Consumer Market Demographics: Welcome to the Dark Side of Statistics

David A. Swanson

A. Development and Evolution

In the latter part of the 20th century, statistics became a commodity independent of government and a statistical services industry developed. This development is pertinent because these services are primarily a business information industry (Starr and Corson, 1987: 416). There are multiple aspects to the statistical services industry and while there are often less-than-clear-cut distinctions among them, several can be identified, including econometrics, data base development and maintenance, survey research, and market demographics. Of interest to us here is the role of this industry as an entity outside of the government. I focus on the US because of its early leadership in this field. I note, however, that today, many of the major players are international and that our comments on current practices apply beyond the United States.

Two factors significant to the development of market demographics are identified by Starr and Corson (1987: 426): (1) the growing interest on the part of major corporations in market segmentation and targeted marketing, which emerged as mass marketing methods began to falter; and (2) technology developed within the public sector that enabled the association of demographic characteristics with addresses.

B. Development and Evolution of Consumer Market Statistics

Two independent methodological streams can be identified as the methodological origins of the current state of the industry dealing with consumer market demographics: demographic and psychographic. The former is based on demographic information aggregated over

areas (ideally down to very low levels of geography such as postal zip codes, blocks, and block groups). "Psychographics" was initially based on survey data in which consumer preferences, attitudes were segmented by demographic information obtained from respondents in these surveys (e.g., age, race, sex, income, education, household size) to create customer profiles. As the field matured, the customer profiles were matched to summaries of aggregated demographic information to generate market segments (Business Dictionary, no date; Morrison and Smith, 2003; Pentilla, 2005; Russell, 1984).

The demographic stream came directly out of the two factors identified by Starr and Corson (1987) and although it came to prominence in the 1970s, its origin can be traced back to 1950, with the founding of "Market Statistics, Inc. (Pentilla, 2005; Morrison and Smith, 2003; Russell, 1984)

The psychographic stream emerged in the 1970s, but its origin can be found in a 1960 report by Arnold Mitchell, Kenneth Cooper, and Hawkins Stern, "Consumer Values and Demands" that was issued by the Stanford Research Institute (Mitchell, 1983), which was drew upon Maslow's (1943, 1971) "Hierarchy of Human Needs". Ideas from the report were nourished at the Stanford Research Institute (SRI) through the 1960s and 1970s and became known as the "Values and American Lifestyles" (VALS) project. In 1978 a full scale VALS project emerged at SRI, sponsored by 39 corporations (Mitchell, 1983) and, in 1983, yielding the book by Arnold Mitchell, *The Nine American Lifestyles*.

While many the ideas underlying today's practice were around in the early stages, both the demographic and psychographic stream required computing power to be successful. This is why they really started to take off in the 1970s as computers became more powerful and accessible. Critical to the demographics stream was the 1970 US census, which saw the first release of truly electronic files, 1st count through 4th count via a private company, called "DUALABS," which was largely comprised of former Census Bureau employees (Swanson and Stephan, 1984: 785). This, coupled with the computer power of the time, led to directly to "zipcode" demographics providers. Critical to the psychographic stream was the advent of analytic programs such as "conjoint" analysis, which also required substantial number crunching ability (Carroll and Green, 1995). This led directly to companies that were providing customer profiles based on ideas that emerged from the "VALS" project.

Not surprisingly, as these two streams matured, they started feeding into one another. As noted by Sternthal (no date): "...the primary analysis of segments and targets is typically performed using

demographic data. Once a target is selected its demographic character is supplemented by psychographic data to offer additional insight about the goals and dispositions of consumers. This information is often useful in developing a brand's positioning and in the execution of creative strategy."

Although "zipcode" demographics could carry a heavy load in terms of market segmentation and targeting, even when supplemented by psychographics, they could not cover all marketing needs. One of these needs was the ability to segment and target down to the level of the individual household. As was the case for the zipcode and psychographic providers, this development in describing "household characteristics," required a substantial leap forward in data storage and manipulation.

The origins of the "household characteristics" stream can be traced to pre-computer times, but these attempts were very limited. In this stream, credit and other data on households held by private companies were assembled with the goal in mind that there could be 100 percent coverage of all households in the U.S. By the early 1970s, these privately-held household records approached the same number of households counted in the 1970 U.S. census and subsequent post-censal estimates of households made by the U.S. Census Bureau (Tripod, no date; Wikipedia, 2011).

C. Data and Methods

Today, "zipcode" providers rely on census data as the foundation for their work. This holds for those whose data are supplemented by psychographics. In contrast, "household" providers relied heavily on credit reporting services to initially build their records (EPIC, no date; Tripod, no date).

To update their records, household providers still rely on credit reporting services, but their reach has extended to many other sources of data, including purchase data, supermarket savings cards, white pages, telephone usage, surveys, sweepstakes and contest entries, medical records, dental records, insurance policies, financial records, property records, motor vehicle data, automatic number information, credit card transactions, phone records, credit records, product warranty cards, the sale of magazine and catalogue subscriptions, customer loyalty cards - and the United States Postal Service (Epic, no date; Tripod, no date; U.S. Postal Service, 2011; Wikipedia, 2011).

While the commercial sector provides characteristics of interest (e.g., age, race sex, recent purchases, income), it still relies on data from the

public sector. The U. S. Postal Service, for example, is a major source of address updates (EPIC, no date; Pitney Bowes, 2008). This is important in the United States, where around 45 million people (about 15 percent) of the population aged one year and over is estimated to have moved in the past year (U.S. Census Bureau, 2009).

As reported by EPIC (no date):

> The National Change of Address Service (NCOA) is administered by the U.S. Postal Service. Groups that subscribe to the NCOA can obtain updates when a current customer makes a permanent change of address request to the Postal Service. This is the one of the primary methods by which companies obtain individuals' addresses after they move. Businesses can purchase the updates for as little as $5 per thousand names.

As the preceding suggests, the private sector statistics industry in the U.S. relies on the support of the public sector. While the U.S. Postal Service provides the private sector with individual-level information in the form of address changes; the aggregated information provided by the U.S. Census Bureau provides benchmarks against which the private sector data can be compared and adjusted if needed. For example, if a private data vendor finds that its estimate of median household income for a given area is less than that of the Census Bureau's estimate, the vendor can adjust its modelling and other processes to match it.

It also is important to note that the commercial demographics industry emerged in a time in which there were few, if any, restrictions concerning the privacy and confidentiality of information held by the private sector. Even though more restrictions have been placed on the commercial demographics industry in regard to privacy and confidentiality, the industry currently enjoys many fewer restrictions than is found in the public sector.

D. Privacy and Confidentiality[1]

As an example of the "penetration" of the commercial sector into individual lives, an independent market consultant named Lawton Howell (undated) advises his actual and potential clients that every household in the United States can be classified according to one of 66 segments, each with specific demographic characteristics and in this way households can be targeted:

> "In America, all households can be coded for 66 segments. In other words, every household in your marketplace can be

assigned a household demographic code. Once you have your...records coded at the household level, and they have been assigned their specific lifestyle code, you can then reach all the desired segments in your marketplace that has the same code....If you find that the majority of your (targets) are defined with 2-3 of the possible segments out of 66, you can then secure a specific mail list of just the households in your marketplace that match your desired household segment."

In closing, Howell (undated) notes that there are vendors who provide not only the information, but the delivery mechanisms employed by marketers and their clients to reach those individuals and households identified by the segments.

Another independent marketing consultant, Brian Teasley (2004) provides a similar story to that of Lawton Howell in regard to the information on individuals and households maintained by two companies, Acxiom and Experian:

"Acxiom Corp. claims to possess information on 111 million U.S. households, covering 170 million people. As the first figure matches the U.S. Census estimate of the current number of U.S. households, I suspect either Acxiom's numbers are off or the government's are. If I had to bet, I wouldn't discount the accuracy of Acxiom's data....Provide Axciom with your customer list, and it'll return the file with a wealth of information about each customer appended. Variables such as household income, length of residency, education levels, neighbourhood ethnicity, and more are available."

Using the same type of information in a different manner, Acxiom can provide you with a list of names and addresses matching "good prospect" demographic information (income, education levels, etc.). It can limit prospects to a geographic region, including proximity to a given location (e.g. a retail store)....

Experian also deals in household data. The company is known for its financial information. It claims to have data on 110 million households, covering 215 million people. It offers some tailored applications, including products targeted at the auto and financial sectors. For those who wonder where the TRW consumer credit scores went, Experian's got them."[2]

Teasely (2004) concludes his advice to marketers with the note that while aggregate information is useful, household-level information is

needed if one wants to target consumers effectively. To this end, he observes that many sources of household information (such as Acxiom and Experian) allow one to pick specific subsets of names from the databases, based on any available variables.

E. The Dark Side

Given the scope and use of information on individuals and households and American hostility toward "big government" maintaining individual-level information, I find it puzzling that Americans have acquiesced to the fact that there are so few restrictions on the private sector statistics industry in regard to privacy and confidentiality.[3] I suspect that this is due to the ability of these companies to operate largely free of the scrutiny that is characteristics of the public sector, especially the Federal Government. However, the maintenance of detailed information on individuals and households in the United States by the private sector (as well as governmental agencies) is of concern to some at least and attempts to obtain stronger safeguards to privacy and confidentiality on the part of the public are on-going. (El-Badry and Swanson 2007; EPIC undated; Garfinkel 1995; Mayer 2002). With the advent of the "Big Data" era, the capabilities of these firms provide an existing foundation for privacy intrusion that literally defies imagination.[4]

Endnotes

1. Privacy is the idea that it is the right of an individual to decide whether and to what extent he or she will divulge thoughts, opinions, feeling, and facts to the government; confidentiality is the idea that there should be restrictions on how information is collected and used and that no data should be disclosed about a respondent that would allow him or her to be either identified or harmed (Mayer, 2002).

2. "TRW" was an American corporation involved in a variety of businesses, mainly aerospace, automotive, and credit reporting. TRW's roots can be traced to the early 20th century. In 2002, it was acquired by Northrop Grumman and a number of corporations were "spun off, including Experian (Wikipedia, 2012)

3. In large part, awareness of the extent of the private sector records on individuals and households and some government interest can be traced to a 1996 story reported by a Los Angeles TV station when the demographic data vendor, "MetroMail" (a subsidiary of market information giant, R. R Donnelly, Inc., which acquired Metromail in 1987) shipped a list of phone numbers, addresses of 5,000 families

with kids to TV reporter using name of convicted child killer, Richard Allen Davis (EPIC, No date II). Metromail subsequently "merged" into Experian within two years of this report (Laughlin, 1998). This case also contributed to hearings opened in 1997 by the Federal Trade Commission on privacy and confidentiality concerns about private sector data (Los Angeles Times, 1997). However, it is clear that the United States lags far behind many countries (notably in Europe) in having regulations and laws governing privacy and confidentiality on private sector records on individuals and households (Dowling and Mittman, 2009).

4. "Big data" is a term coming into vogue. A description of it is provided by Robert Groves (2012) and a detailed description on its elements and its importance to marketing is found in a report by the McKinsey Global Institute (2011).

References

Business Dictionary. "Psychographics." (available online at www.businessdictionary.com/definition/psychographics.html).

Carroll, J.D., and P. Green. 1995. "Psychometric Methods in Market Research: Part I, Conjoint Analysis." *Journal of Marketing Research* 32 (November): 385-391.

Dowling, D and J. Mittman. 2009. International Data Protection and Privacy Law. Practising Law Institute (available online at www.whitecase.com/files/Publication/367982f8-6dc9-478e-ab2f-5fdf2d96f84a/Presentation/PublicationAttachment/30c48c85-a6c4-4c37-84bd-6a4851f87a77/article_IntlDataProtectionandPrivacyLaw_v5.pdf).

EPIC. No date. "Privacy and Consumer Profiling."(available online at www.scribd.com/doc/20289541/EPIC-Information-on-Consumer-Profiling).

EPIC. No date II. "Largest database marketing firm sends phone numbers, addresses of 5,000 families with kids to TV reporter using name of child killer." (available online at http://epic.org/privacy/kids/KCBS_News.html).

Groves, R. 2012. "National Statistical Offices: Independent, Identical, Simultaneous Actions Thousands of Miles Apart." Online Blog of the Director of the U.S. Census Bureau (available online at http://directorsblog.blogs.census.gov/2012/02/02/national-

statistical-offices-independent-identical-simultaneous-actions-thousands-of-miles-apart/).

Howell, L.no date. "Household Level Marketing." (available online at http://ezinearticles.com/?Household-Level-Marketing&id=3730203).

Laughlin, M. 1998. "Experian, Metromail Work on Merging." (available online at www.dmnews.com/experian-metromail-work-on-merging/article/58957/).

Los Angeles Times. 1997. "FTC Opens Hearings on Public's Right to Online Privacy." (available online at http://articles.latimes.com/1997-06-11/business/fi-2197_1_online-privacy).

Maslow, A H. 1943. A Theory of Human Motivation. *Psychological Review*, Vol. 50: 370-396. http://psychoclassics.yorku.ca/Maslow/motivation.hmt).

Maslow, Abraham H. 1971. *The Further Reaches of Human Nature*. New York: Viking Press.

Mayer, T. 2002. Privacy and Confidentiality Research and the U. S. Census Bureau: Recommendations Based on a Review of the Literature. Research Report Series. Survey Methodology Report #2002-01. Statistical Research Division, U.S. Census Bureau. Washington, D.C. US Census Bureau (www.census.gov/srd/www/byyear.html).

Mckinsey Global Institute. 2011. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. (the full report can be downloaded at no cost at http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation).

Mitchell, A. 1983. *The Nine American Lifestyles*. New York, NY: MacMillan

Morrison, P A and S Smith. 2003. "Business Demography." pages 106-108 in P. Demeny and G. McNicoll (eds.), *Encyclopedia of Population*. Farmington Hills, MI: Macmillan Reference.

Pentilla, C. 2005. "A Brief History of Demographics." *Entrepreneur Magazine* (May) www.entrepreneur.com/article/77152-4#).

Pitney Bowes. 2008. "Pitney Bowes Group 1® Software Unveils Aura™, Industry's Most Comprehensive Address Automation Solution." (available online at www.pbinsight.com/about/newsroom/press-releases/detail/2466_pitney-bowes-group-1reg-software-unveils-auratrade/).

Russell, C. 1984. "The Business of Demographics." *Population Bulletin*: 39 (3).

Washington DC: Population Reference Bureau

Starr, P. and R. Corson. 1987. "Who will have the numbers? The Rise of the Statistical Services Industry and the Politics of Public Data." pp. 415-447 in W. Alonso and P. Starr (eds.) *The Politics of Numbers*. New York, NY: The Russell Sage Foundation

Sternthal, B. no date. "Class Case Studies" (Available online at http://www.kellogg.northwestern.edu/faculty/sterntha/htm/module2/6.html).

Swanson, D. A. and G. E. Stephan. 2004. "A Demography Timeline." pp. 779-786 in J. Siegel and D. A. Swanson (eds). *The Methods and Materials of Demography, 2nd Edition.* Amsterdam: Elsevier Academic Press.

Teasly, B. 2004. "Data Sources, Part 1: Consumer Level." (available online at http://www.clickz.com/clickz/column/1705566/data-sources-part-consumer-level).

Tripod. No date. "Experian (Formerly TRW) Requesting all insider information (ASAP)." (available online at http://nobillsman.tripod.com/experianmain.html).

U. S. Census Bureau. 2009. "Geographic Mobility by Selected Characteristics in the United States ." S070. 2009 American Community Survey. (available online at http://factfinder.census.gov/servlet/STTable?_bm=y&-qr_name=ACS_2009_1YR_G00_S0701&-geo_id=01000US&-ds_name=ACS_2009_1YR_G00_&-_lang=en&-format=&-CONTEXT=st).

U. S. Postal Service (USPS). 2011. "NCOALink® Systems" (available online at www.usps.com/ncsc/addressservices/moveupdate/changeaddress.htm).

Wikipedia. 2011. "Acxiom" (available online at http://en.wikipedia.org/wiki/Acxiom).

Wikipedia. 2012. "TRW" (available online at http://en.wikipedia.org/wiki/TRW).

David A. Swanson, Department of Sociology, University of California Riverside, Riverside, CA 92521, USA.

Email: david.swanson@ucr.edu