Watching the Watchers: Designing a National Survey of Retail Investors After the Surveillance Debate

Fictitious Teaching case for Business Research Methods. Meridian Insight Group, its staff and the federal task force are fictional. The regulatory backdrop draws on real 2026 developments, simplified for classroom use. All appendix data are illustrative.

A $2 Million Question

Dana Whitfield had eleven days to tell the United States government how to ask tens of millions of Americans a question many of them did not want to answer.

It was a grey Monday in November 2026 at the Bethesda, Maryland, offices of Meridian Insight Group (MIG), a 240-person market research firm best known for consumer finance and public-sector studies. Three weeks earlier, MIG had won a federal contract worth $2 million. The client was the Interagency Task Force on Market Data and Investor Privacy, a working group convened by the Treasury Department and the Securities and Exchange Commission (SEC). Its brief to MIG ran just two pages. The central question fit in one line: *How have retail investors in U.S. equity markets been affected by the recent debate over government surveillance of their trading?*

Whitfield, managing director of MIG's public-sector practice, had run national surveys on student debt, health insurance and cryptocurrency adoption. This one felt different. "Usually the hard part is getting people to talk about money," she told her team. "Here we need them to talk about money *and* about being watched, and we are the government's contractor asking. That's a trust problem stacked on a measurement problem."

The research design memo was due to the task force on December 1. It had to say who would be surveyed, how they would be found, how they would be reached, what would be measured, and how $2 million would be spent. Once the task force approved it, the design would be locked. Fieldwork would start in February 2027, and the findings would feed a report to Congress in mid-2027.

Around the conference table sat four people with four different views on how to do it.

Background: The Watchers and the Watched

At the centre of the debate sat the Consolidated Audit Trail (CAT). Approved in 2016, the CAT was built to give regulators a single, cross-market view of every order and trade in U.S. equities and options. After the 2010 "flash crash", regulators argued they could not reconstruct market events quickly without such a system.

To link trades to people, the CAT originally collected investors' personal information: names, addresses, years of birth and, in early designs, Social Security numbers. Critics called it a de facto government registry of everyone who owns a brokerage account. Civil-liberties groups raised Fourth Amendment objections, and lawsuits followed. Costs ballooned too. The SEC's 2016 estimate for annual operating costs was about $55.8 million; by 2025 the figure exceeded $248 million.

The year 2026 brought the issue to a head:

- **Early 2026:** Plan participants scaled back the personal data reported to the CAT, though a customer identifier linking each investor's trades remained.
- **April 16, 2026:** The SEC issued a concept release launching a comprehensive review of the CAT, with options ranging from SEC ownership to eliminating it altogether. Comments were due June 22.

- **Mid-2026:** An inspector-general finding about sensitive data handling inside the SEC's own systems gave broker-dealer groups fresh ammunition to demand that retail identifiers be stripped out entirely.

Comment letters poured in. Some retail investors wrote that the CAT was the only tool that could catch manipulation and protect "the little guy". Others called it mass surveillance of law-abiding savers. Media coverage swung between "data breach waiting to happen" and "Wall Street wants fewer cops on the beat".

What no one had was systematic evidence. Had awareness of surveillance changed how ordinary investors traded, which brokers they used, or whether they invested at all? Had it eroded trust in markets, or strengthened it? Were some groups, such as older investors, active traders or immigrants, more affected than others? Comment letters were a self-selected, highly motivated sample. The task force wanted to hear from the silent majority.

Sources: National Law Review, Apr 2026 · InvestmentNews · Federal Register notice of the concept release

The Brief

The task force's statement of work was two pages long and short on specifics. It restated the one-line question, set a fixed price of $2 million for design, pre-testing, fieldwork, analysis and a final report, and asked for "nationally credible evidence". It did not define "affected", say which investors counted, or set any precision targets.

Because a federal agency was the sponsor, any survey of more than nine members of the public needed approval from the Office of Management and Budget under the Paperwork Reduction Act. That process could take 60 to 120 days and would scrutinise the burden on respondents, the confidentiality pledge and the expected response rate. In practice, the December design memo would also become the core of MIG's clearance package.

The kickoff meeting in late October added more voices than clarity. Whitfield's notes recorded four people, each with a different worry:

- **The Treasury representative** wanted to know whether households were "quietly leaving the market". Equity ownership underpinned retirement security, and she had heard stories of investors moving to crypto, gold or foreign platforms.
- **The SEC representative** cared about market integrity. "If people believe the CAT protects them from manipulation, that's worth something. If they think it's spying on them, that's a cost. I want to know which story they believe, and whether they even know it exists."
- **A privacy scholar on the advisory panel** worried about the people least likely to complain: older investors, immigrants and low-income savers who might simply stop investing without saying why. "Absence of protest is not absence of harm."
- **A congressional liaison** was blunter. "My boss needs one number for a hearing. What percentage of investors changed what they do because of this?"

Driving back, Whitfield reflected that the four had asked four different questions. Some might not be answerable with a survey at all. Turning them into research questions was MIG's job, not the client's.

One line in the brief worried Whitfield more than the rest: *"Findings must withstand scrutiny from stakeholders on all sides of the CAT debate."* Brokers, exchanges, privacy advocates and investor-protection groups would all read the report. Any of them would attack a design that looked tilted.

Four Voices Around the Table

"We don't yet know what we're measuring"

Priya Raman, MIG's head of qualitative research, spoke first. "'Affected by surveillance' is not a variable. Is it fear? Anger? Indifference? A broker switch? We don't know how investors even think about this. Most have never heard

of the CAT." She proposed starting with 40 in-depth interviews and eight online focus groups across investor types. Only then would MIG write a questionnaire. That would take about eight weeks and $180,000.

She also pointed to a framing problem in MIG's quick pre-test (Appendix 2). When the same question used the word "surveillance", concern was 27 points higher than when it said "regulatory record-keeping". "Whoever picks the words picks the headline," she said.

"Only a probability sample will survive Congress"

Marcus Oyelaran, MIG's chief statistician, agreed the constructs needed work but insisted the core study be a probability sample. He wanted address-based sampling (ABS): drawing addresses from the U.S. Postal Service file, mailing invitations with a $5 prepaid incentive, and pushing people to a web survey with a paper option for holdouts. Coverage would be nearly complete. But only about 38% of households held equities outside retirement plans (Appendix 1), so MIG would need to screen roughly three households for every eligible one.

With no precision targets in the brief, he set his own: ±2 points nationally and ±5 points for key subgroups. His numbers were sobering. A simple random sample needs about 2,400 completes for ±2 points nationally. With a design effect of 1.6 from weighting and clustering, that rose to about 3,850. Adding eight subgroups at ±5 points each pushed the target to roughly 6,000. At an estimated $210 per eligible complete, ABS alone would consume about $1.26 million (Appendix 3).

Oyelaran's deeper worry was nonresponse bias. "The people most worried about government surveillance are exactly the people least likely to answer a government survey about it," he said. "If we miss them, we'll report that nobody cares, and we'll be wrong." He wanted to budget for a nonresponse follow-up: a subsample of non-respondents re-approached by priority mail with a $40 incentive and a shorter questionnaire.

"Go big, go fast, go online"

Jake Tran, who ran MIG's digital panel partnerships, thought the ABS plan was a museum piece. "For $400,000 I can get you 20,000 completes from opt-in online panels in three weeks. Add another 3,000 active traders recruited through investing forums and trading apps. That's where the action is: these are the people who post about the CAT."

Tran argued that weighting to benchmarks could fix the demographics. Oyelaran disagreed sharply. Opt-in panels could be balanced on age and income but not on attitudes toward privacy. People who join online panels for gift cards may be systematically less privacy-conscious. Tran's own data showed opt-in respondents were twice as likely as ABS respondents to say they "don't mind" being tracked (Appendix 3). Still, Whitfield noted, a large opt-in sample could power a **survey experiment**. Respondents could be randomly assigned to read different neutral descriptions of the CAT, letting MIG estimate the causal effect of information on trust and trading intentions.

"Compared with what?"

Tom Keller, MIG's director of operations, raised a design question nobody had settled. "The brief says *impact*. Impact means change. What's our baseline?" MIG had no pre-2026 data on these exact questions. The team listed three options:

- **Retrospective questions** ("Compared with two years ago, do you trade more, less or the same?"). Cheap, but vulnerable to recall error and motivated reasoning.
- **Secondary benchmarks** such as earlier national financial capability and investor surveys, which measured trust in markets and trading frequency before 2026. Comparable only where wording matched.
- **A two-wave panel**, re-interviewing respondents in late 2027 after the SEC decides the CAT's future. This would capture real change, but a second wave would cost $350,000 to $450,000 and attrition could run to 30%.

Keller also flagged the irony of one proposal: linking survey answers to respondents' actual brokerage records, with consent, to validate self-reported trading. "We'd be fixing a measurement problem by doing the very thing we're asking people about," he said. "Imagine that headline."

Confidentiality and trust

Everyone agreed that the confidentiality pledge would shape response. MIG could collect data under the federal statistical confidentiality law, CIPSEA, which bars use of responses for enforcement. That protection would need to be explained plainly. MIG would also have to commit to collecting no account numbers and to holding all data on its own servers, not the government's. Raman's interviews suggested a sponsor-disclosure dilemma. Naming the Treasury and SEC would raise legitimacy for some investors and suspicion for others. Yet OMB rules required the sponsor to be disclosed.

The Decision

By Thursday, Whitfield had boiled the debate down to three designs, costed in Appendix 4:

- **Design A, "Gold Standard":** a probability-only ABS survey of about 6,000 investors, with a nonresponse follow-up and minimal qualitative work.
- **Design B, "Scale and Speed":** a 20,000-person opt-in online sample plus forum recruits, a large framing experiment, and a second wave in late 2027.
- **Design C, "Hybrid":** qualitative exploration first, then a smaller ABS core of about 3,900 completes for headline estimates. An opt-in sample of 8,000 would carry the experiment and subgroup analysis, calibrated against the probability core.

Each had a weakness that a hostile reader could exploit. Design A might meet the precision targets but lacked a baseline and could still miss the privacy-anxious. Design B was cheap per interview but would struggle to claim national representativeness. Design C tried to do everything, and Keller warned that it left almost no contingency if OMB clearance slipped or response rates disappointed.

Whitfield looked at the whiteboard, now covered in sample sizes, unit costs and arrows. Something was missing.

For three days her team had argued about *how* to ask. Nobody had written down, in plain words, *what* the study had to answer, or which of the task force's worries a survey could address and which it could not. Until that was settled, there was no way to judge which design was better. Each was best at answering a different question.

"Whatever we choose," she told the team, "someone will say we rigged it. Before we pick a design, we need to agree on what we're trying to find out. Then we need to show why this design, and no other, answers it."

Assignment Questions

The case does not state the study's research questions. Deriving them from the setting is part of the task. Treat the appendices as evidence to reason with, not as answers.

1. What is the management dilemma facing the task force, and how does it differ from the research problem facing MIG? Derive the research questions the study must answer and the investigative questions beneath them. Which stakeholder concerns can a survey address, and which can it not?
2. What kind of knowledge does each of your research questions call for? What does that imply for the overall logic of the research design?
3. Each voice at MIG's table rests on assumptions about the population, about measurement and about what counts as evidence. Surface and critique those assumptions, using the appendices.

4. Propose and defend a research design within the $2 million budget. Show how each element maps to a research question you derived, and state what your design chooses not to find out.
5. "Asking people whether they feel watched is itself a form of watching." Discuss the ethical and methodological implications for this study.

Appendix 1: Who Is a Retail Investor?

Of roughly 132 million U.S. households, MIG estimates about 50 million (38%) hold equities directly, outside employer retirement plans. These are illustrative working estimates compiled from secondary sources, not official statistics.

Exhibit 1a. Investor households by investing style

Segment	Households (millions)	Share of investor households	Typical trades per year
Buy-and-hold, mainly ETFs and index funds	21.0	42%	1–5
Occasional stock pickers	14.5	29%	6–50
Robo-advisor or managed accounts only	5.0	10%	Automated
Active traders	4.0	8%	50+
Options or margin users	3.0	6%	50+
Employee stock plan only	2.5	5%	0–2
Total	**50.0**	**100%**	

Exhibit 1b. Priority subgroups: expected vs. required completes

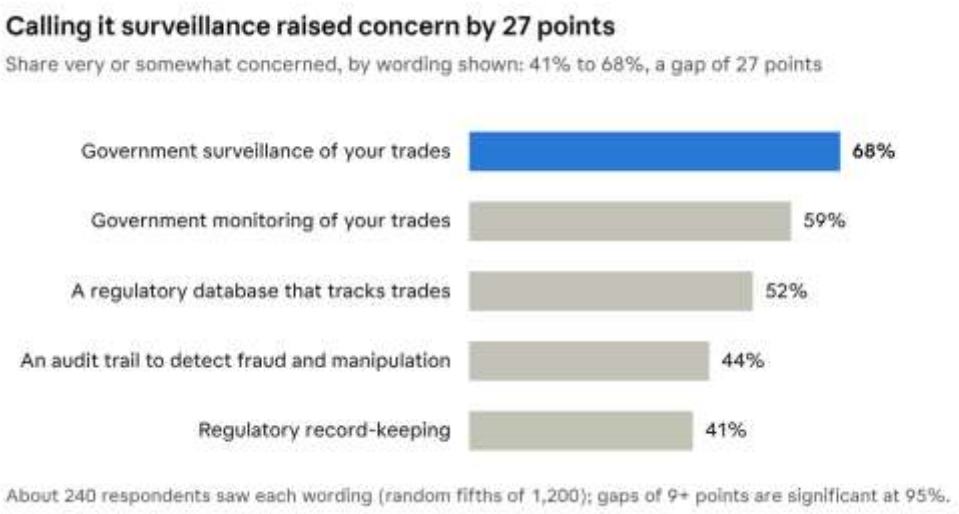
Required completes assume ±5 points at 95% confidence and a design effect of 1.6 (385 × 1.6 ≈ 615). Expected completes assume a proportional sample of 3,850 investors.

Priority subgroup	Share of investors	Expected completes (proportional)	Required completes	Shortfall
Household income under $75,000	27%	1,040	615	None
Age 65 and over	24%	924	615	None
Age 18–34	21%	809	615	None
Foreign-born	12%	462	615	153
Hispanic	9%	347	615	268
Asian American	8%	308	615	307
Active traders	8%	308	615	307
Black	7%	270	615	345

Appendix 2: Pre-test Findings

In October 2026 MIG ran a quick pre-test with 1,200 retail investors from an opt-in panel, plus 12 cognitive interviews. Wording alone moved concern by 27 points.

Exhibit 2a. Framing experiment: how the CAT was described



Illustrative data · MIG pre-test, opt-in panel of 1,200 retail investors, October 2026

Exhibit 2b. Other pre-test signals

Indicator	Result
Had heard of the Consolidated Audit Trail	14%
Answered at least 2 of 3 knowledge items correctly	9%
Said they trade less because of government monitoring	11% overall; 4% among those who had never heard of the CAT
Item nonresponse on total portfolio value	19%
Breakoff rate, sponsor named vs. not named in invitation	9% vs. 6%
Median completion time (target: 15 minutes)	17 minutes

Cognitive interviews: several respondents asked whether the survey itself was "part of the tracking". Two older respondents read "audit trail" as a tax audit.

Appendix 3: Comparing Sampling Frames and Modes

The cheapest frames are the least representative of privacy attitudes. The "Don't mind being tracked" column comes from a comparable item in MIG's 2025 consumer-data study, fielded on each source.

Frame and mode	Coverage of investor households	Expected response rate	Cost per eligible complete (US$)	Field time	"Don't mind being tracked"	Main risk
Address-based sample, mail push-to-web, paper option	~97%	18%	210	12–14 weeks	21%	Nonresponse among privacy-anxious investors

Frame and mode	Coverage of investor households	Expected response rate	Cost per eligible complete (US$)	Field time	"Don't mind being tracked"	Main risk
Telephone, random-digit dial	~90%	5%	260	10 weeks	23%	Very low response; interviewer effects on sensitive items
Probability-based online panel	~94%	8% (cumulative)	95	4–6 weeks	27%	Panel conditioning; thin subgroups
Broker customer lists (partnership)	~35% (partner brokers only)	9%	40	6–8 weeks	24%	Broker conflict of interest; respondents may see it as surveillance
Opt-in online panels	Unknown	Not calculable	18	2–3 weeks	42%	Self-selection; professional respondents
Investing forums and trading-app recruits	Active traders only	Not calculable	12	2 weeks	15%	Highly motivated respondents; bots and duplicates

Notes: The ABS response rate is for the screener; about 38% of responding households are eligible. A nonresponse follow-up subsample would add roughly $55 per case attempted.

Appendix 4: Three Candidate Budgets

All three designs fit within $2 million but buy very different evidence. Design B leaves $564,000 unallocated, while Design C keeps only a 2.5% contingency.

Line item (US$)	A: Gold Standard	B: Scale and Speed	C: Hybrid
Qualitative exploration (interviews, focus groups)	40,000	40,000	180,000
Questionnaire design and cognitive testing	60,000	60,000	60,000
OMB clearance support	70,000	70,000	70,000
ABS probability core	1,260,000 (6,000 completes)	0	819,000 (3,900 completes)
Targeted oversampling of priority subgroups	0	0	130,000
Nonresponse follow-up	82,500	0	55,000
Opt-in panel sample, incl. framing experiment	0	360,000 (20,000	144,000 (8,000

Line item (US$)	A: Gold Standard	B: Scale and Speed	C: Hybrid
		completes)	completes)
Forum and trading-app recruits	0	36,000 (3,000 completes)	12,000 (1,000 completes)
Fraud and bot screening	0	60,000	30,000
Second wave, late 2027	0	400,000	0
Weighting, calibration, analysis and reporting	220,000	260,000	300,000
Project management and data security	150,000	150,000	150,000
Contingency or unallocated	117,500	564,000	50,000
Total	**2,000,000**	**2,000,000**	**2,000,000**
Probability-sample completes	6,000	0	3,900
All completes, wave 1	6,000	23,000	12,900

Unit costs follow Appendix 3: $210 per ABS complete, $18 per opt-in complete, $12 per forum recruit and $55 per nonresponse follow-up case attempted.