# The Cantor Institute

**Institutional Policy Response Department**
(1) Siddharth Mahajan*
Mihir Gupta

[redacted]

**FEDERAL TRADE COMMISSION**

***Agency Information Collection Activities; Proposed OMB Generic Clearance; Comment Request; Generic Clearance for Information Collection Using Voluntary Surveys for Studies Conducted by the Federal Trade Commission Bureau of Economics To Support the FTC's Missions To Protect Consumers and Competition***

**Consultation ID: FTC-2025-0132-0001**

---

*Prepared by:* ***Mihir Gupta*** *and* ***Sid Mahajan***
*Dated: 19 August 2025 (America/LosAngeles)*

---

## Executive summary

The Federal Trade Commission (FTC) seeks a generic clearance to run ~10 voluntary social/behavioral studies per year with **20,000 total interviews/surveys (30 minutes each; 10,000 respondent hours)** to support consumer protection and competition missions. We agree the evidence need is real, but the current notice understates respondent burden, leaves statistical performance goals unspecified, and lacks commitments that safeguard data quality, replicability, and privacy at scale.

**Key findings (quantitative):**

- **Burden likely understated by 5–38%** once screening/partial completes are included. With realistic eligibility and completion rates, total public burden rises from **10,000 hours** to **10,500–13,800 hours** (details below).

- Given 10 studies/year (≈2,000 completes/study), typical two-arm A/B tests will have **minimum detectable effects (MDEs) ≈ 6.3 percentage points** for binary outcomes at 80% power, α=0.05 (p=0.5 worst case). With 5 arms (≈400/arm), MDEs inflate to **≈10 p.p.**; weights or clustering further increase MDEs by √DEFF.

- With multiple outcomes (e.g., 10), familywise error control (Bonferroni) raises α-adjusted critical values, enlarging MDEs by **~30%**, threatening actionable precision absent larger samples or pre-specified primary endpoints.

- Using a conservative **value-of-time range $35–$75/hour**, the direct respondent-time cost of the program is **$350k–$750k/year** as stated; with screening burden included the implied range is **≈$368k–$1.04M/year**.

**Core recommendations:**

1. **Publish per-study Statistical Analysis Plans (SAPs)** with **a priori power/MDE targets**, primary outcomes, and multiple-comparison control.

2. **Commit to effective sample sizes** (accounting for weights/design effects) sufficient to meet MDE goals for primary outcomes.

3. **Quantify and disclose screening and partial-complete burden**; update the annual hour estimate when realized screening rates exceed assumptions.

4. **Adopt minimum standards** for sampling frames, response-rate reporting (AAPOR), weighting, nonresponse bias evaluation, instrument pretesting, and preregistration.

5. **Strengthen privacy** (data minimization, limited retention, public technical docs, release of de-identified microdata when feasible with formal disclosure risk assessment).

6. **Build an outcomes dashboard** reporting realized MDEs, precision, and enforcement-relevant insights per study.

---

# 1) Context and what the notice proposes

The FTC proposes a **generic clearance** to run voluntary interviews, focus groups, surveys, and experiments to understand consumer perceptions/behaviors and support enforcement and education; it estimates **20,000 responses/year**, **30 minutes/respondent** (10,000 hours), and seeks 60-day comments by **September 8, 2025**. The notice clarifies the work supports law-enforcement investigations and program improvement rather than direct policymaking.

---

# 2) Quantitative burden analysis

### 2.1 Reported burden

- **20,000 completes × 30 minutes = 600,000 minutes = 10,000 hours (FTC estimate).**

### 2.2 Screening & partial completes (often omitted)

To net 20,000 completes, FTC will contact more than 20,000 people. Let:

- **e** = eligibility rate among those who start screening
- **c** = completion rate among eligible

**Invites needed = 20,000 / (e × c).** Reasonable scenario ranges:

- High-efficiency: e=0.8, c=0.7 ⇒ invites ≈ **35,714**
- Low-efficiency: e=0.5, c=0.3 ⇒ invites ≈ **133,333**

Non-completes = invites − 20,000 ⇒ **15,714 to 113,333** people may spend **screening time** (assume 2 minutes avg).
Additional burden minutes ≈ **31,428 to 226,666** ⇒ **≈ 524 to 3,778 hours**.

> **Adjusted annual burden** = 10,000 + (≈524 to 3,778) = **≈10,500 to 13,800 hours**
> **Implied cost at $35–$75/hour** = **≈$368k to $1,035k**.

**Recommendation:** The final Supporting Statement should explicitly account for **screeners and partials** and commit to **annual true-up** of burden hours based on realized e and c.

---

# 3) Statistical performance (power/MDE) and design effects

Assume 10 studies/year with ≈2,000 completes each. For an equal-split A/B test on a binary outcome with p=0.5:

- **Standard error** = $\sqrt{[p(1-p)(1/n_1 + 1/n_2)]}$ with $n_1=n_2=1{,}000$ ⇒ $\sqrt{[0.25*(1/1000+1/1000)]} = \sqrt{(0.0005)} \approx$ **0.02236**.
- **MDE (80% power, α=0.05)** ≈ (1.96 + 0.84) × 0.02236 ≈ **0.063** (**6.3 p.p.**).

**More arms = lower per-arm n and higher MDEs.** With 5 arms (n≈400/arm), pairwise SE ≈ √[0.25*(1/400+1/400)] = √(0.00125) ≈ **0.03536** → MDE ≈ **~10.0 p.p.**

**Design effects (DEFF)** from weighting/clustered designs inflate SE by √DEFF:

- DEFF=1.25 (e.g., CV(weights)=0.5) → MDE × **1.118**
- DEFF=1.50 → MDE × **1.225**

**Multiple outcomes.** If 10 outcomes are tested with Bonferroni, α becomes 0.005 ⇒ two-sided z≈2.81; MDE inflates by **(2.81+0.84)/(1.96+0.84) ≈ 1.30** (≈+30%).

**Recommendation:**

- **Set per-study primary outcomes** and **target MDEs** (e.g., ≤4 p.p. for key binary outcomes) and adjust **sample sizes** accordingly.
- **Report realized DEFF** and **effective n** (n/DEFF) in all public technical appendices.
- **Limit arms or increase total n** when experiments are multi-arm or multi-endpoint.

---

# 4) Sampling, response rates, and bias mitigation

- **Frames & coverage.** Specify frames (e.g., address-based sampling, RDD, web panels, intercepts) and coverage gaps (language, device, platform).
- **Response rates.** Report AAPOR-standard **RR**, **COOP**, **REF**, **BREAKOFF**, with disposition codes.
- **Nonresponse bias.** Conduct bias analyses using frame covariates or administrative benchmarks; justify post-stratification / raking marginals.
- **Weighting.** Publish weight construction, trimming rules, design effect, and replicate weights for variance estimation.

**Recommendation:** Include a **minimum methods checklist** in each Information Collection (IC) under the generic clearance.

---

## 5) Instrument quality & measurement error

- **Pretesting/cognitive testing**: For any new/complex items (e.g., fraud typologies, sensitive financial questions), require **cognitive interviews** and **pilot tests** with predefined acceptance criteria (skip logic failure rates <2%, comprehension issues <5%).
- **Recall periods**: Fix and justify recall windows (e.g., 12-month fraud experience) to reduce telescoping.
- **Common metrics**: Standardize definitions (e.g., loss amounts, contact channels) for comparability across studies and years.

---

## 6) Privacy, confidentiality, and transparency

The notice contemplates collecting sensitive experience data (e.g., fraud). We recommend:

- **Data minimization** and explicit **PII handling** (collect only what's necessary for linkage or deduplication; purge linkers quickly).
- **Disclosure risk assessment** before any data sharing; use **noise injection or synthetic microdata** where appropriate.
- **Public documentation**: Post instruments, SAPs, codebooks, weighting documentation, and analytic code (with redactions as needed).
- **Retention limits**: Commit to strict retention schedules for directly identifying information.

---

## 7) Benefits and use to enforcement

To ensure surveys inform enforcement (as the notice intends), specify **learning objectives** per study (e.g., detect a ≥4 p.p. change in scam susceptibility after X intervention; estimate market share of a deceptive claim with ±2 p.p. precision). Publish an **outcomes dashboard** showing realized precision, effect sizes, and whether pre-specified decision thresholds were met.

---

## 8) Alternative designs to reduce burden or raise value

- **Sequential/multi-wave designs**: Start small, expand only if interim looks promising (group sequential boundaries).
- **Embedded field experiments**: Where lawful and ethical, test consumer education framings in-situ (site/app/email) to reduce measurement error.
- **Data linkage**: With consent, link to platform telemetry or complaints data to validate self-reports and reduce survey length.
- **Sample re-use panels**: Carefully managed short panels can improve power on change outcomes with fewer completes.

---

# 9) Concrete redlines & commitments for the final package

1. **Burden accounting:** Add a table separating **completes vs. screeners/partials** with scenario bounds; commit to **annual burden re-estimation** when realized rates differ by ≥10%.
2. **Power/MDE standards:** For each IC, publish an SAP with **primary endpoints**, **target MDEs**, **α/β**, **arm counts**, and **DEFF assumptions**.
3. **Methods minimums:** Require instrument pretesting results; AAPOR-standard response rate reporting; weighting & variance documentation; nonresponse bias analysis plan.
4. **Transparency:** Post instruments, SAPs, codebooks, and (where feasible) **de-identified microdata** with disclosure control documentation.
5. **Privacy:** Include data-minimization, retention limits, and risk-assessment language; prohibit unnecessary collection of sensitive identifiers.
6. **Performance dashboard:** Commit to an annual public summary with precision achieved, realized DEFF, response metrics, and enforcement relevance.

---

# Appendix: Worked calculations

### A. Burden (with screeners)

- **Base**: 20,000 completes × 0.5 hr = **10,000 hours**.
- **Invites** (range): 35,714 – 133,333 ⇒ non-completes 15,714 – 113,333.
- **Screening time** (assume 2 min):
    - Low: 15,714 × 2 = **31,428 min = 524 h**
    - High: 113,333 × 2 = **226,666 min = 3,778 h**
- **Adjusted burden**: **10,524 – 13,778 hours** (rounded in body text to 10,500–13,800).

### B. MDE (two-arm, binary, p=0.5, 80% power, α=0.05)

SE = √[0.25*(1/1000+1/1000)] = √0.0005 = **0.02236**.
MDE ≈ (1.96 + 0.84) × 0.02236 = **0.063** (6.3 p.p.).

### C. MDE inflation from multi-arm and DEFF

- Five arms: per-arm n=400 ⇒ SE ≈ √[0.25*(1/400+1/400)] = **0.03536** ⇒ MDE ≈ **~10.0 p.p.**
- DEFF=1.5 ⇒ SE × √1.5 = **×1.225** ⇒ two-arm MDE grows from 6.3 to **~7.7 p.p.**

### D. Multiple comparisons (10 outcomes, Bonferroni)

α* = 0.05/10 = 0.005 ⇒ two-sided z≈**2.81**.
MDE inflation factor ≈ (2.81+0.84)/(1.96+0.84) ≈ **1.30**.

---

# Conclusion

The FTC's generic clearance is a sensible vehicle to support enforcement with timely behavioral evidence, but the package will be significantly stronger—and better aligned with Paperwork Reduction Act principles—if it (i) fully accounts for real-world burden, (ii) sets and meets quantitative precision targets, (iii) standardizes high-quality methods and transparency, and (iv) fortifies privacy protections. Adopting the concrete commitments above will yield more credible, decision-useful evidence per respondent hour while honoring public burden.

*End of response.*

**Please contact the corresponding author above.**

---