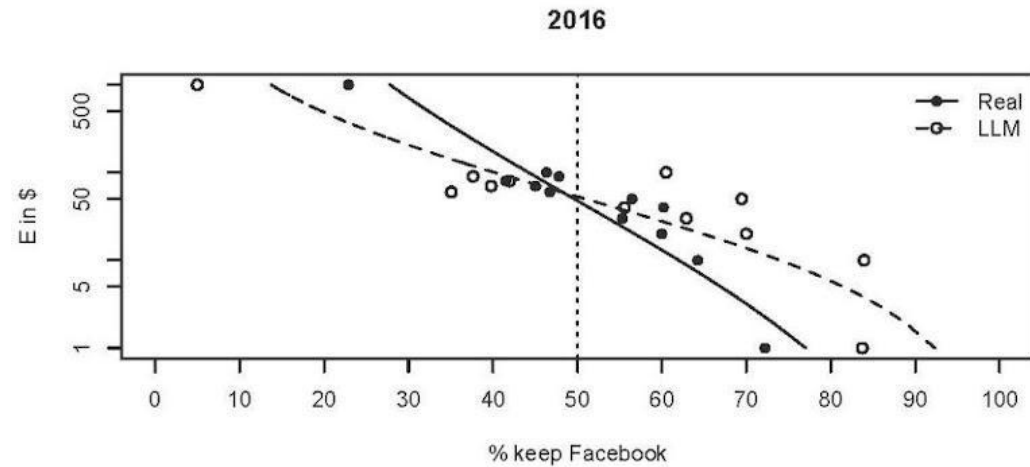
- 8 waves of data collected from 2016-2024 , measuring how much users value Facebook
  - Representative samples of US population recruited on Prodege
  - N = 1,500/yr, 12k total
  - Incentive - compatible design
  - Results from 2016, 17 published previously in our PNAS paper

## LLMs

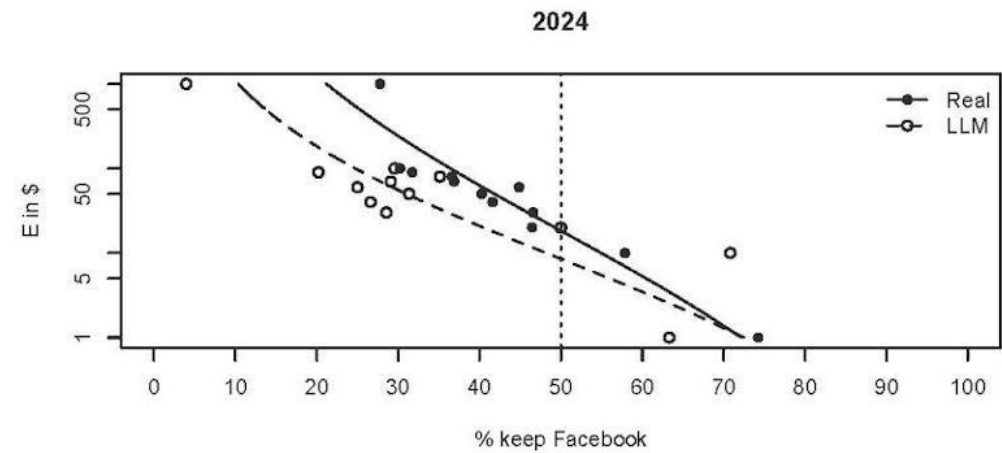
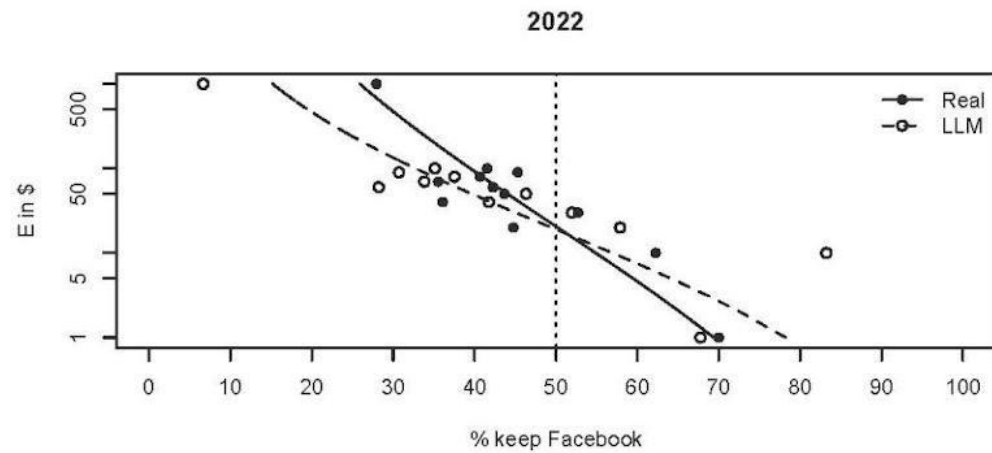
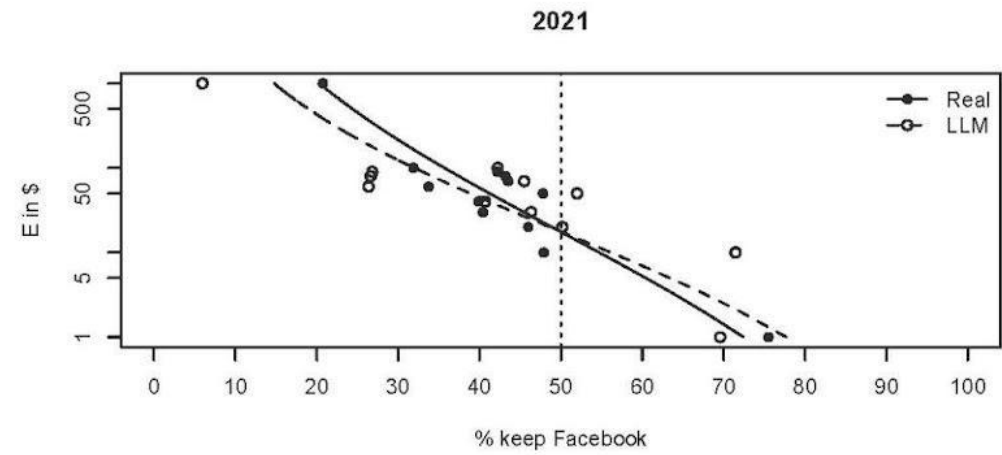
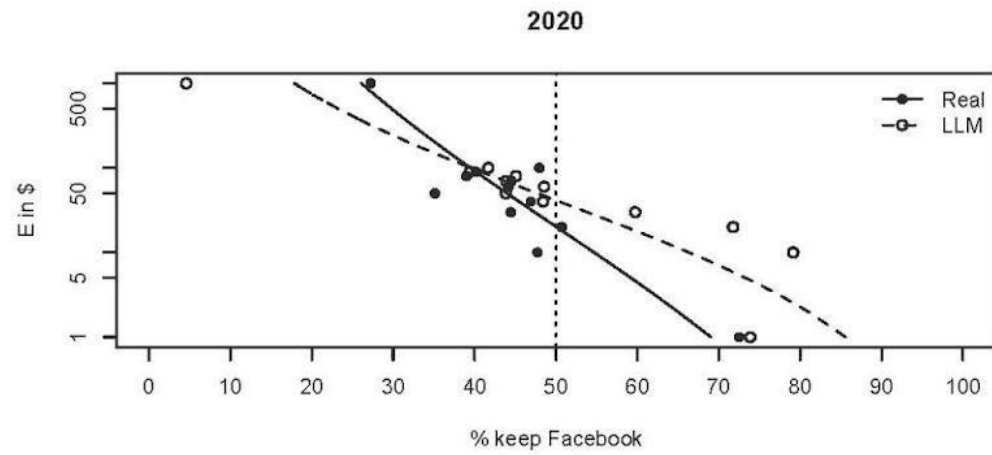
- GPT 4 Turbo, temperature of 1.5
- 5 paragraph prompt (see next slides)
- Define different personas to match socio-demographics of human subjects every year
- Generate decision for each specific persona
  - 0 for not giving up Facebook or 1 for quitting Facebook and accepting the monetary offer.

# Results:

No significant differences in median valuations of Facebook for humans vs. LLMs

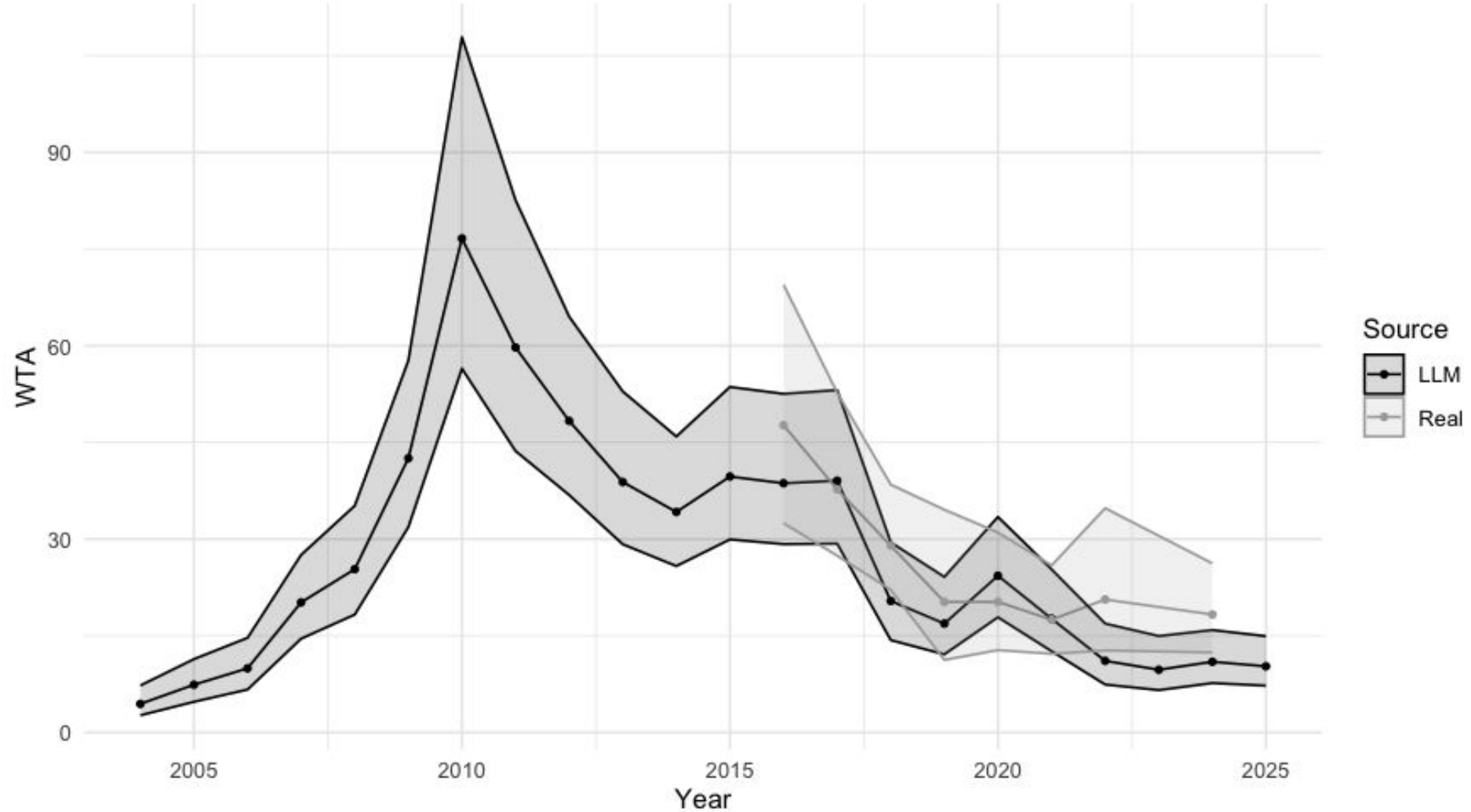


# Results

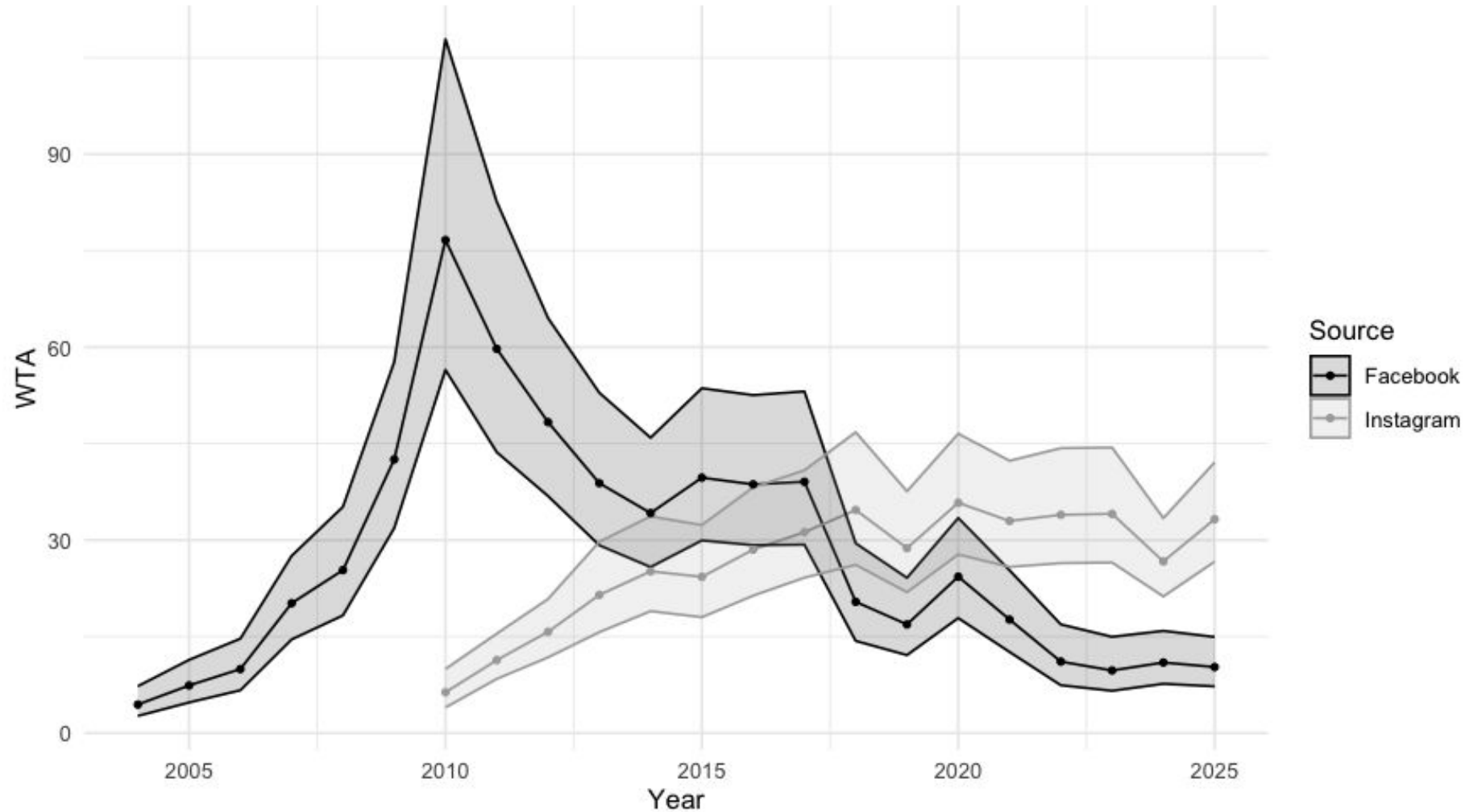


# LLM Time Machine

Can we create longitudinal data going back in time?



# LLM Time Machine: FB vs. Insta



# Summary

- Demonstrate the feasibility of LLMs as synthetic respondents in online choice experiments
  - LLMs seem to do well for estimating median valuations and capture temporal trends
  - Incentive compatibility and random utility theory are important in prompting
- Generate longitudinal data going back into the past, even if human data was never collected
  - What we are more interested in is changes in valuation over time, rather than the absolute valuations today
- Useful not just for digital goods, but also for other non-market goods that lack data on pricing

# ***5. Externalities, Limitations, Extensions***

# Taxonomy of externalities

We built a structured taxonomy of the hidden costs users and society incur when using "free" digital goods.

We classify these costs along two dimensions:

- 1) direct (individual) versus indirect (societal);
- 2) short-term versus long-term

# Taxonomy of externalities

Table 1: Direct Costs: Major Direct Costs of Digital Goods

Category	Brief Description	Code
Attention & Time	Excessive screen time, reduced productivity	D0
Addiction & Mental Health	Increased dependency, anxiety, depression	D1
Privacy Loss	Personal data exploitation, surveillance, loss of control over personal information	D2
Information Processing	User fatigue, decision paralysis, reduced ability to process large amounts of information	D3
Relationships	Reduced in-person interactions, echo chambers	D4
Physical Health & Safety	Eye strain, distracted driving	D5
Other	A catch-all category	D6

# Taxonomy of externalities

Table 2: Indirect Costs: Major Indirect Costs of Digital Goods

Category	Brief Description	Code
Misinformation	Spread of false or misleading content, undermining trust in media	I0
Political Polarization	Heightened ideological divides, reduced civic discourse, radicalization	I1
Social Cohesion	Erosion of common ground, weaker community bonds	I2
Market Concentration	Dominance by a few large platforms, barriers to entry, monopoly/oligopoly concerns	I3
Environmental	Large data centers, high energy consumption, e-waste	I4
Other	A catch-all category	I5

# Taxonomy of externalities

<b>Horizon</b>	<b>Definition / Notes</b>	<b>Code</b>
Short-term	Impacts felt in the immediate to near future (days to months); often more observable and direct	H0
Long-term	Impacts that accumulate or manifest over extended periods (years); often more systemic or diffuse	H1
Both	Costs or issues that begin quickly and also persist or worsen over time	H2

# Taxonomy of externalities

## Product-specific mapping: Facebook

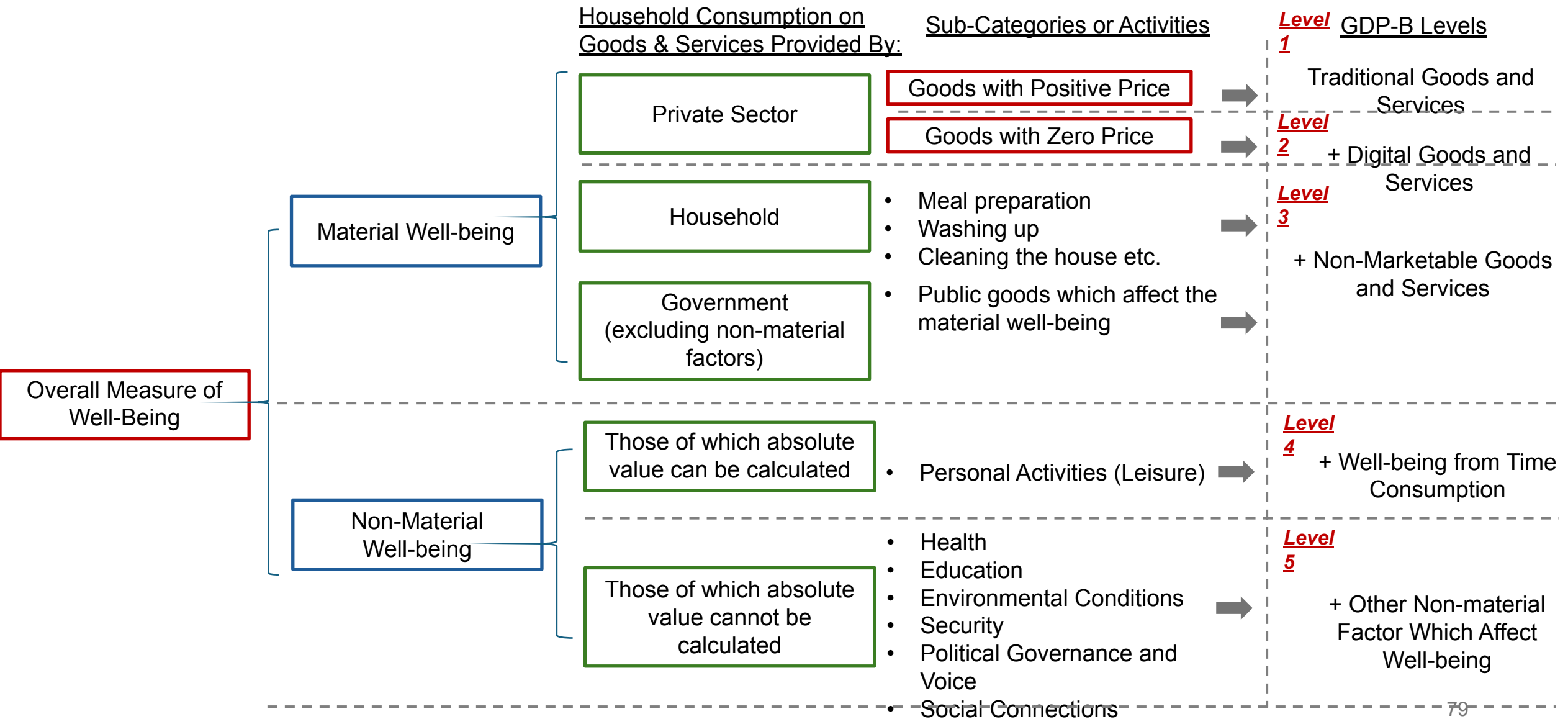
Product	Benefits (GDP-B)	Costs (GDP-C)
Facebook	<ul style="list-style-type: none"><li>• Estimated consumer surplus: \$40–\$50/month</li><li>• Social connection and news updates</li><li>• Entertainment, event planning, personal branding</li></ul>	<ul style="list-style-type: none"><li>• <i>Attention &amp; Time (D0)</i>: Approx. \$10–\$30/month in lost productivity</li><li>• <i>Privacy Loss (D2)</i>: Approx. \$5–\$15/month in data value and psychological costs</li><li>• <i>Misinformation (I0)</i>: Approx. \$1–\$10/month in indirect societal harm</li></ul>

How to estimate costs?

Allcott, Gentzkow & Song (2022) “Digital Addiction”.

Main finding: lack of self-control causes 31% of social media use.

# Taxonomy of contributors (and detractors?) of wellbeing

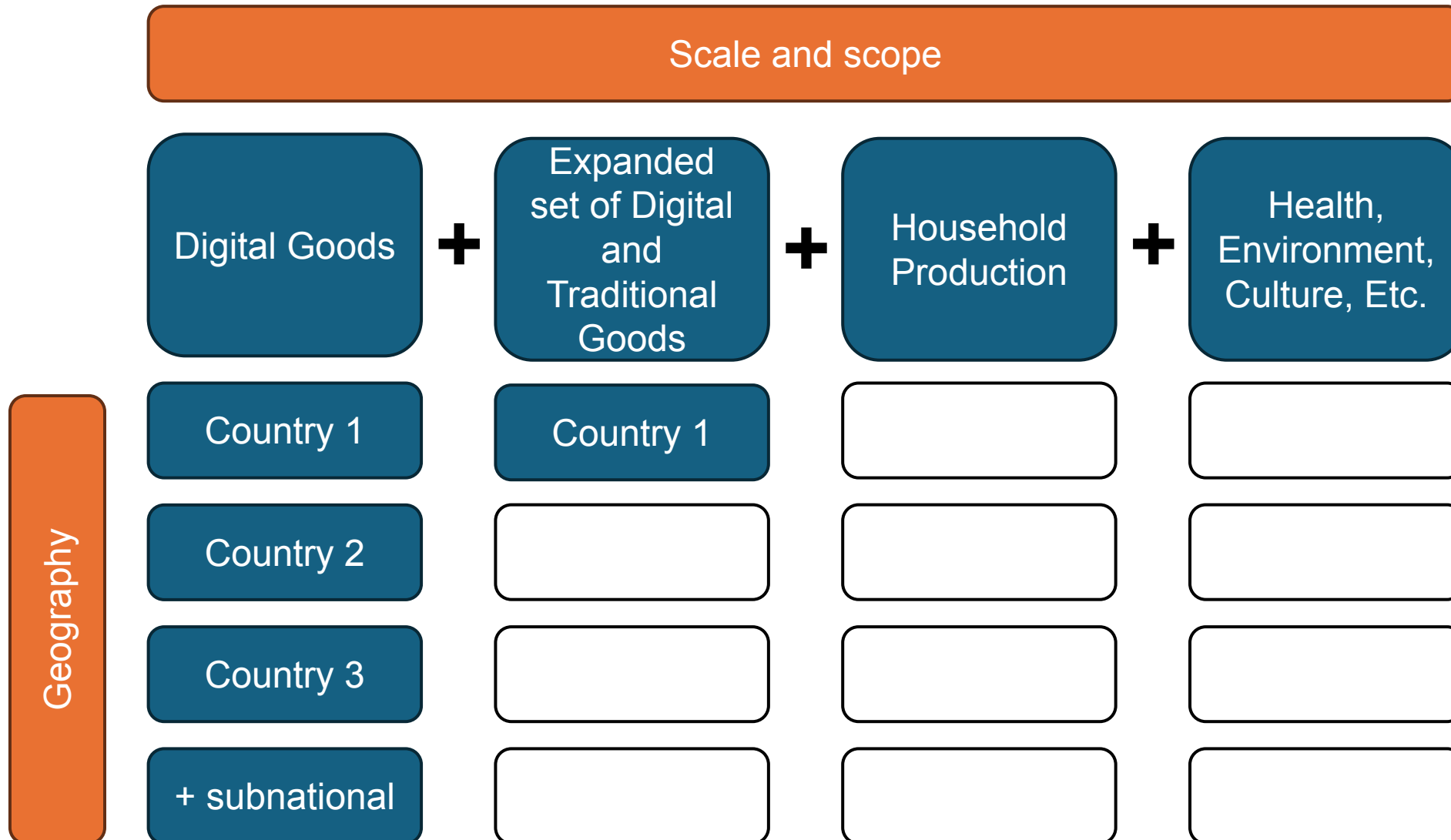


# ***6. Building the infrastructure for future work – call to arms***

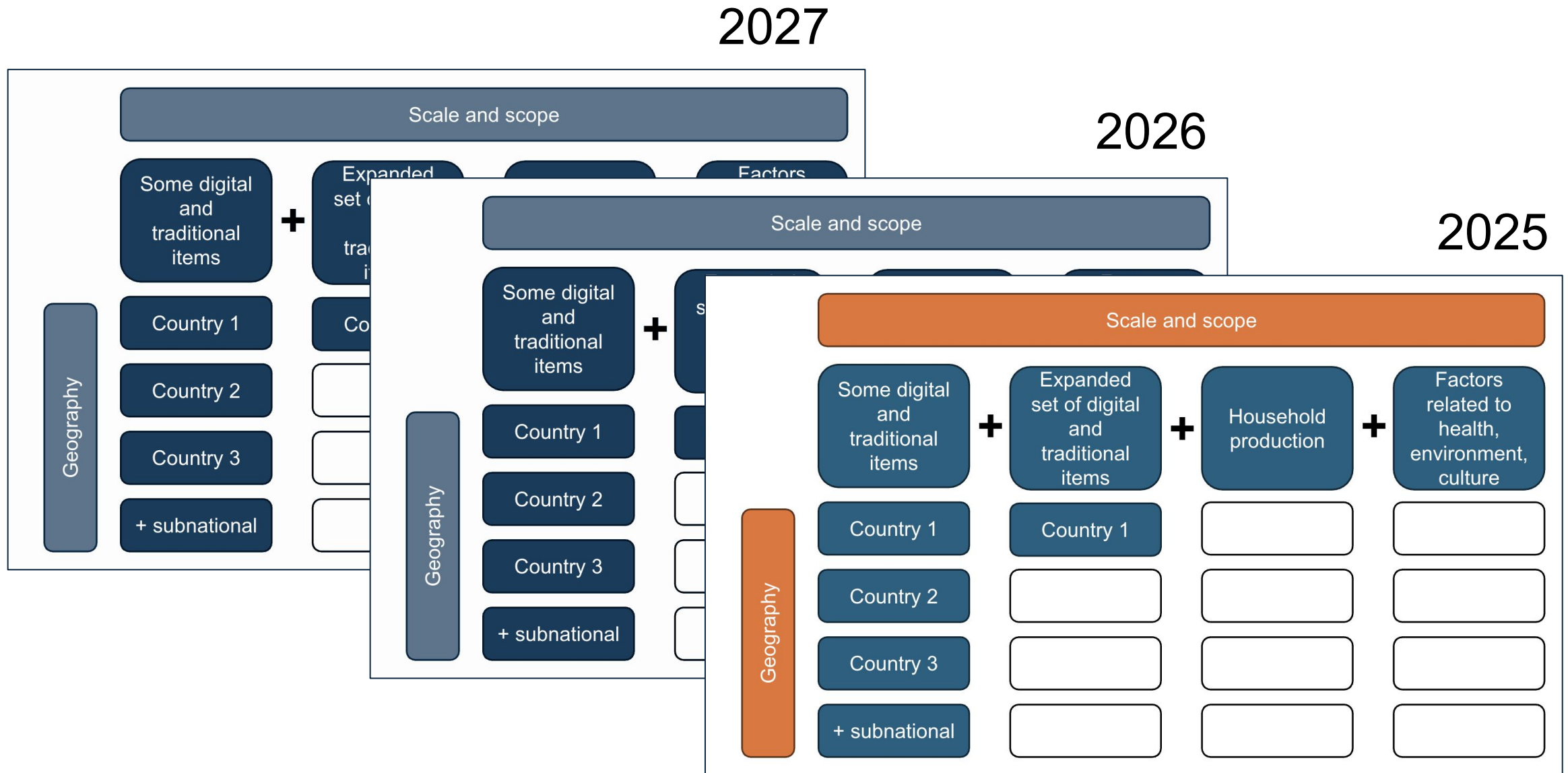
# A Research Agenda for GDP-B

1. Producing and **publishing regular GDP-B numbers**. First annually then, match the quarterly GDP release schedule.
2. **Expanding the number of items** (traditional and digital) that are included. Allows for disentangling substitutability issues.
3. **Adding quality of life factors** by developing research on measuring people's health, environmental assets, household production, culture and social connection.
4. **Explore variation across groups and geography**. Quantify types of survey biases.
5. **Augment surveys with LLM agents** from complementary projects at the lab. This should enhance efficiency, speed, and lower costs of surveying.

# GDP-B next steps: Expanding geography and scale and scope of basket



# GDP-B next steps: A Recurring Time Series (Panel) of Data



# Conclusion: Updating our Measures for the 21st Century

1. **GDP, developed in 1930s, remains the *de facto* metric of production, expenditures and economic growth**
2. **Conceptually, GDP-B can provide a better metric of well-being (benefits), especially for goods with zero price**
3. **Massive online choice experiments have the potential to supplement the measurement of economic welfare.**
  - We have developed methods to do this at scale
  - The digital welfare of nations can be measured
  - LLM Agents can (sometimes) be used for survey data
  - Can be extended to new categories and geographies

***A rich toolkit for improving the measures of the economy!***

# Papers

1. Brynjolfsson, Erik, Avinash Collis, and Felix Eggers. "**Using massive online choice experiments to measure changes in well-being.**" *Proceedings of the National Academy of Sciences* 116.15: 7250-7255, 2019.
2. Brynjolfsson, E., & Collis, A. (2019). **How Should We Measure the Digital Economy**. Harvard Business Review.
3. Brynjolfsson, Erik, Avinash Collis, W. Erwin Diewert, Felix Eggers, and Kevin J. Fox. "**GDP-B: Accounting for the value of new and free goods in the digital economy**". *American Economic Journal: Macroeconomics*. 2025.
4. Brynjolfsson, Erik, et al. "**The digital welfare of nations: New measures of welfare gains and inequality**". NBER Working Paper No. w31670, 2023.
5. Brynjolfsson, Erik, Avinash Collis, Sophia Kazinnik, David Nguyen. "**Measuring Consumer Surplus at Scale.**" Working Paper, 2025.
6. Brynjolfsson, Erik, Sophia Kazinnik, José Ramón Enríquez, and David Nguyen. "**Augmenting Human Survey Responses with Generative AI: An Application to Economic Research**". Working Paper, 2025.
7. Collis, Avinash Felix Eggers, Erik Brynjolfsson. "**LLM Time Machines: Valuing digital goods over time**". Working Paper, 2025

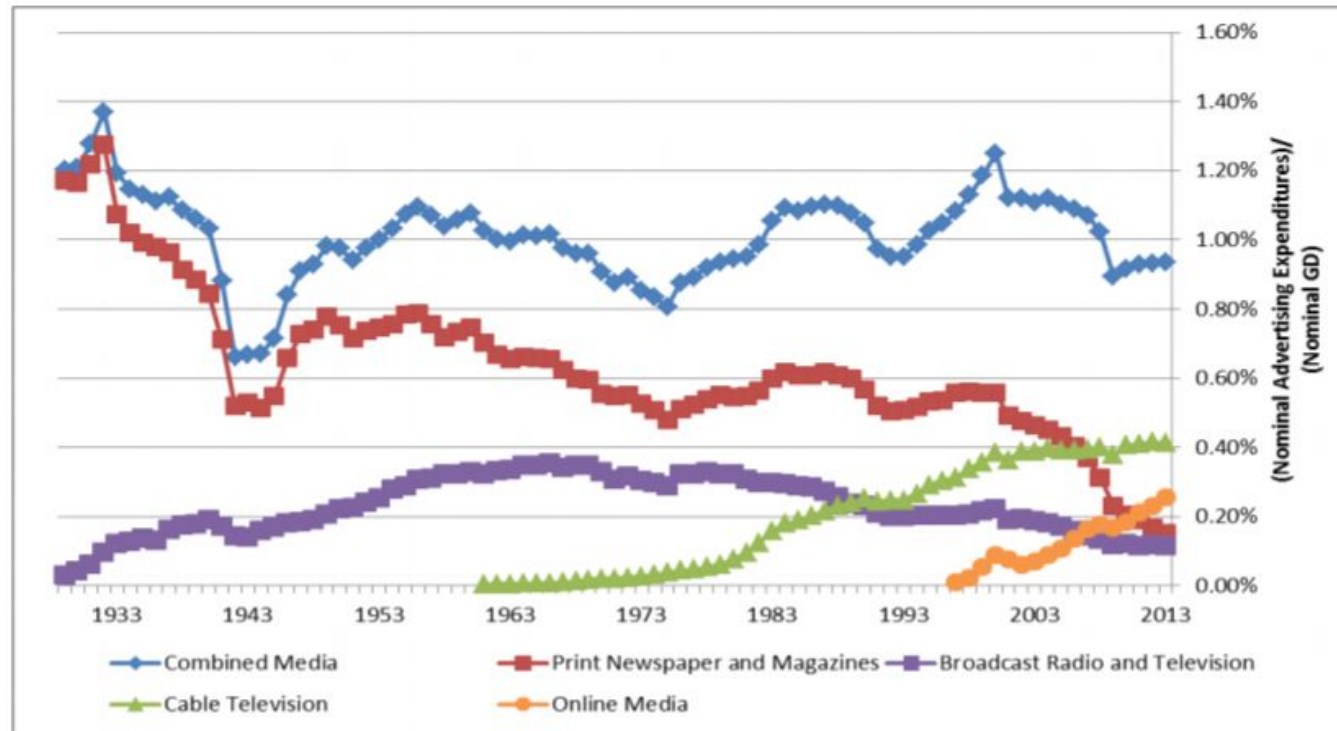
8. **More information:**

<https://www.gdp-b.org/>

<http://digitaleconomy.stanford.edu>

# Some of these goods have ads

Figure 4: Advertising Revenues Over Time



Ref: Nakamura, Samuels and Soloviechik (2017)

**Advertising revenues are generally not proportional to consumer surplus. (*Spence and Owen 1977*)**