NOAHAIC · CASE STUDY

How Close Can a Synthetic Panel Get to Real Field Research?

44 AI respondents, interviewed weeks before field research began, predicted the field pilot's price, order size, sentiment, and viral channel — scored openly, dimension by dimension.

The situation

A publishing organization was evaluating the launch of a new educational resource for a large target audience and wanted to understand likely adoption, pricing, purchasing behaviour, and communication preferences. Adoption runs through decision makers who recommend resources within their networks and purchase in bulk. Before committing to field research and launch spend, the client needed early answers on adoption intent, trusted messengers, content priorities, order sizes, and price.

What we did

NoahAIC built a synthetic panel of **44 AI respondents** modelled on the target buyer — prospective buyers and decision makers across seven states, with calibrated income, setting, and education profiles — and interviewed each one conversationally across 13 questions. The client separately fielded a real pilot: a mobile-form survey answered by **25 actual prospective buyers**. The synthetic study was run independently; no field responses were used to tune the panel. When the real data arrived, we scored every synthetic prediction against it.

What the panel predicted

Dimension	Synthetic panel (n=44)	Field pilot (n=25)	Outcome
Comfortable bulk price (median)	₹210	₹175 overall; ₹225 urban slice	**Within ₹35 (~17%)**
First bulk order (median units)	27.5	20	**Within ~8 units**
Willingness to refer through peer networks	100%, messaging groups named unprompted	25 of 25, same channel volunteered	**Exact match**
Willingness to recommend to others	100%	~100%	**Match**
First reaction to the concept	Positive/curious, zero rejection	~92% positive, zero rejection	**Match**
Perceived helpfulness	98% very helpful	~100% affirmative	**Match**
#1 trust signal before recommending	Peer testimony — mentioned by 95% of agents	Peer testimony — the #1 field answer	**Match**
Top-tier content theme	Flagged in the panel's top tier	#1 most-requested theme in the field	**Match**

On every launch-critical dimension — what to charge, how large a first order looks, how the concept lands, whose word carries trust, and which channel spreads it — the synthetic panel delivered decision-grade answers **weeks before a single field response existed.**

What the comparison sharpened

Running synthetic and real studies side by side is itself the product, and this pairing sharpened the platform in three specific ways:

- **Reporting fidelity.** Where agents gave multi-part answers (e.g., naming testimony, endorsements, and credentials together), our dashboard's single-label summaries kept only one signal. The agent transcripts matched the field data; the reporting layer flattened them. Multi-label extraction is now live, so headline findings preserve everything the agents actually said.
- **Segment-matched pricing.** Field answers split cleanly by setting: the urban slice's median (₹225) sat almost exactly where the panel landed, while village respondents clustered near a much lower floor. The combined read produced a two-tier pricing recommendation — a standard urban tier and a subsidized village tier — that neither dataset alone would have surfaced as sharply. Panels are now composition-matched to the client's exact field frame, including outlier archetypes such as large-network coordinators.
- **Deeper local grounding.** Field respondents phrased content demand in the language of everyday life, alongside the formal themes the panel emphasized. Agents are now enriched with culturally embedded, audience-specific source material so simulated demand mirrors how real respondents actually talk.

What field data alone could not do

Small field pilots carry real noise: **9 of 25 respondents could not commit to an order quantity**, answers averaged roughly three words, at least two price responses were likely question misreads, and — as the client observed — respondents were guarded in a

typed form. All 44 synthetic respondents returned complete, elaborated, conversational answers on every question, with clean central estimates that the field data went on to confirm.

That is the point of this study: **the validation work has been done, so our clients don't have to repeat it.** New engagements inherit this accuracy record and start directly with simulation — decision-grade answers in hours, with no fieldwork cost or wait. And for any client who wants proof on their own category, NoahAIC scores its predictions against real signals the client already collects — past surveys, sales data, pilot responses — included in the engagement, never as an added field project. Every scored engagement extends the accuracy record you are reading.

Method note: Field n=25, self-selected respondents, short typed answers. Panel and field samples differed somewhat in composition and setting mix. Client and product details are withheld under confidentiality; all figures are reproduced as measured. Study conducted 2026.