Article from

**Long Term Care News**

August 2015
Issue 38

# Soft Data: Another Side of the Story

By David Weinsier, Dean Kerr and Helen Duzhou

## ACTUARIES AND DATA

When actuaries think of data, we traditionally think of hard data. Hard data is information that is directly observed and measured. For example, for long-term care (LTC) insurance, actual reported losses are observed and used to calculate incidence rates.

Hard data is useful and reliable; however, an overreliance on hard data can narrow the risk viewpoint and lead to potentially costly errors.

Actuaries are less familiar with the use of soft data and its applications. Because soft data is not acquired through robust processes, many actuaries may dismiss soft data as insufficiently viable. However, since the 2008 financial crisis, there has been a change in attitudes towards soft data. Utilizing or thinking about soft data allows risk managers and actuaries to broaden their risk perspective, which can help organizations identify and proactively mitigate emerging risks. Using soft data is particularly critical for LTC insurance as there may be limited credible experience based on hard data.[1]

This article will focus on several applications of how soft data can be applied to problems where either the amount of credible (hard) data is inadequate to draw conclusions or the process of data collection has not been appropriately robust.

## CASE STUDY: DEMOGRAPHIC SHIFT IN AGE GROUPS OVER THE NEXT DECADE

One of the questions to be asked concerning LTC insurance is: what is the true demand for LTC insurance? More importantly, what will be the demand for LTC insurance in the future?

An actuary following a purely hard data approach might plot historic premiums over time and extrapolate those values into the future. However, history has proven the past to be an unreliable predictor of the future, as the sales of LTC insurance ballooned in the 1980s and then dwindled in the 2000s.[1]

**HARD DATA** is empirical *information* that has been observed, collected and measured using a systematic process.[10]

**SOFT DATA** is based on empirical *observations*. Unlike hard data, soft data is often not collected through robust processes, and may represent a proxy variable.[10]

**A PROXY VARIABLE** is a variable that is not directly relevant to, but has a close correlation with the true variable being measured.

See *Soft Data Example: A Ship Captain's Journey* to see how soft data may be applied.

## SOFT DATA EXAMPLE: A SHIP CAPTAIN'S JOURNEY

A ship captain sailing from North America to Africa relies on an advanced navigational system to show exactly where the ship is on a world map. One day, the captain discovers that the system no longer works.

When the captain is relying on data gathered from the navigational system, the captain is using hard data that has been systematically collected and measured. However, when hard data is unavailable, the captain must resort to soft data, with a goal of blending together several sources of information in order to navigate the ship.

For example, the captain may navigate by constellations and the sun, weather patterns and cloud formations, and the presence of certain bird species to determine the correct course. The captain may also assign a sailor to watch for any physical obstacles. Finally, the captain may ask other sailors on the ship to recount their traveling experiences to determine the safest route. ∎

**Table 1**

Comparison of hard and soft data[2]

| HARD DATA | SOFT DATA |
|---|---|
| • Quantitative; easy to store and transmit in impersonal ways, such as electronically | • Difficult to store and transmit for future use without understanding what that use will be |
| • Almost always recorded in the form of numbers | • Often communicated in non-measureable ways |
| • Independent of the collection process, meaning it is possible to separate the collection and evaluation (or use) of the data | • Not possible to separate the context under which the data is collected and the collector of the data—the collector is part of the data |
| • Can be collected in an impersonal fashion (has the added benefit of removing bias) | • Although one can potentially create a numerical score for soft data, it must be interpreted in the exact same way by different individuals in order to be considered hard data |

Another actuary may look at this problem as being one of supply and demand. The true demand for care—and related insurance products—has always been driven by humans growing older and falling ill. As baby boomers age, demand theoretically increases. However, the supply of insurance has decreased as claims experience has exceeded expectations and insurers have exited the market. Further, LTC insurance premiums have increased, making policies less affordable for those seeking insurance.

Clearly, the above is a simplistic explanation of the market. In reality, the true demand or supply of LTC insurance is dependent on a myriad of factors. A model reflecting all of these factors is infeasible. Thus, soft data or anecdotal evidence can be used to help frame the answer.

Consider the following: In the United States, 66 percent of long-term care received is informal.[3] In Canada, Statistics Canada reported that nearly 70 percent of long-term care is provided by close family members, while 60 percent of informal long-term care is provided for an aging parent.[4,5] The likelihood of requiring long-term care in Canada is one in ten by age 55, and one in two by age 75.[6] This is certainly useful hard data.

One clear take-away from the above facts is that a common situation occurs where children become the primary caregiver to their parents. Applying a soft data mindset, perhaps this experience will educate individu-

> **INFORMAL CARE** is unpaid care provided by family, friends and volunteers.
>
> **FORMAL CARE** is paid care provided by professionals.

als on the cost and level of care needed, and result in a realization of the value of purchasing LTC insurance (and of the fact that the cost increases by delaying purchase).

Using the above soft data (which, at the core, was drawn from hard data), an insurer may arrive at the conclusion that the aging of the baby boomers will cause an increase in demand for LTC at not only the older but also the younger ages. The insurer may then decide to target this source of demand for LTC, in the form of familial caregivers, and direct marketing activities at this segment of the population.

## USING SOFT DATA TO ANSWER OTHER QUESTIONS

Soft data can be used to provide insight into several other LTC insurance challenges and weave together seemingly disparate sources.

## COST OF LTC INSURANCE BENEFITS

Historically, benefits paid by many insurers have been higher than initially expected. However, looking forward, it is not immediately apparent whether improvements in medicine will a) prevent policyholders from going on claim or enable them to come off claim faster, or b) only be adequate enough

to permit those going on claim to live much longer (but not recover). Thus, in the context of LTC it is not clear to an insurer how to interpret and react to this trend of medical improvements.

Consider the dilemma in the context of Alzheimer's disease, which is a major LTC insurance claims driver. Based on analysis of hard data, Alzheimer's is expected to become more prevalent in the future.[7] Accurate data related to diagnoses of Alzheimer's and other LTC claims drivers benefits an insurer in multiple ways; however, there is currently no single method to definitively diagnose Alzheimer's aside from an autopsy on brain plaque after death.[8] Thus, it would seem to be in an insurer's best interest to keep abreast of progress made in diagnosing this disease. One soft data approach to this problem is to monitor medical ventures and take note of any improvements in diagnosis standards. For example, one company has recently built a team of experts geared towards diagnosing Alzheimer's.[9] Proactively monitoring these advancements using a soft data approach may one day yield additional hard data that can be used to estimate future claims from Alzheimer's cases.

## LEARNINGS FROM OTHER PRODUCTS

A classic example of hard data application is the extrapolation of Canadian Term to 100 insurance lapse experience to predict policyholder behavior on other lapse-supported products, such as Secondary Guarantee Universal Life (SGUL) insurance. Clearly the original collectors

**ENDNOTES**

1. Goldman Sachs, "A Little TLC for LTC," 15 May 2014. [Online]. Available: *http://www.annuitywatchusa.com/long-term-care-insurance-a-little-tlc-with-ltc/.* [Accessed 25 May 2015].

2. M. A. Petersen, "Information: Hard and Soft," Northwestern University, Jul 2004. [Online]. Available: *http://www.kellogg.northwestern.edu/faculty/petersen/htm/papers/softhard.pdf.* [Accessed 31 May 2015].

3. Family Caregiver Alliance, "Long Term Care Statistics," 15 Jan 2015. [Online]. Available: *https://caregiver.org/selected-long-term-care-statistics.* [Accessed 24 May 2015].

4. Long Term Care Canada, "Our Care-Years," 2011. [Online]. Available: *http://www.longtermcarecanada.com/long_term_care_resources/care_years_pg_11.html.* [Accessed 24 May 2015].

5. Statistics Canada, "Long Term Care Facilities Survey," 10 Jun 2014. [Online]. Available: *http://www.statcan.gc.ca/eng/survey/business/5203.* [Accessed 04 Feb 2015].

6. Benefits Canada, "Canadians Ill-prepared for Long-term Care Costs," 22 Jun 2012. [Online]. Available: *http://www.benefitscanada.com/benefits/health-wellness/canadians-ill-prepared-for-long-term-care-costs-29917.* [Accessed 2015 24 May].

7. Alzheimer's Association, "Long Term Care," 2015. [Online]. Available: *http://www.alz.org/join_the_cause_long_term_care.asp.* [Accessed 31 May 2015].

8. National Institute on Aging, "Alzheimer's Disease Fact Sheet," Sep 2012. [Online]. Available: *http://www.nia.nih.gov/alzheimers/publication/alzheimers-disease-fact-sheet#diagnose.* [Accessed 31 May 2015].

9. Amarantus, "Alzheimer's Disease Diagnostics SAB," Feb 2015. [Online]. Available: *http://www.amarantus.com/about/alzheimers-disease-diagnostics-sab.* [Accessed 24 May 2015].

10. A. Samad-Khan, S. Guharay, B. Fischtrom, M. Scanlon and P. Shimpi, "SOA Monograph – A New Approach For Managing Operational Risk, Ch 8.," Chicago, Society of Actuaries, 2009, 2010.

11. Actuarial Standards Board, "Data Quality," Dec 2004. [Online]. Available: *http://www.actuarialstandardsboard.org/pdf/asops/asop023_097.pdf.* [Accessed 24 May 2015].

of the Term to 100 lapse experience data were completely independent from the SGUL actuaries who later evaluated and relied on the data to help predict experience on portfolios of SGUL business.

Like hard data, soft data can also be used to generalize learnings from one type of product to another. A characteristic of many LTC insurance products is that they are highly customizable and non-standardized, featuring a range of benefit and rider options. Using a soft data approach, actuaries can look to other highly customizable products, such as certain Universal Life insurance or Variable Annuity designs, for insight into how in-force blocks of LTC insurance could be modelled and managed.

## ENCOURAGING ACTUARIES TO BE "SOFTIES"

Insurers should encourage actuaries and other staff members to start thinking in terms of soft data in the following ways:

1. Create actuarial working groups and brainstorming sessions, paring down the resulting findings and key learnings, and disseminating the information across the firm.

2. Place different teams working in different functions close together to promote "cross-pollination" of ideas. For example, placing underwriters, marketers, and pricing actuaries together may result in creative new ideas that balance consumer demand and internal pricing and underwriting protocols.

## CONCLUSION AND WORDS OF CAUTION

As with hard data, soft data should not be overly relied upon. Experts advise using soft data to complement hard data, and vice versa.[10] Further, Actuarial Standard of Practice (ASOP) 23 should be consulted: although soft data may not be acquired through robust methods, it should still be reviewed by a qualified actuary and vetted for quality and consistency.[11]

At a minimum, soft data can prove highly useful when defining the boundaries for potential answers to a problem or developing hypotheses. Where the deficiencies in hard data are too broad or complex, soft data can be used to make informed decisions.

*The views expressed are the authors' own and may not represent the views of Oliver Wyman.* ∎

David Weinsier, FSA, MAAA, is a partner with Oliver Wyman and leads the U.S. Actuarial Life Practice. He can be reached at *David.Weinsier@ oliverwyman.com.*

Helen Duzhou, ASA, is an analyst with Oliver Wyman and specializes in long-term care. She can be reached at *Helen.Duzhou@ oliverwyman.com.*

Dean Kerr, FSA, MAAA, ACIA, is a principal with Oliver Wyman and specializes in life, annuity and long-term care. He can be reached at *Dean.Kerr@oliverwyman.com.*