

CONSIDERATIONS IN QUANTITATIVE DATA COLLECTION IN CHINA

A Market Strategies International White Paper

January 2007

Commercial market research is a relatively recent industry in China, blossoming only since the 1980s. As a young and growing industry, operation practices have been undergoing changes and improvements. In this white paper, we lay out considerations in data collection of quantitative studies specific to the Chinese market, with a focus on technology research as opposed to other product sectors.

MODES OF DATA COLLECTION

All methods of data collection such as web surveys, phone interviews, face-to-face interviews, central location tests, mail surveys and so forth exist in China as in other markets. What differentiates China from other countries is the relative popularity of certain methods and the twist put on certain methods to suit the local context. The 3 most common methods are described below.

Face-to-face interviews. Face-to-face interviewing has been the historical default mode of data collection from the beginning of commercial market research in China. With the gradual popularization of phone and internet access among businesses and consumers, more and more studies are now implemented by phone and by web. Face-to-face is, however, still widely used in China and has its merits in at least 2 situations.

When the target respondent is an executive in a large company or a quasi-government official in a state-owned enterprise, establishment of rapport becomes the preeminent factor in determining the success of an interview, and face-to-face interviewers become necessary. These respondents tend to come from a more formal and conservative business background. They demand human contact, and interviewers must employ proper identification, business cards, documents, invitation letters and, most importantly, politeness and respect.

When a study is fielded in the second- or third-tier cities, the norm for data collection and quality control practices for local field providers is face-to-face. Low labor costs in these geographies mean that web and phone surveys offer no cost and time advantage over face-to-face.

Phone interviews. Phone interviews are currently the most prevalent mode of data collection for business-to-business research in China. Most research companies in China are now equipped with a CATI center. A special consideration in China is the need to have local interviewers make phone calls to local respondents; international and even national call centers are not preferred. There are linguistic and cultural reasons for this. In oral communication, accents and other linguistic elements must be considered. The Chinese language is composed of dozens of mutually unintelligible spoken varieties. Though

Putonghua (standard spoken Chinese) is used nationwide, people from different regions speak with easily identifiable accents and local vocabularies. Interviewers calling from Guangzhou, for example, can be detected right away by a Beijing respondent as a “foreigner” which can lead to lower cooperation rates and response error. For social and cultural reasons, respondents sometimes are cautious about calls from strangers in other cities. As a result, a foundation of good telephone interviewing in principle is to use local interviewers as much as possible.

Even more fundamentally, non-local interviewers can face practical language barriers. For example, English product names are sometimes pronounced differently by Chinese people in different regions and might not be understood by non-local interviewers. In prior studies, we’ve seen interviewers spend more than a minute trying to understand a single response due to language “incompatibility.” In some cases, the correct answer is never recorded simply because of this language barrier.

Web surveys. Pure web surveys make use of web panels and other online channels to provide sample frames. These sample sources are rarely verified for target population representativeness and respondent authenticity (see more in the next section). Many researchers have serious doubts about the error inherent in internet panel design in China and are hesitant to do pure web surveys. A common solution for web surveys in China is to separate the sample decision from the data collection decision, often yielding phone-recruit-to-web surveys. Recruitment is performed via phone using traditional, offline sampling sources, and qualified respondents are then directed to the web to complete the interview.

Even so, a related concern in implementing web surveys in China is the impossibility of respondent authentication. One cannot be certain that the recruited respondent is filling out the web survey him-/herself. Proxy respondents, usually subordinates of busy executives and IT managers, are commonly asked by their bosses to fill out surveys. This concern often drives researchers back to the need for some form of face-to-face interviewing to establish the identity and legitimacy of the respondent.

In response to these issues, many quality-oriented research companies in China tend to discourage the use of web surveys. Others exercise more robust quality assurance measures, such as adding a follow-up phone call or administering a proctored web survey, than would be used for other modes of data collection. In the proctored approach, an interviewer meets a respondent face-to-face, authenticates the identity of the respondent, and literally escorts the respondent to administer the web survey. Proctored web surveys can take place either in the respondent’s home or place of business, or the respondent is invited to the research company’s office. This approach offers a high degree of control over sample quality while allowing the researcher to host a multi-country study in a single, web-based mode. But in many ways, proctored web surveys are really not web surveys and sacrifice the time and cost advantages offered by “true” web surveys.

Despite the reluctance of research providers to conduct web surveys, in the last few years, the ineluctable trend—compelled in part by international research standards established in developed markets—is to do research online. As the Chinese research market becomes more mature, we expect more studies will be administered online. And when questionnaires become complex or require visual stimuli, web surveys become more frequent and

advantageous. In the meantime, more stringent quality control and data cleaning procedures are the remedy to quality concerns.

SAMPLE SOURCES

Sample sources generally fall into the following 3 broad categories: customer lists, conventional sources and field provider databases.

Customer lists. In studies that target existing customers of the research client, the sample source is a list extracted from a customer database provided by the client. It is a frame-based sample source from which a probability sample can be drawn. (Read more about probability vs. non-probability sample in the Doxus white paper “A Structured Approach to Choosing and Using Web Samples.”) The precondition is that the list must be updated and accurate, which unfortunately is often not the case in China. When employing customer lists in China, we find it is best to assume a high rate of “non-working” customer information and thus that a large list is needed to yield the target sample size.

Conventional sources.

Phone directories. In China, because of certain criminal incidents in the past, the government does not generally allow a comprehensive listing of phone numbers to be published. Residential phone directories basically do not exist now. Business phone directories are available in most cities; however, these directories represent mainly or only paid listings and exclude phone subscribers who are not advertisers. It is worth noting that adoption of mobile phones has been growing at a double-digit rate and approached 40% penetration in late 2006. However, mobile phone numbers tend to be less listed and trackable as well. Furthermore, cold calls or text messaging that targets the mobile platform is still largely unacceptable to respondents. No appropriate methodologies have yet been created by researchers.

Commercial databases. Commercial business databases are published and sold by business information and intelligence providers such as Dun & Bradstreet. For business-to-business studies, these databases represent important resources to use in developed countries. Given the complicated company organization of state-owned, quasi state-owned and other businesses in China, however, many of these databases have relatively low accuracy for the Chinese market compared to developed markets. Some directories published by companies affiliated with official bodies such as the National Bureau of Statistics or some industry associations are usually biased as they are likely to contain only “related” companies. And the rapid pace of change in the market means these databases are usually outdated by the time they are in wide circulation.

Commercial panels. In China, panels are widely used in fast-moving consumer goods research, less so in other product sectors. This is in part because consumer goods research makes up the bulk of all research activities. Technology research panels and sub-panels do exist, but many of these panels are nascent and growing. Panel providers generally do not disclose the exact procedures or history of panelist recruitment, and the benchmarks of quality panel maintenance (and thus the “representativeness” of the panel) remain murky.

A significant concern in internet and qualitative panel samples is identity fraud. It is commonplace for fake respondents to participate in qualitative research that they are not qualified for and thus earn incentives as their full-time job; sometimes contract researchers participate in a kick-back scheme that benefits both the recruiter and the “participant.” Identity fraud exists in all markets, but given the current Chinese economic and social context, industry organizations such as the China Market Research Association (CMRA) have identified fraud as a more serious problem in China than in most markets.

Field provider panels and databases. As with focus group facilities in other markets, each field provider in China aggregates contacts of people to form a proprietary respondent database that can be used for various research modalities. The way in which this database is built is often unclear to the research buyer, and vendors are often unwilling to provide specifics. Network sampling seems to be the usual mechanism: A field provider accumulates former respondents from projects it has conducted historically or recruits potential respondents via other channels. Sometimes a search is conducted on an ad hoc basis based on each project’s specifications. Because respondents are highly interconnected (A refers B, who refers C and D, etc.), no probability sample can be obtained and selection of respondents into the pool essentially comprises a convenience sample.

Mix and match. In practice, under time pressure and cost considerations, research providers in China tend to draw a sample by combining various sample sources mentioned above. This reduces risk by ensuring that if any one sample source has significant flaws, it is only part of the picture—not the sole source on which an entire project is reliant. While this approach has pragmatic appeal, it raises serious questions. Can we combine sample sources in a statistically acceptable way? If not, will choosing and relying on one particular sample source generate a heavily biased sample? Can we measure how biased this resultant sample is?

CONCLUSION

As the China market research industry approaches its fourth decade of history, we are seeing the co-existence of several common modes of data collection with strong growth in online methods. On the surface, China supports the same sampling and data collection methods used in developed markets, yet the day-to-day “rules” for operationalizing these methods can be completely different. The challenges and opportunities discussed in this paper will continue to face China in the near future, and research practitioners will need to be diligent in updating their understanding of the market and customizing their studies accordingly.