SOCIETY OF ACTUARIES

Article from:

Small Talk

March 2014 – Issue 41

Experience Studies—Big Data for Small Companies?

By Al Klein

What is "big data"? This term has become quite prominent over the last couple of years. I did an Internet search to get the answer. The first website I clicked offered a free report titled "Big Data in Big Companies." I took a pass.

Wikipedia must have the answer. "Big data is the term for a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications." That's a mouthful and not a definition I like either. It sounds like even big companies can't deal with "big data."

SAS Institute, Inc. had a definition I liked: "For most organizations, big data is the reality of doing business. It's the proliferation of structured and unstructured data that floods your organization on a daily basis—and if managed well, it can deliver powerful insights." I like both the beginning, that it is "the reality of doing business," and the end, that "if managed well, it can deliver powerful insights." This to me is what big data is all about and, with these pieces, something that is for small companies as well as big ones. In this article, I am going to explain how you can use experience studies as part of "big data" to help give you "powerful insights" and "manage well" your pricing, profitability and risk, all keys to successfully running and managing your business. This can be done, even with your more limited data, and I will explain how as "it is the reality of doing business." First, I am going to spend time discussing experience studies and some of the issues we find in completing ours. I will conclude with some of the steps you should take beyond the experience study analysis that will provide you with the "big data" boost, without resorting to consumer data that many have come to associate with "big data."

Much of what I am going to say regarding the development of experience studies many of you have heard before, but hopefully it will be a helpful reminder with the upcoming principle-based reserves (PBR) requirements. I plan to explain some of the key elements and considerations in conducting experience studies and how to avoid some of the obstacles you face or will face.

Why Are Experience Studies Important?

Let's start our discussion on experience studies with why a company should do them. There are a number of reasons:

- First, the data needed for PBR is the same as that needed for experience studies, so you will need to do something in this regard if you haven't already. The Smaller Insurance Company Section has done a nice job of preparing you for PBR so I won't spend more time on this subject, but rather focus on experience studies themselves. A link to this information can be found at *www.soa.org/Professional-Interests/Smaller-Insurance-Company/pbr-corner.aspx.* Two other sources of information for you are the slides from a presentation on what you should know about PBR and the March 2013 issue of *Small Talk* has an article on preparing for PBR.

- While we can price products with what we believe to be reasonable assumptions, unless we look at the results through experience studies, we really don't know how well we did. And it is important to know how well we are doing from an experience standpoint because our business must be managed to certain profitability levels for our stakeholders.

- With financial statements, we may learn that changes are needed, but without experience studies, we won't know where to make the changes that are needed.

- With experience studies, we can more quickly make necessary changes before a problem gets out of hand.

How to Conduct an Experience Study

Now let's move on to how to do an experience study. My focus in this article will be on mortality studies. Lapse stud-

ies would be done similarly, but I will point out where there are some differences.

Cleaning the Data

In real estate, you have heard it is all about location, location, location. With experience studies, the first three steps are most important.

1. Clean the data.

2. Clean the data.

3. Clean the data.

Without good data, nothing else matters. I am not talking about credible data here because, as a small company, at least some of your data is likely not to be credible. I will discuss credibility later. What I am referring to here is that you need to make sure the data that will go into the study is accurate and reasonable.

I am responsible for most of the experience studies Milliman conducts, and I can tell you that over half of the time conducting the study is spent working with each contributing company on cleaning their data. And we generally go back and forth with a company at least three times to resolve all of the data issues. I like to look at our work with companies on cleaning the data as an extra benefit of participating in our studies—that is, the companies can fix the errors we find on their systems, helping with the accuracy of their own internal experience studies. We find the errors through a series of programs that look for potential anomalies in the data.

> … over half of the time conducting the study is spent working with each contributing company on cleaning their data.
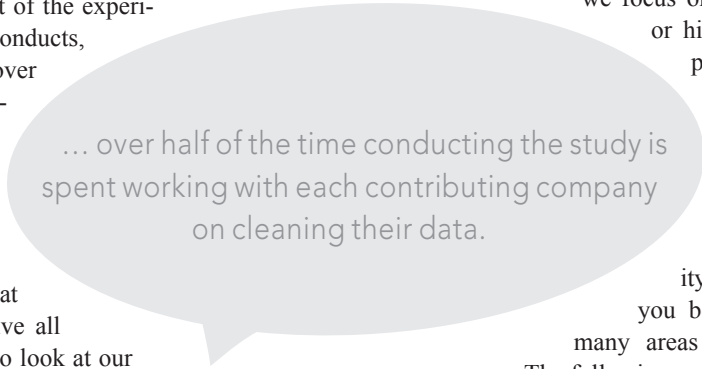
There are a number of things that we look at in this process:

• Check for valid dates.

- The date coming from your system should be a valid date.

- The date of birth should not be after the issue date.

- The termination date should not be before the issue date.

- These are all basic items, but often not checked. The problem with ignoring these or other errors

is that you do not know how the system will react to bad data. It could use it incorrectly and distort the true results without you even knowing it. With the smaller amount of data small companies have, only one or two errors could be enough to seriously distort the results.

- Does a whole block of business all have the same issue date or termination date? Is that because it is a default value? If so, can the correct date be determined?

• Check for valid face amounts.

- If your maximum face amount is $2,000,000 and there is a $15,000,000 policy in the data, is this valid? It may (through an exception) or may not be. Since this one large policy is likely to skew your results, it is important to make sure you verify that it is valid or correct it, if not. In our studies, we focus on every policy $5,000,000 or higher and work with companies to make sure every one of them is correct as we do not want to distort the results of our studies.

• Do some reasonability checks of the data before you begin the study. There are many areas where a problem could occur. The following are three examples.

- If you know that sales have been increasing by a large amount, let's say for the last five years, and at a rate that exceeds the lapse rate, then you should see an increase in business by both issue year and study year for at least each of the last five years.

- If you issue an even distribution of female and male business and you see data that shows 75 percent of the issues were males, a problem likely exists in the data (or maybe the not-taken rate for females is extremely high, which would be worthwhile to investigate). As you can see from this example, you may find issues that go beyond experience studies that need further investigation and action.

- If you primarily issue to the 35- to 55-year-old market, and 25 percent of your issues are at 70, there is likely a problem.

These are just a few examples of the items to look out for. The point of this is to alert you to an activity that should be done rather than to provide you with an exhaustive list of items to review and investigate. Investigation of these issues is not fun and often time-consuming, but it needs to be done in order to have the data to do a study. If you just cannot resolve an issue, my recommendation would be to delete that record from the study, but hopefully you will not have to be deleting too many records. You will need to judge how much time to invest in the investigation of a record or group of records. The answer to this will likely be driven by the level of investigation needed to resolve the issue, how much time you have to devote to it, and when the study is due. In some instances, it makes sense to do the investigation after the study and include the corrected data in the next study. Make sure you document this, especially if it is a whole block of data you are excluding from the study.

One other issue related to the data is the effort being required of companies to search for deaths on some of the older blocks of business in order to pay the beneficiary the death claim, even though the death has not been reported. My understanding is that the findings from this research vary significantly from company to company. One challenge that everyone must face is that the best source for checking for potential claims is the Social Security Death Master File; however, a couple of years ago certain states stopped reporting deaths, making this research more difficult and impossible to fully complete. What does this have to do with experience studies? You may want to recognize that there might be some underreporting of deaths at the older ages and later durations of your business. My suggestion would be to do the experience study as you normally would and then possibly make some adjustment at the end, or at the very least, comment on this issue in your report/documentation.

Completing the Study

Now that you have completed the first three steps and have clean data, it is time to complete the study. Most mortality studies are calculated on an actual-to-expected (A/E) basis. To calculate the A/E ratio, we need to determine the actual and expected exposure.

"Actual" represents the claims. Claims can be by amount or policy count. I say claims because you may have an insured who has three policies with you. When that individual dies, you have three claims, but only one death. Most companies study claims because it is generally time-consuming to determine how many deaths are in the records.

"Expected" can be your pricing assumption, a standard industry table, or something else (e.g., cash flow testing assumptions, a percentage of a standard industry table that better reflects your experience, etc.). Using the pricing assumption allows you to determine how your experience did relative to how it was priced. Using a standard industry table allows you to compare your results to others in the industry. You would do this by comparing the percentage(s) of the standard industry table that you came in at relative to the industry results from an industry experience study. The choice depends on the study, and many companies use more than one basis to look at results.

In a mortality study, exposure is typically determined assuming a full year of exposure in the year of death and a partial year of exposure for any other decrement (e.g., new issue during the year, lapse, surrender, conversion, move to non-forfeiture option, etc.). For a lapse study, we would assume a full year of lapse exposure in the year of lapse and treat death as another decrement. When the decrement is one that is not specifically being studied, an exact calculation of the exposure is used. This follows a Balducci assumption, which is typically used in experience studies. A description of the Balducci assumption is beyond the scope of this article. A good source for a description of the Balducci approach and exposure calculations is *Mortality Table Construction*[1] by Robert W. Batten.

Before I provide a few examples to help you better understand the exposure calculation, there is one other item that needs to be discussed: whether the study should be done on a calendar-year or policy-year basis. The Society of Actuaries (SOA) does its mortality studies on a policy-year basis; however, with PBR there is a move to completing mortality studies on a calendar-year basis. We use a calendar-year basis. Lapse studies are typically done on a policy-year basis.

The following are a few examples to help you better understand the nuances of exposure calculations. The examples assume the mortality study is done on a calendar-year basis and the lapse study is done on a policy-year basis. Lapse exposure for a particular year is determined by the exposure up to the anniversary ending in that year. Let's also assume we want to know the exposure for 2013 and 2013 is the last year of the study.

1. Policy issued April 1, 2005 and still in force December 31, 2013:

    a. Mortality exposure in 2013 would be three months for duration 8 and nine months for duration 9.

    b. Lapse exposure would be 12 months for duration 8 (from policy anniversary April 1, 2012 to policy anniversary April 1, 2013).

2. Policy issued April 1, 2005 and terminated by lapse March 1, 2013 (before the 2013 anniversary date):

    a. Mortality exposure in 2013 would be two months for duration 8.

    b. Lapse exposure would be 12 months for duration 8 due to the decrement being a lapse.

3. Policy issued April 1, 2005 and terminated due to death March 1, 2013 (before the 2013 anniversary date):

    a. Mortality exposure in 2013 would be three months for duration 8 due to the decrement being death and the other nine months of duration 8 exposure being counted in 2012.

    b. Lapse exposure would be 11 months for duration 8 (from policy anniversary April 1, 2012 to the termination by death March 1, 2013).

4. Policy issued April 1, 2005 and terminated by lapse May 1, 2013 (after the 2013 anniversary date):

    a. Mortality exposure in 2013 would be three months for duration 8 and one month for duration 9.

    b. Lapse exposure would be 12 months for duration 8 (from policy anniversary April 1, 2012 to policy anniversary April 1, 2013). The lapse is ignored because it is beyond the policy anniversary in 2013 and lapse exposure in 2013 only is counted until the policy anniversary in 2013.

5. Policy issued April 1, 2005 and terminated due to death May 1, 2013 (after the 2013 anniversary date):

    a. Mortality exposure in 2013 would be three months for duration 8 and 12 months for duration 9. This may appear to be an anomaly because there is more than 12 months of exposure in 2013, but it is just the outcome of the Balducci assumption, commonly used in actuarial practice.

    b. Lapse exposure would be 12 months for duration 8 (from policy anniversary April 1, 2012 to policy anniversary April 1, 2013). The death is ignored because it is beyond the policy anniversary in 2013 and because the study period ends December 31, 2013.

One important consideration is the observation period that should be studied. There is not one answer to this question. You will want the data studied to be current and relevant, but also to cover as much time as possible to provide a more robust database. Generally, five-year studies are good, although shorter or longer studies could also make sense for you. One important consideration is if and when any significant underwriting changes that impact mortality were made. You will likely want to look at the experience issued both before and after that change. Note that these changes often take place in the middle of a year, and it may or may not be feasible to break your study at the point of the change.

A consideration for when to pull the data to be studied is how long it typically takes for your claims to be reported. Claims that happen over the year-end holiday often take longer to report than claims other times during the year. So, if you decide to pull your data mid-January, for example, it is likely you will be missing some of the previous year's claims. We generally like to give companies at least three months after the calendar year before they report the most recent year-end data to us, but more typically six months to allow all or virtually all of the claims to have been reported and put into the system.

Another consideration for you to make is what to do with pending claims. Do you include them with paid claims when calculating "actuals"? The answer to this may be to simply include or exclude all pending claims. Or you could take a more sophisticated approach and determine the percentage of pending contestable claims that are typically resisted and make an adjustment to reflect this. Note that if you take this approach, you will need to stay on top of any company changes with respect to claims practices. Resisted claims should be excluded from the study. They should be treated as if the policy never existed. Otherwise, you will be including extra exposure that has no possibility of a corresponding claim and this would distort results.

Looking at the Results of the Study

Now that you are a master of getting the data in order, how to make the appropriate calculations, and have decided the years you are going to study, let's discuss how to look at the results of the study. I know small companies do not typically have as much data as larger companies, but that doesn't mean that you cannot study your own business and be able to draw some conclusions, including what actions may be needed. I find putting the output in an Excel pivot table to be most useful for analyzing results. Much of the data will need to be grouped to get meaningful results. Some examples include:

• Issue years: Group older business by decade; more recent years in smaller groups or individually.

• Issue ages: Create five or 10 age groupings, and the youngest and oldest ages are typically in a larger group (e.g., 0 to 29, 80+).

- Duration: Durations 1 and 2 could be grouped or shown individually, but should be separated from the other durations because they represent the contestable period. Then durations 3-5, 6-10, 11-15, etc. could be used up to the ultimate period, which would be shown as 16+ or 26+. Note that although the older issue ages often have a shorter select period, they are typically shown in the same format as the younger ages.

- Face amount bands: Group by points where the premium changes, where there are differences in underwriting, etc. You may also want other groupings to be able to study differences in experience.

I will typically start my studying of the mortality results by a single element whether the analysis is for a big or small company. I initially like to look at items such as:

- Study year
- Issue year
- Issue age
- Duration
- Gender
- Risk class
- Policy size
- Product

I look for problem areas within each of these broad groups first, and then I try to determine why there may have been a problem. So, for example, let's say females are showing poor experience and issue ages 45 to 59 are showing poor experience. I would try to dig deeper to find the reason. I will look at cause of death if it is available. If possible, I would review the specific claims involved, if the number of them is limited and the data is available. Let's say there were more diabetes claims than would be expected—I would then look at the underwriting and determine whether or not it is providing sufficient protection with respect to diabetes.

Although, as a small company, you have limited data, I think it is "big" enough to unlock the mysteries of your larger-than-expected claims. Is it too many claims or a few large claims? Through further digging, you should be able to determine the reason and take appropriate action. This further digging is where my definition of big data comes in, enabling you to gain the knowledge needed to take proper action.

Here are a couple of important points to consider in drawing your conclusions:

- If you are using an industry table as your expected and find that female experience has a higher A/E ratio than male experience, it doesn't necessarily mean that your female experience is worse than your male experience. This result may be due to the underlying table rather than your experience. If you see this in your results, check the industry experience that covers the same period you are looking at to see if your results are consistent with those of the industry. This same issue may happen with issue age and duration as well as gender.

- You will need to determine the credibility of your results. Two methods more commonly used and discussed in recent literature are the limited fluctuation method and the Bayesian or Bühlmann method. A description of these approaches and credibility theory in general is beyond the scope of this article. A good source for information on credibility theory is "Credibility Theory Practices" by Stuart Klugman et al. I recommend you study these approaches to determine which of these, or perhaps another method, is best for your circumstances. It should be noted that if the data is not credible, you will need to be careful about the conclusions you draw. However, if it is not credible, the results may still be indicative of good or poor mortality experience. You can sometimes make the results more credible by studying a broader group. For example, you may not have credible results if you look at gender by issue age, but may have credible results if you look at results by gender alone.

Both your reinsurers and consultants can help you with much of the items covered in this article. The SOA and Smaller Insurance Company Section also have resources to help. The experience studies Milliman does are for both small and large companies. One of the benefits of participating in our studies is that you receive a complete experience study based on the data you submit. We will be completing our second final expense study in 2014. If you have any questions on the article or our studies, please feel free to contact me.

Remember, "big data" is within your grasp. Take advantage of it as just described and you will be well on your way to "powerful insights" and a "well-managed" business. ●

ENDNOTE

[1] Batten, Robert W. 1978. *Mortality Table Construction.* New Jersey: Prentice-Hall, Inc.

Al Klein, FSA, MAAA, is a consulting actuary at Milliman, Inc. in Bannockburn, Ill. He can be reached at *al.klein@milliman.com.*