



NAVIGATING THE ONLINE SAMPLE MARKET

The pursuit of quality

By Jeremy McNamara, Nancy Brigham and Frank Kelly | November 2019

IPSOS VIEWS

GAME CHANGERS



EXECUTIVE SUMMARY

Gaining access to the right respondents has been the foundation of effective market research since the discipline began. Originally conducted on a face to face (FTF) basis with pencil and paper, market research interviewing has progressed through different access channels in response to a digital transformation that has changed consumer habits and expectations as well as the need to reduce costs. Telephone surveys (CATI) first became more prevalent, followed by online surveys. Today, in developed markets such as the USA and Western Europe, the majority of surveys are conducted online, both through web and mobile access routes. And as internet penetration increases globally, the online channel becomes more significant. Although FTF and CATI will probably always exist for specific survey types and niche use-cases, the ability to access and engage respondents online is now, more than ever, an important focus for market research.

The research market itself has undergone a transformation and today there are many different kinds of “insight provider” including traditional research firms, technology companies, and small boutique agencies. Not all of these have the resources or the need to manage their own panels of respondents, instead buying access to respondents on a project basis from the market. In this way, access to respondents has grown as a distinct market segment, and there is an array of providers of pools of respondents.

Commonly referred to as “sample” (the required target in a research survey), the market is as intriguing as it is complex. The commoditisation of consumer sample has led to a quest for a combination of speed, quality and truthfulness in respondent access.

This paper will explain the dynamics of the online sample market as it has developed and how researchers can best equip themselves to respond to challenges and demands in this evolving landscape.

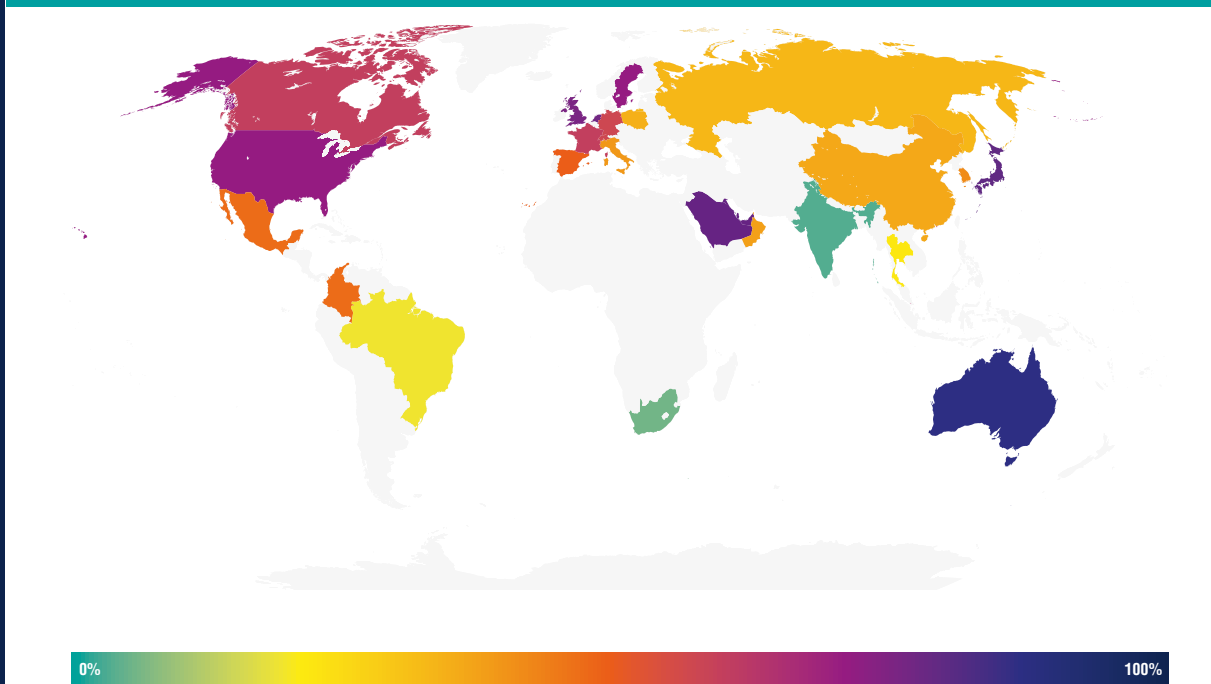
The overall growth of the survey research industry and the rate at which fieldwork transitions online are the key driving forces for online sample. In major Western markets, projects transitioning from offline to online has slowed to a trickle as everything that is suitable for online has already moved. But for many parts of the world this process is ongoing.

The ability to access and engage respondents online is now, more than ever, an important focus for market research

The size and status of the global online sample market is affected by country infrastructure. In countries that have been slow to offer affordable internet access, CATI and F2F research have remained the principal research methodologies. Latin America has long favoured F2F research due to high urbanization and a population that is highly stratified in income. This means that they are still in the process of transitioning research from F2F to online and are experiencing high growth rates. In addition, selected Asian, African and Central European countries are also experiencing 10%-30% annual growth in demand for online sample. (See figures 1 and 2 for a global view of online transition.)

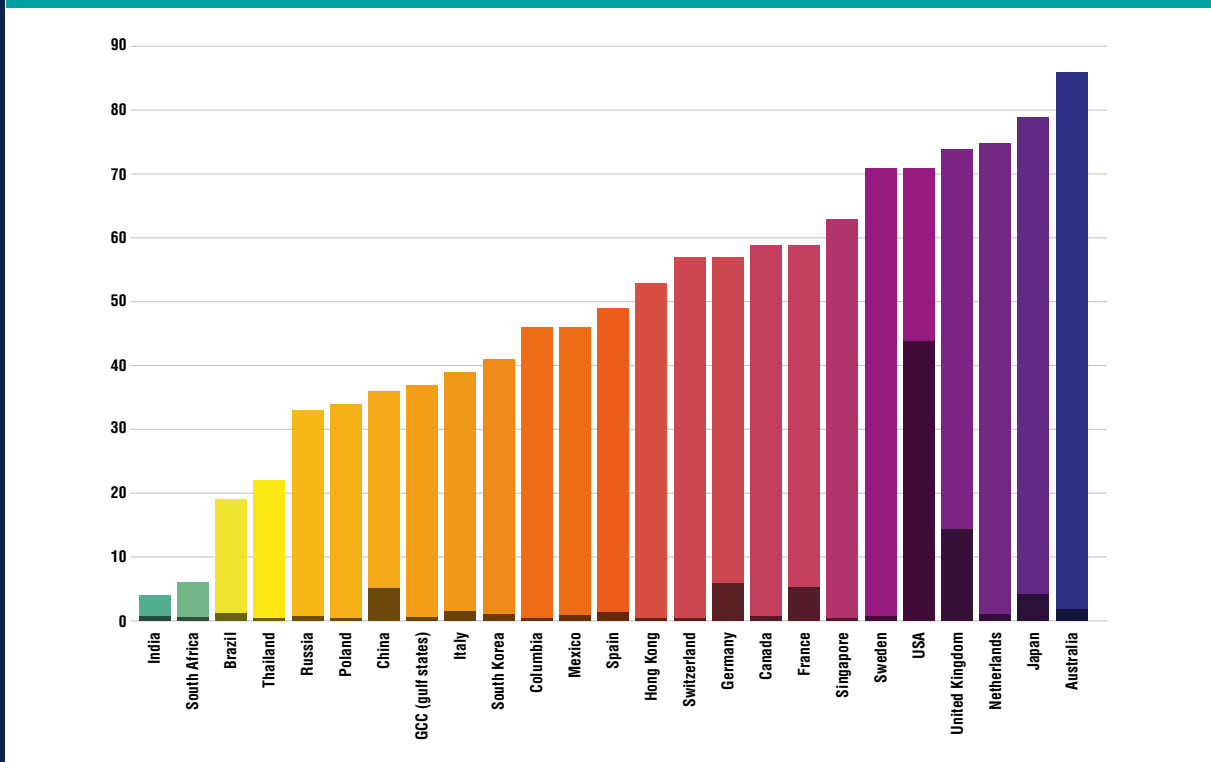
A challenge in transitioning online is that F2F interviews tend to be quite long, lasting more than 30 minutes on average, but for optimal use in mobile browsers, they should be 15-20 minutes or less. Structural design changes to questions are also required. Expertise in handling this transition is key to obtaining accurate research results in these markets.

Figure 1 Global rate of survey transition to online



Source: ESOMAR Global Market Research Industry Report 2019. pp.138-139 & 154-155

Figure 2 Global rate of survey transition to online (% of total), showing total market spend



Source: ESOMAR Global Market Research Industry Report 2019. pp.138-139 & 154-155

THE ONLINE SAMPLE MARKET: AN OVERVIEW

The online sample market as we know it today is less than 20 years old and is estimated to be worth around \$1.5 billion. During this time, there have been four key factors driving the choice of sample supplier: speed, capacity, cost and research quality. Buyers of sample want capacity, or the ability for one supplier to handle all sample requirements. They also want a quick turn-around and competitive costs, and everyone wants to be assured of research quality – even though few can define what this is.

While, initially, research principles and capable proprietary technology were key constituents of being competitive in the sales of online sample, digital marketing and modern technology that can easily integrate with external platforms are also part of the mix today.

Research principles and heritage have become less compelling as the industry has shifted towards offering clients speed and innovation. This has meant that the market is awash with cheap and quick sample suppliers while research-oriented, quality sample is in short supply. New sources of respondents enter the market every day, but few of them have successfully differentiated themselves as a higher quality product.

Suppliers to the market comprise retail players, wholesale providers and aggregators and many companies in the growing DIY sample market space. Readily available sample enables companies to create surveys in whole or in part without engaging a professional research organisation, which has the advantage of reducing cost and, potentially, time.

Many companies looking to cost-reduce their research activities have attempted to “unbundle” sample as a discrete component, which has also fuelled market growth.

Everyone wants to be assured of research quality – even though few can define what this is.

SUPPLY

The initial participants in the online sample market were internet start-ups and fieldwork companies. Then, research companies, direct marketing firms, loyalty businesses and sample exchanges entered the market at a later stage.

Most sample providers are aggregators, meaning they combine multiple sources of sample into one stream of respondents to send to a survey, and their own panels vary in size and scale. The aggregator is responsible for managing all aspects of the sample, including relationships and communication with sample sources, ensuring respondents meet the specific needs of the study, and troubleshooting if there are issues with field and data – including quota fulfilment. Both before and during field, sampling involves a lot of careful handling of the sample and dialogue with researchers. As their single point of contact, the aggregator must not let the extra layer of management slow down the field process.



As most sample sources are blends, buyers of sample today are much less aware of the origins of their sample than they were five or more years ago.

Obscurity in the market means that there is a lot of misinformation about the size and reach of respondents that providers have access to, which can cause widespread confusion about the different types and sources of sample supply.

While very few sample providers manage their own panels, Ipsos does both: it owns and operates survey panels but is also an aggregator as it sources sample from a set of preferred partners to enhance the richness and diversity of its proprietary panels.

DEMAND

Market research agencies (MRAs) make up most of the market for research sample. Only a handful of the largest research companies maintain their own panels, while the rest buy sample from other companies that maintain panels, exchanges or intercept sources. The larger MRAs have several preferred vendors and often have a service level agreement with those vendors.

Panel companies buy sample from each other to fulfil difficult ad hoc jobs or to provide sample consistency throughout the year on large scale trackers. Much of this work tops-up or completes the target sample population when their own panel is exhausted, but some of it can be profitable if the target audience is well profiled within the panel.



MRAs often find themselves under pressure to deliver against tight deadlines and may fall into the trap of compromising quality by valuing fieldwork speed and price as the top two criteria.

SOURCES

Research panels continue to be the best source of stable and engaged research participants. Their benefits include:

- More is known about their participants due to profiling data that is compiled over time.
- Respondent performance can be assessed over time and bad members removed.
- Panel members build trust with a panel brand over time and are more willing to undertake a broad range of research activities.

However, due to their managed access platforms and volumes of studies offering rewards for completion, research panels are also more vulnerable to many types of research fraud such as bots and click farms. A panel requires a solid quality control process, just as much as dynamic sources do.



Research panels are also more vulnerable to many types of research fraud such as bots and click farms.

The key quality differentiator for panels is sample stability. While respondents change over time within a panel as with any source, panels can be carefully managed to minimize variation over time. Stable panel sources are often helpful for trackers or research products that include benchmarks.

DELIVERY

The industry is quickly transitioning to online quotation and ordering (including programmatic methods) so that clients no longer have to use email or telephone to obtain a quote for basic sample requests.

We have also seen a move from project-based buying through bulk-buying to full programmatic buying via sophisticated pricing algorithms. Many of the major players now adapt pricing based upon supply and demand at a given moment; creating a real-time automated bidding process, much like the way that digital ad space is sold.

This “programmatic sampling”, essentially the automatic bidding, buying, selling, and/or fulfilment of sample without human intervention, connects buyers and sellers throughout the sample supply chain. The integration of multiple sources makes sample access faster and invisible to the user, enabling quicker feasibility decisions to be made. This technology can also minimize project costs.

Integration can be either spec-based, where buyers submit their study specifications, or individual-based. For the latter, the buyer can select and invite individual respondents through a mirror copy of the seller’s panel database.

The online sample market is large and complex. Clients need a partner who can help them to navigate this complex and dynamic market.

NOT ALL SAMPLE IS CREATED EQUAL

Industry studies have shown that sample/panel sources are not completely interchangeable¹, and this is a basic principle for online survey samples.

Different panels will give different results, which means that control and transparency are increasingly important. For example, new respondents tend to rate product concepts more favourably than longer-tenured respondents, so this effect will be visible in panels with a higher proportion of new respondents.

Given the wide variety of sample sources and their benefits and drawbacks in cost and quality, some clients struggle with the question, “How do I choose the right blend for my research without impacting my data?”

Some of the factors to consider are:

- Different sample sources are created, managed, and accessed differently, which in turn attracts different types of people and affects the outcomes. There is variance in how sample sources respond to surveys, but there can also be variance within the same source, if that source is an aggregator.
- A key point of difference among types of sample relates to the level of commitment that a person makes when

completing a research survey. Those who sign up to a panel are making a longer-term commitment, so tend to be willing to take longer surveys and do different kinds of research tasks. But respondents who are intercepted while engaged in another online activity are unlikely to be willing to devote the same amount of time to it. The sample types must be selected carefully based on the type of research.

- It is sometimes necessary to combine multiple data collection methods together into one study to complete the sample (mixed mode). This raises the challenge of eliminating mode effects, but it does mean that demographic differences can be accounted for.
- Even within one widely-used mode, there are multiple streams of sample sources which often need to be combined (aggregated) in one study to furnish all the desired respondents. Blended sampling is becoming the norm today. This increased complexity means that successful sampling requires expertise in managing multiple sources and streams of sources.
- Even when samples are balanced carefully for demographic and socioeconomic variables, there are attitudinal or “cultural” differences among respondents that can influence results.



- The many different data collection modes/sample sources and the need to combine them in one study can lead to inefficiencies and inconsistency in data. Inconsistency can be managed and minimised through sampling design and careful handling. Inefficiencies are typically handled by optimising the data collection operations such as sample aggregation.

For both the sample provider and the researcher, there are three key elements for successfully managing multiple sample sources:

1. An understanding of the quality and responsiveness of the different sample sources, and how the respondents are being recruited and incentivised.
2. Deep expertise in sampling, such as in blending sources, sample balancing, quota management and sample design.
3. A sense of urgency and timely response to sample and feasibility requests from researchers, including 24-hour accessibility for field sample requests.

Without these three components, research costs and timing will increase, e.g. incidence may be lower so study costs will increase, quotas will take longer to fill if sample is not sent in a timely manner, and data/insights will be compromised. For example, respondents from one source may be incentivised by the study's topic, artificially inflating scores, or blend proportions may not be well-maintained across tracker waves, not allowing data to be trended.

So, while sample aggregators can introduce efficiencies into sample sourcing, data quality within this process is key. For Ipsos, this is even more important than efficiency, as the sample itself is the foundation of the insights that we deliver. Ensuring data quality means having control over sample sourcing and transparency with each supplier. We don't want to simply access respondents without regard to the quality of those respondents.

Quality is driven by where and how respondents are recruited, incentivised, managed (at an individual study level and over multiple studies), and sampled (before being sent to Ipsos). To deliver quality it is essential that we understand and manage them appropriately within a study, or across studies (in the case of trackers/normed studies). This is a quality differentiator for Ipsos.



THE FOUR PILLARS OF QUALITY CONTROL

We believe in the mantra that “Good data in means good data out”, so building in quality control is vital. This starts with the respondents. Quality respondents that generate reliable survey data must be Real, Unique, Engaged and Fresh. Here are some examples of how Ipsos approaches these four pillars:

REAL

Ensuring that a respondent is who they say they are. Some approaches to determine this include:

- A double opt-in approach for email validation and confirmation of match between device settings and geo-location, anonymous proxy detection.
- Detection of robots via the CAPTCHA code validation.
- Removal of records with email addresses which are available only for a few minutes or for a limited number of messages.
- Checking against the Ipsos blacklist, detection of data anomalies and patterns.

UNIQUE

We detect duplicates for panel and survey integrity.

De-duplication detects people that are trying to complete the same survey more than once, from multiple accounts or on different devices. Ipsos uses an industry leading “fingerprinting” solution for de-duping.

Ipsos’ online sample quality program has a strong competitive advantage compared to others who might only eliminate offenders post-field, or use custom algorithms based on pilot samples.

ENGAGED

Immediately after joining, a panelist’s survey-taking behaviour is evaluated to detect fraudulent or inappropriate behaviour, for example:

- “Speeding” respondents, who are inattentive and complete surveys too quickly.
- “Straight lining” respondents, who choose the same answer to all statements in a grid.
- Quality evaluation of responses to open-ended questions.

Survey-taking behaviour is tracked in real time through self-adjusting algorithms, and the panelist’s history is monitored across all surveys and used for purging procedures to remove bad or inactive respondents.

FRESH

Controlling the number of surveys that respondents are invited to by using exclusion rules. These take into account the numbers of surveys, type of study, the category researched and the number of surveys respondents have already participated in.

While the quality of the end data relies heavily on the respondents, this is underpinned by technology that enables the right “chemistry” between respondent and questionnaire, namely, automated systems that are invisible to the user and do not interfere with survey experience, working in real-time.

All these factors need to be accounted for to minimize bias as the sample plan is built and respondents are sourced and delivered to a survey.

THE IMPORTANCE OF 'REAL'

As respondent fraud is increasing with the continuous pace of technological change, it is important to continuously develop new ways of checking survey behaviour. This also helps us to deeply understand how respondents take part in studies.

Ensuring authenticity and avoiding fraud requires a leading-edge technology. True fraud doesn't happen very often at all, but when it does, it can devastate the results of a study. Due to proprietary algorithms we have developed in this area, and strategic use of specialised industry solutions, Ipsos' quality approach helps to protect the integrity of data.

Our checks dedicated to identifying and removing fraudsters include:

- **Removal of blacklisted email domains and IPs.**
The list includes email domains of all clients, competitors, and employees of Ipsos. No-one who is remotely associated with you or your survey can enter the study.
- **Anonymous open proxy detection.** Someone connected from an anonymous open proxy is hiding her/his device identity and geo-location. This behaviour is highly correlated with deliberate fraud attempts and is screened out immediately.
- **Removal of copy/paste or robot answers.** Answers pasted into a text box from the respondent's clipboard or inserted by an automated script are flagged and removed.
- **Detection of suspicious patterns and anomalies.** These can be found in the name, email, IP, and/or demographic information collected at registration. Accounts having multiple elements in common are deactivated. Semi-automated procedures created by Ipsos and maintained in-house are employed at an early panel stage.
- **Artificial Intelligence and Machine Learning techniques.** We are incorporating the power of AI into fraud detection in our proprietary panels to better

predict and detect new fraud patterns as they occur.



QUALITY AND INNOVATION

At Ipsos, we believe that cost should not come at the expense of quality. Investing in high quality sample will be beneficial in the long-run, as respondents who care more about the survey will respond more readily, give more thoughtful answers, complete the surveys, and not drop out or be removed from the sample.

This extends to quality partnerships. Knowing who to partner with is highly important to researchers. In other words: whose sample can you trust?

At Ipsos, while our own sources provide the foundation of our sampling, in situations where we need to supplement with other sample sources, we use a rigorous vetting process based around control and transparency to ensure that the sample partners we work with can deliver respondents—and survey results—that are just as accurate and trustworthy as our own.

Ipsos invests significantly in research on research (RoR) in order to ensure survey results are stable and meeting standards. This informs us about the impact of respondent and marketplace changes on sampling, sample design,

surveys, quality, data and business decisions, research processes, and other elements. On any given day we might examine questionnaire changes, new sampling technologies, mixed modes, sampling algorithm changes, field timing changes, respondent access points, and so on. RoR also helps us establish best practices in new areas, such as audio and video open-ends.

Ipsos was recently recognised as the most innovative research company in the 2019 GRIT 'Top 50 Most Innovative Suppliers' list.²

Clients need a provider who understands the dynamics of the online sample market and how to optimise sources for the highest levels of quality, security, speed and consistency.

SO WHAT?

Seeking, securing and engaging real consumers to undertake research surveys has never been as challenging as it is today. Capturing and retaining the attention of the survey respondents that research firms want to hear from is a continual battle for success in the online research space.

In these times of pressure on operating costs, the research industry should never abandon the quest for quality.

Changes in the online sample market have resulted in a complex respondent ecosystem as there are now many sample sources, and types/streams of sources. This, along with the noise and proliferation of media that consumers are

exposed to, requires a combination of skill and technology to master. Not all players are able to create the systems that are properly equipped to do this, so there are invariably questionable or even fraudulent activities which need to be safeguarded against.

It is vital that those wishing to run surveys are aware of the risks and pitfalls of the respondent sample that they use and work with a sample provider who can navigate through this perfect storm, ensuring that the data provided is accurate, fit for purpose, and a reliable foundation on which to base their business decisions.

REFERENCES

1. R. Walker, R. Pettit, and J. Robinson. "The Foundations of Online Research Quality, Knowledge Brief 3: Inter Study Comparability and Benchmark Analysis." New York: Advertising Research Foundation, 2009.
2. GRIT 'Top 50 Most Innovative Suppliers' List 2019: <https://greenbookblog.org/2019/07/18/the-top-50-most-innovative-suppliers-for-2019/>

NAVIGATING THE ONLINE SAMPLE MARKET

The pursuit of quality

Jeremy McNamara Global Service Line Leader, Observer, Ipsos

Nancy Brigham Head of Sampling & Research, Global IIS, Ipsos

Frank Kelly Head of Operational Activation & NPD, Total Operations, Ipsos

The **Ipsos Views** white papers are produced by the **Ipsos Knowledge Centre**.

www.ipsos.com
[@Ipsos](https://twitter.com/Ipsos)

GAME CHANGERS

