# National Accounts in the AI Era

## Measuring the Progress and Diffusion of AI

*Carol Corrado*

Center for Business and Public Policy McDonough School of Business Georgetown University

ESCoE May 21, 2026

# What I'll be talking about

**Framing**

- What is AI?
- Does SNA 2025 solve its measurement?
- National accounts: what's there and what's missing

**Production and Capital**

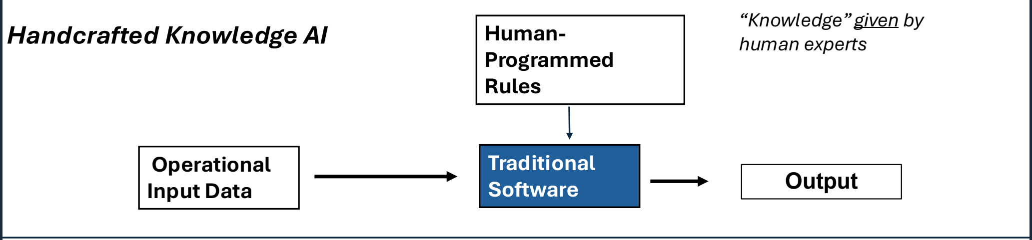
- Two-sector model
- AI production stack
  - Cloud boundary
  - AI prices (approach and proof of concept)
- A measurement agenda for AI (prices and services)

**Consumption**

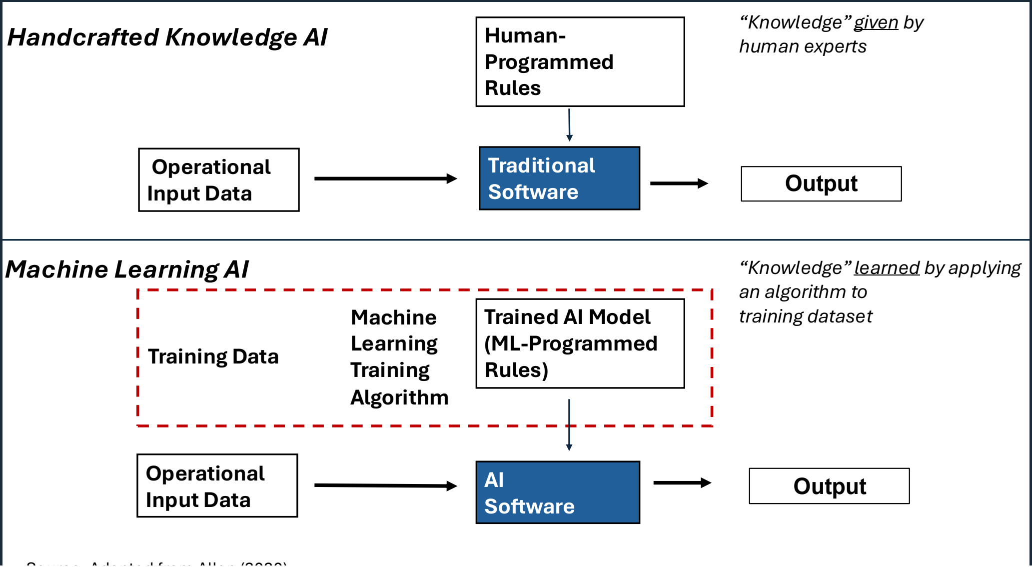
- How much of the annual decline in the price of intelligence reaches consumers?
- Free goods: Taking GDP-B seriously

# Framing

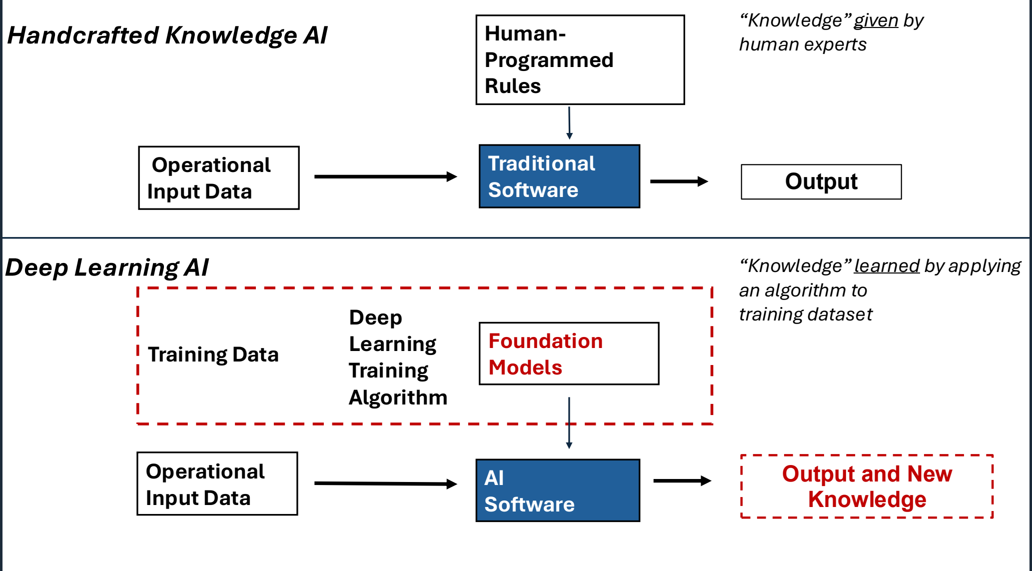
# Handcrafted AI: Expert systems coded by humans



# Machine Learning AI: Shock/Transformation in software production



# AI as innovation in the method of innovation



# The AI measurement problem: does SNA 2025 solve it?

## The prompt

The 2025 SNA — released March 2025, first update since 2008 — addresses cloud computing, AI, and data assets explicitly. Does it solve the AI measurement problem?

## Three sources of AI mismeasurement

1. Cloud asset boundary: cloud expenditures are not intermediate consumption in all sectors of the economy
2. Price index gaps: AI price indexes absent from official statistics
3. Investment concentration: AI spending concentrated in frontier firms; aggregate patterns reflect producers, not broad-based diffusion

### What SNA 2025 does

- Recognises cloud services as a distinct category
- Clarifies that some capacity-building expenditure should be capitalized (e.g. migration costs)
- Strengthens guidance on data assets and own-account software

### What SNA 2025 does not do

- No guidance on API expenditure as shadow capital rental
- No methodology for capability-adjusted AI price indices

# Production and Capital

Based on
“Is Software Eating the World: Measuring the Progress and Diffusion of AI” (April 29, 2026)
and ”AI as an Innovation in the Method of Innovation: Implications for Productivity Growth.” AEA Papers and Procedings, 116: 38-40 (May 2026)
both by
F. Bontadini, C. Corrado, J. Haskel, and C. Jona-Lasinio, April 29, 2026.

# The AI measurement problem — and what to do about it

### Signals are visible...

- Rising capital intensity in software-producing sectors
- Software capital (production and use) explains ∼**half** of the post-2017 acceleration in US labor productivity and total factor productivity growth
- Software industries averaged just **6.4%** of the US NFB economy

### ...but measurement is not

1. Cloud asset boundary
2. Price index gaps
3. Industry Concentration

All three **understate** true AI capital deepening

### What the "Eating the World" paper does

1. Develops a two-sector framework connecting AI production and diffusion to national accounts through two observable diagnostics: factor shares and the user cost of AI capital
2. Formalizes the cloud asset boundary problem and its implications for both diagnostics
3. Constructs a proof-of-concept capability-adjusted price index for frontier AI: $-38\%$/yr decline in rental price of AI capital
4. Sets out a measurement agenda for all AI prices
5. Estimates multi-input Morishima elasticities on EU KLEMS &INTANProd data; simulates factor share consequences of AI-driven price declines

# A two-sector framework for AI

**Upstream sector $S$ (AI producer)**

$$X_S = A_S F_S(K_S^{AI}, K_S^O, L_S)$$

- Produces AI capital: software, trained models, data assets
- Uses AI in its own production $\Rightarrow$ recursive
- Superior TFP $\dot{a}_S$ drives $P_S$ down

**Downstream sector $C$ (AI user)**

$$Y_C = A_C F_C(K_C^{AI}, K_C^O, L_C, Z_C)$$

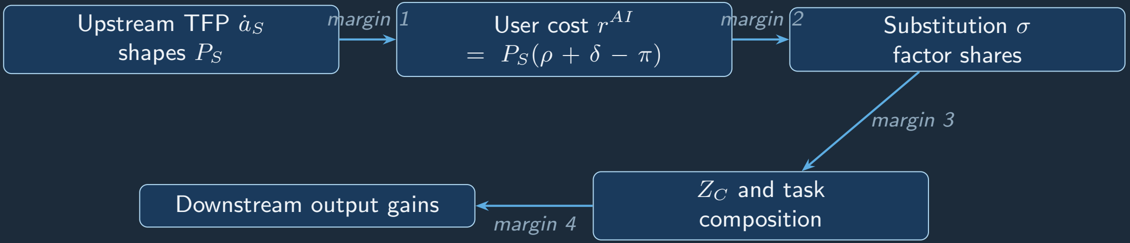
- Uses AI capital as an input
- $Z_C$: organizational capital that must co-evolve
- Falling $r_C^{AI}$ $\Rightarrow$ substitution toward AI capital, governed by $\sigma$

**Two distinct substitutions**

1. **Across the economy**: AI replaces/augments labor in downstream production — the **diffusion** channel
2. **Within sector $S$**: AI used to produce AI — the **recursion** channel

Both show up in aggregate capital shares but are conceptually distinct. Upstream TFP $\dot{a}_S$ and downstream TFP $\dot{a}_C$ have different determinants and different policy implications.

# Two diagnostics, four linked margins




**Diagnostic 1 — Factor shares**

$\ln(s_K^{AI}/s_L)_j = \text{const}_j + (\sigma_j - 1)\ln(w/r^{AI})_j$

Observable signal of AI diffusion in sector $C$

**Diagnostic 2 — AI capital costs**

Rate of decline of quality-adjusted relative $P_S$

Signal of upstream productivity advantage; $\pi \approx \dot{a}_S - \dot{a}_C$

# Constraints on downstream output gains

- Observed shares capture net substitution *after* investment in $Z_C$
- Output bounded by tasks not yet automated ("weak links" constraint)
- Coordination of production is a **meta-task**: new coordination methods (e.g., Amazon's delivery system) a driver of $\dot{a}_C$

High $\sigma$ and falling $r^{AI}$ are necessary but not sufficient.

**~~Three~~ Four measurement problems**

1. Cloud asset boundary
2. Price index gaps
3. Concentration

All four bias the two diagnostics **downward**

4. Intangible asset boundary — Investment in $Z_C$ not included in national accounts

# The AI production stack

| Layer | Definition · Representative vendors | Revenue (bn) 2025 | Revenue (bn) 2035E |
|---|---|---|---|
| **L1 Applications** | AI-native software: consumer chatbots, enterprise copilots, coding assistants, vertical AI (health, legal, finance). *e.g. GitHub Copilot · ChatGPT/Claude.ai · Microsoft 365 Copilot* | **$19** | $650 |
| **L2 Foundation Models** | API access to large language and foundation models, sold to developers and enterprises. *e.g. OpenAI API · Anthropic API · Gemini API · Mistral · Cohere* | **$20** | $470 |
| **L3 Platforms & Tooling** | MLOps, vector databases, model orchestration, fine-tuning infrastructure. *e.g. Pinecone · Weights & Biases · LangChain · Hugging Face* | **$6.4** | $77 |
| **L4 Hyperscale Cloud AI** | AI-attributable cloud revenue (AWS, Azure, GCP) plus neocloud GPU-as-a-service. *e.g. Azure OpenAI · Vertex AI · CoreWeave · Lambda* | **$27** | $715 |
| **L1–L4 total** | **(software and services)** | **$72** | ∼$1.9T |
| **L5 Compute** | AI chip revenue: NVIDIA, AMD, and custom silicon (memo item). *e.g. NVIDIA H100/B200 · AMD MI300X · Google TPU · AWS Trainium* | **$180** | ∼$1.1T |

## Why this matters

- **$72bn** frontier model market today; **$1.9T** by 2035 — a market that barely existed in 2020
- Asset boundary: L4 cloud spend is expensed as intermediate consumption, not capitalised ⇒ capital deepening understated
- Price index: L2 API prices are the observable services price counterpart to declining $P_S$ — the basis for our index

# The cloud asset boundary: shadow capital and factor shares

## Shadow reclassification (Type 2a)

$$M_C^{AI} = r^{AI} \cdot K_{shadow}^{AI}$$

The cloud expenditure of sector-$C$ firms *is* the rental payment for an unrecorded capital stock

## Corrected factor shares

$$\left(\frac{s_K^{AI}}{s_L}\right)_C^{true} = \frac{\Pi_C^{AI,obs} + M_C^{AI}}{w_C L_C}$$

$$\left(\frac{s_K^{AI}}{s_L}\right)_S^{obs} = \frac{\Pi_S^{AI,true} + M_C^{AI}}{w_S L_S}$$

**Mirror-image errors**: $M_C^{AI}$ missing from sector $C$, spuriously present in sector $S$

| Effect | Direction |
|---|---|
| AI-capital income in sector $C$ | Understated |
| Operating surplus in sector $S$ | Overstated |
| Rental rate $r^{AI}$ | Overstated |
| Elasticity $\hat{\sigma}_C$ | Biased down |
| TFP in AI-using industries | Biased down† |

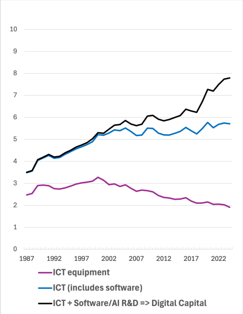
†During periods of rapid expansion of cloud AI expenditure, when the output channel dominates.

### Aggregate vs sectoral

Errors roughly cancel in aggregate (capital income shifts $S \rightarrow C$, not lost). But **aggregate stability is correct for the wrong reason** — it masks the location of AI capital use. Breaks down if sector $S$ earns rents.

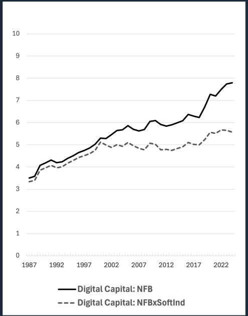
# Descriptive evidence: digital capital factor shares, US NFB sector

## Nonfarm Business Sector Digital Asset Income Shares, 1987–2024



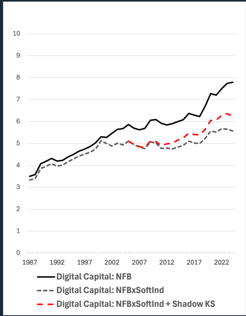

**Panel 1: NFB aggregate**

Digital capital share rising sharply from 2017. ICT equipment share declining; software & AI R&D rising.




**Panel 2: Software industries vs. rest of NFB**

Gap between NFB and NFB ex. software industries shows concentration in software-producing sector.




**Panel 3: Shadow cloud adjustment**

Adding shadow cloud capital services narrows the gap — diffusion is further advanced than official data imply.

**Key finding**

Concentration is real, but diffusion is further advanced than official data show

**How to inform policy?**

Requires disciplined empirical estimates of $\sigma$

$\Rightarrow$ AI services satellite account

# Measuring AI capital prices: Why existing methods fail

**The IT era approach (no longer applicable)**

- Hardware and software priced separately
- Semiconductor price indices: transistor counts, clock speed
- Software deflators: input costs or matched-model
- "What Intel giveth, Microsoft taketh away" — separability was a genuine feature of the technology

**Why AI breaks separability**

- Capability gains emerge from **interaction** of all stack layers: architecture (software, data), training scale (compute)
- No single layer can be priced independently
- Improvements are often **discontinuous** — architectural breakthroughs, not smooth attribute progress
- Hedonic methods require stable measurable characteristics: AI systems fail this requirement

# Measuring AI capital prices: An integrated approach

## The right unit of measurement

The integrated system — **the deployed model** — rather than any individual input. The index should measure:

"What the frontier model delivereth"

Cost per unit of productive AI output, regardless of which stack layer generated the gain

## Connection to user cost framework

**Services price, not acquisition price**: sector-$C$ firms observe only a services price (API / cloud). Under competitive equilibrium this is a valid proxy for the declining $P_S$ that drives $r^{AI}$.

**Shadow capital link**: the API/cloud expenditure underlying the index is precisely $r^{AI} \cdot K^{AI}_{shadow}$ — the rental payments for AI capital services downstream firms do not own

# The capability-adjusted price index: construction

**Data**: 21 model-release observations, Mar 2023 – Mar 2026
Base: GPT-4 launch, Mar 2023 = 100

**Token price index**
Blended price: $p_t = 0.5p_t^{in} + 0.5p_t^{out}$ per million tokens
Token price ≈ unit value index of AI capital services price

**Capability-adjusted index**
Deflate by Epoch Capabilities Index (ECI):

$$\text{Cap-adj}_t = \frac{p_t/\text{ECI}_t}{p_{base}/\text{ECI}_{base}} \times 100$$

ECI: Item Response Theory composite across reasoning, coding, knowledge benchmarks; acceleration from 8 to **15 pts/yr** post-Apr 2024

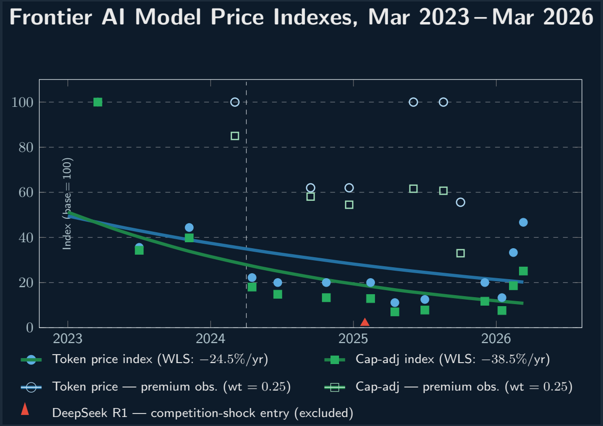
**Observation classification**

| Type | $w$ | $n$ |
|---|---|---|
| Mainstream frontier | 1.0 | 13 |
| Premium launch spike | 0.25 | 6 |
| Competition shock | excl. | 1 |
| Pro-tier outlier | excl. | 1 |

**Proof of concept**

Approach designed to be extensible as panel accumulates. Could be operationalized as a satellite price index for AI capital services within the national accounts framework.

# Results




## Token price index

**$-24.5$ percent per year**

$(p = 0.096,\ R^2 = 0.15)$

Wide CI reflects sawtooth pattern of mainstream/premium releases

## Capability-adjusted index

**$-38.5$ percent per year**

$(p = 0.008,\ R^2 = 0.35)$

Statistically significant; robust to exclusions

## Quality-adjustment gap

**$\approx 14.5$ percent per year**

No counterpart in official statistics

# The proof-of-concept in context: the full software market

| Segment | 2025 | 2035E |
|---|---|---|
| AI stack software (L1–L4) | 6% | 79% |
| AI pricing uplift* | 9% | 2% |
| Agentic AI* | 2% | 20% |
| Pure legacy (incl. trad. ML) | 83% | 0% |
| **Total market** | $1.24T → $2.4T | |

*AI pricing uplift and agentic AI refer to AI-featured and AI-agent tiers of legacy enterprise platforms.

AI segments: **17%** of market today → **100%** by 2035. The window to develop methods before distortions reach macroeconomic scale is **narrow**.

## Where the proof-of-concept fits

The capability-adjusted index covers AI stack software (L1–L4) — today the smallest AI segment by revenue, but the fastest-growing and the one with the largest quality-adjustment gap. No counterpart exists in official statistics.

## What remains

Each other segment has its own measurement challenge:

- Pricing uplift: AI premium tractable, but intra-tier quality improvement unmeasured
- Agentic AI: new pricing units (per task/outcome); methodology unsettled
- Legacy: official deflators adequate as lower bound

# Estimation results and factor share simulations

## Morishima elasticities of substitution

All $< 1$: firms substitute toward cheaper inputs but not enough to raise that input's cost share

| Row \ Col | ICT | Tang. | Intang. | Labor |
|---|---|---|---|---|
| ICT | · | 0.54 | 0.89 | 0.39 |
| Tang. non-ICT | 0.61 | · | 0.82 | 0.38 |
| Intang. excl. SW | 0.65 | 0.61 | · | **0.55** |
| Labor | 0.59 | 0.46 | **0.75** | · |

- Highest: intangibles $\leftrightarrow$ labor (**0.55–0.75**)
- Lowest: tangible non-ICT $\leftrightarrow$ labor (0.38–0.46)
- **Weak-links**: no single input cheaply replaces another

## Three simulation scenarios

**Scenario 1 — Baseline (observed prices)**

Labor share drifts down; intangibles rise. Reflects **biased technical change** ($\hat{\beta}_i$), not price effects.

**Scenario 2 — ICT prices $-10\%$/yr**

ICT share $\rightarrow$ near zero in 30 periods. **Labor share rises**: $M_{ij} < 1$ throughout — rival input's share increases when an input cheapens.

**Scenario 3 — ICT $-20\%$/yr & intangibles $-10\%$/yr**

Both input shares fall. **Labor share rises more substantially**: joint cheapening reinforces the labor-share result.

Estimates based on pre-AI ICT data — **ICT simulations, not AI forecasts**. Translog parameters are reduced-form, not structural primitives.

# Tasks, bottlenecks, and the limits of estimation

### Two-level production structure

- Within-task ($\sigma_F$): capital vs. labor *inside* each task
  e.g. human dancers vs. AI-generated performance
- Across-task ($\sigma_T$): substituting one task for another
  e.g. automated coding replacing human programmers
- Without task-level data, cost function recovers a **combination** of $\sigma_F$ and $\sigma_T$

### The bottlenecks special case

If $\sigma_F = \infty$, each task belongs to one factor; cost function recovers $\sigma_T$ **alone**. Even if capital displaces labor within every task, output growth still depends on substitution *across* tasks.

### Coordinating tasks is itself a task

Ideally $Z_C$ represents the **coordination** task. Its accumulation is a precondition for substitution gains to reach output.

### Overall interpretation

Results support both conjectures: Factor shares are governed by limited substitution *across* tasks, not within them. Estimates embed existing **meta-technology of coordinating tasks**. If AI shifts Morishima elasticities above unity, labor-friendly results could reverse.

Static elasticity estimates cannot capture structural change in the production possibility set — which is precisely why the **measurement agenda** remains indispensable.

# The measurement agenda: four priorities (1 of 2)

1. **Asset boundary reclassification (Type 2a / 2b)**

   Requires statistical agencies to decompose cloud service fees by economic use — capacity-building vs. current production vs. own-account investment. Feasible by combining industry information with targeted additions to business investment surveys. **Highest near-term payoff.**

2. **Capability-adjusted price index for AI stack software (L1–L4)**

   The ECI-based approach developed here is a proof of concept. Operationalization requires a public-private partnership with benchmark index providers (Epoch AI, METR) and NSOs. Compilation into a satellite price index program for AI capital services. L1 has growing consumer segment: Tie consumer AI price to the AI stack price, to be discussed momentarily

# The measurement agenda: four priorities (2 of 2)

3. **Price index for the AI pricing uplift segment**

   Compare AI-featured and base tiers of the same platform (Microsoft 365, Salesforce, SAP) at successive points in time. Current approach Identifies the AI premium without a capability adjustment. But the AI tier itself improves over time, so current indexes are a **lower bound**: can captures price compression, not intra-tier quality improvement. Data to close that gap do not yet exist.

4. **Task-completion index for agentic AI**

   **Leading edge of the agenda**. METR time-horizon and SWE-bench are candidate capability benchmarks. The unit of output — task, conversation, or outcome — must be defined before index methodology can be settled.

# Consumption

# Consumer AI services price:

- L1 layer of production stack has growing consumer segment $\Rightarrow$ tie consumer price to the AI stack price.
- How much of the consumer price reflects genuine non-model costs versus rent capture?
- Need consumer prices to evaluate forthcoming results on consumers' valuation of AI vis GDP-B project.

# Consumer AI Price Index: Approach

## Motivation

Upstream: API prices for frontier AI fall sharply ($-38.5\%$/**yr** capability-adjusted).

Do consumers benefit? Existing CPI does not include AI services (yet). We ask: what would a prototype consumer AI price index show?

## Conceptual chain

API price declines
↓
Consumer-facing firms' costs fall
↓
Pricing response ($\lambda$)
↓
Consumer unit price

$\lambda$ reflects competition, margins, and other costs of consumer-facing firms.

## Unit price concept

Define a *standardised AI service unit*:

$$Q_j = U_j \times I_j \times d \times 30$$

$U_j$ = paid subscribers, $I_j$ = interaction intensity, $d$ = prompts/day.

$$P_j^{\text{nom}} = \frac{\text{Revenue}_j}{Q_j} = \frac{\text{Price}_j}{I_j \cdot d \cdot 30}$$

Subscriber counts cancel: unit price depends only on *list price*, *intensity*, and *utilization*.

## Capability adjustment

$$P_j^{\text{cap}} = \frac{P_j^{\text{nom}}}{\text{ECI}_j/\text{ECI}_0}$$

Deflate by the Epoch Capabilities Index.

**Gap** $= \Delta P^{\text{nom}} - \Delta P^{\text{cap}}$ = unmeasured decline in CPI for AI services from quality change.

# Consumer AI Price Index: Implementation

## Panel dataset: 23 tier-observations

Three providers × multiple tiers × time, 2023 Q1 – 2026 Q2.

| Provider | Tier | Price | Class |
|---|---|---|---|
| OpenAI | Plus | \$20/mo | S |
| OpenAI | Go | \$8/mo | B |
| OpenAI | Pro \$100 | \$100/mo | P |
| OpenAI | Pro \$200 | \$200/mo | P |
| Anthropic | Pro | \$20/mo | S |
| Anthropic | Max \$100 | \$100/mo | P |
| Anthropic | Max \$200 | \$200/mo | P |
| Google | AI Pro | \$20/mo | S |
| Google | AI Ultra | \$250/mo | U |

S = Standard (wt 1.0) P = Premium (wt 0.25)
B = Budget (wt 0.25) U = Ultra (excl. from trend)

## Utilization scenarios

| Assumption | Light | Central | Heavy |
|---|---|---|---|
| Prompts/day (paid user) | 2 | 4 | 8 |
| OpenAI intensity (base) | 1.0 | 1.0 | 1.0 |
| Claude intensity premium | +25% | +50% | +100% |
| Gemini relative intensity | 50% | 70% | 90% |

## WLS regression — same method as upstream

Regress $\log(\text{index})$ on time, weighting observations by class weight. Base period: ChatGPT Plus, Feb 2023.

Two series estimated in parallel:

- Nominal unit price index
- Capability-adjusted unit price index

BUDGET and ULTRA observations excluded from trend regression; annotated as in upstream paper.

# Consumer AI Price Index: Results

## WLS trend estimates (central scenario)

| Series | Trend | $p$ | $R^2$ |
|---|---|---|---|
| Nominal unit price (Standard only) | $+2.9\%$/yr | 0.687 | 0.015 |
| Cap-adj unit price (Standard only) | $-13.5\%$/yr | 0.084 | 0.247 |
| **Gap** | $-17.8$ **pp/yr** | | |

Nominal price **flat to slightly rising** — subscription prices held at \$20/month while capability improved rapidly.

Cap-adjusted price **falls** $-13.5\%$**/yr** (p = 0.084), significant at 10%.

$n$ = 13 Standard obs; ECI base = GPT-4 Mar 2023 = 85. Utilization: central scenario (4 prompts/day).

## Comparison with upstream output price

| Series | API (%/yr) | CPI (%/yr) |
|---|---|---|
| Nominal price | $-24.5$ | $+2.9$ |
| Capability-adjusted | $-38.5$ | $-13.5$ |
| **Gap (pp/yr)** | $-14.1$ | $-17.8$ |

## Interpretation

- **Pass-through near zero**: API nominal price declines are *not* reaching consumer prices.
- **Quality improvement pass-through real**: consumers receive better AI at flat nominal prices.
- Consumer-side gap ($-17.8$ pp/yr) *exceeds* upstream gap ($-14.1$ pp/yr) — CPI mismeasurement may be larger than the producer-side distortion.

# Consumer AI prices and GDP-B

- **GDP-B should be taken seriously.** Byrne and Corrado (2021) showed that quality-adjusted prices for consumer digital services yielded a consumer surplus estimate that closely matched the value obtained from Brynjolfsson et al.'s stated-preference approach.
- The capability-adjusted consumer AI price index constructed here extends that logic to AI services: the $-13.5\%$**/yr** cap-adjusted decline implies rapid growth in consumer surplus that GDP and CPI both miss.
- Together, the upstream and downstream gaps — $-14.1$ and $-17.8$ pp/yr respectively — point to a consistent direction of bias: national accounts **understate** the welfare contribution of AI.

Byrne and Corrado (2021): "Accounting for Innovations in Consumer Digital Services: IT Still Matters," in Corrado, Haskel, Miranda, and Sichel (eds.), *Measuring and Accounting for Innovation in the Twenty-First Century*, NBER Studies in Income and Wealth vol. 78, University of Chicago Press, pp. 471–517.

# Thank you

Carol Corrado

carol.corrado@georgetown.edu