Retirement Income Analysis
with scenario matrices

William F. Sharpe

10. Fixed Annuities

Annuities

The previous chapter briefly discussed the idea of pooling mortality or longevity risk. During our working years, the greatest financial risk is that associated with *mortality* when future incomes are lost. But insurance companies make it possible to pool such risk with others. For example, an insurance policy can diminish the financial impact of early death: if you die prematurely, your heirs will receive a cash payment that may compensate at least in part for lost future income. Dying early has adverse financial consequences. Insurance contracts that pay off when one dies are termed (rather curiously) *life insurance* policies. An insurance company can issue such policies to many people of the same age, pooling their mortality risks and, in effect, making it possible for those who live longer to pay those who die prematurely.

After retirement, the situation is starkly different. Premature death reduces future *cost*, not *income*. The financial risk is that you will have a long life in retirement, not a brief one. Hence it may be desirable to pool *longevity* risk, with people who die prematurely paying those who live to ripe old ages. An insurance policy that provides such risk sharing is typically called an *annuity*. Here is dictionary.com's rather formal definition:

> *"... a specified income payable at stated intervals for a fixed or a contingent period, often for the recipient's life, in consideration of a stipulated premium paid either in prior installment payments or in a single payment."*

The root is the Latin term *annu* or *annus*, meaning *year*, although the insurance companies typically make payments monthly.

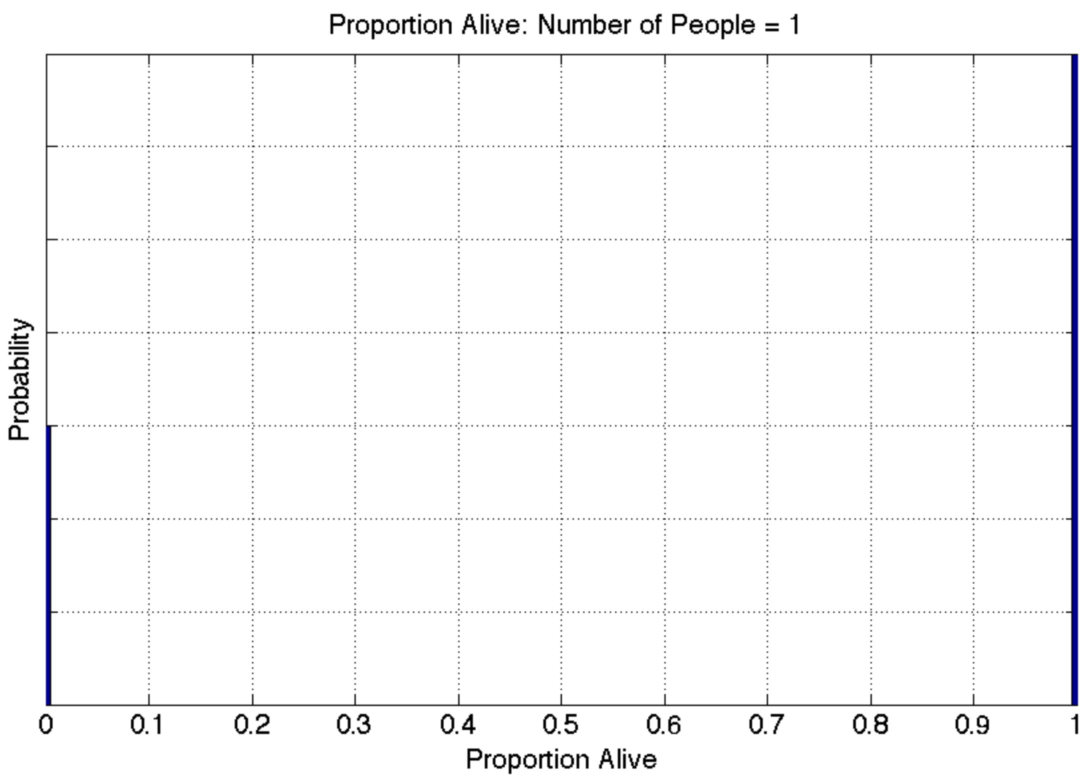
Our focus is on annuities that are paid for with an initial lump sum. Variations include those for which payments begin as soon as possible (*single premium immediate annuities*, or SPIAs) and those for which payments begin at some specified future date (*single premium deferred annuities* or SPDAs). We will not deal with investment vehicles provided by insurance companies, often called variable annuities, that allow for contributions to be made over an extended *accumulation period,* nor with taxes on earnings deferred until payment is made. Our interest is in annuities for retirement – the *decumulation period* of one's life.

As we will see, there are many types of such annuity policies. This chapter and the next focus on policies that provide payments on a fixed schedule, agreed upon in advance, with amounts contingent only on whether one or more named individuals is alive. The amounts paid may be the same each year in either real or nominal terms, or they may vary according to a predetermined schedule (for example, increasing 2% each year). But as long as the real or nominal amounts are specified in advance, such a policy can be considered a *fixed annuity*.

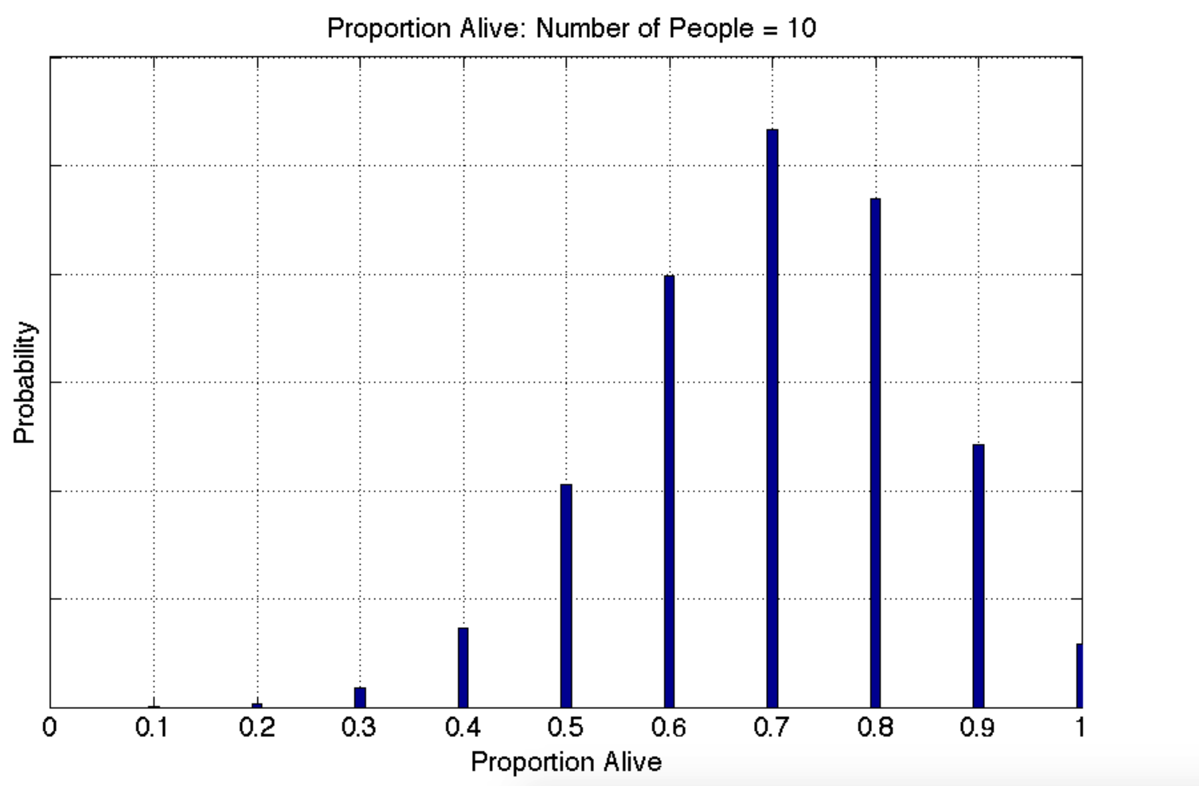
Of course it is possible to pool longevity risk using investments with uncertain returns. Virtually any investment and spending strategy can underlie an annuity, allowing those participating to pool longevity risk while taking investment risk. We will cover examples of such annuities in later chapters. This chapter deals with fixed annuities issued by private insurance companies, the next with those issued by governments.

Pooling Longevity Risk

Risk pooling is a relatively simple concept but it is useful to see how well it can work when longevity risk is concerned. Consider Sue Smith who is now 65 years old. Our mortality tables show that there is a 70% chance that she will still be alive to celebrate her 85'th birthday 20 years hence. This is shown in the figure below, which indicates the probabilities of different proportions alive in 20 years a pool in which she is the only member. Clearly there is great uncertainty about how much money will be required to fund her consumption at the time.

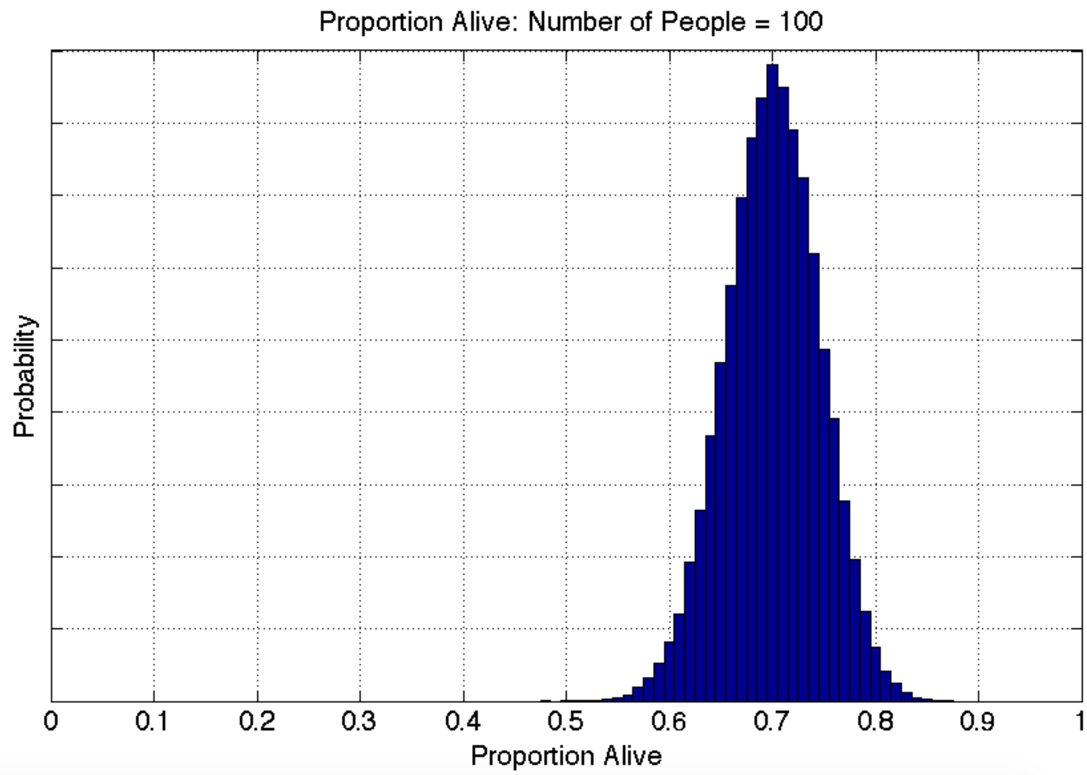


Now let's consider a pool of 10 people, all 65 years old and female, with health characteristics similar to Sue's. We can find the proportions of this pool that might be alive in 20 years by generating a number of scenarios (100,000 here). The next figure shows the results.

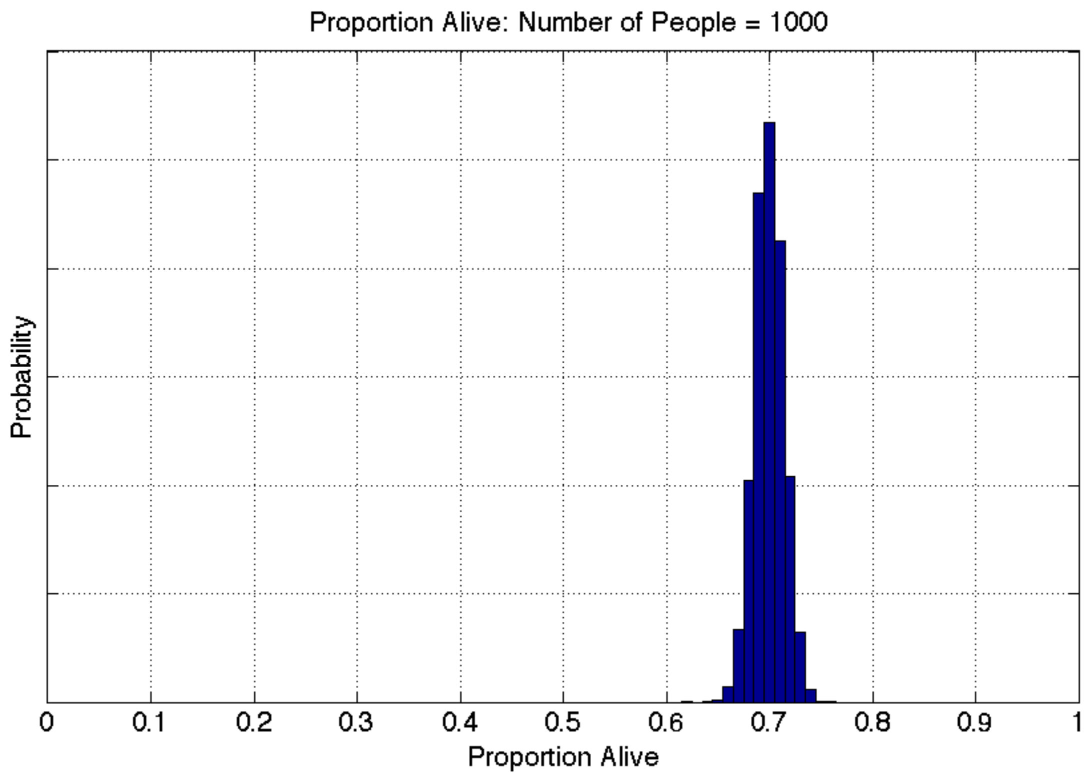


Even with this very small pool, it is clear that there is little need to have enough money to cover the expenses of all 10 people in that future year. To be sure, there is still a wide dispersion of possibilities, but the likely range has been significantly reduced.

The next figure shows the results if 100 people like Sue enter the pool. The range of outcomes is significantly reduced. A pool of money that can finance 80 people would almost certainly cover the needs of the survivors, saving everyone 20% of the amount that would have been required had their longevity risks not been pooled.



One more graph makes the point even more dramatically. In the figure below there are 1,000 people in the pool. Each can contribute say, 75% of the amount needed 20 years hence, knowing that this is almost certain to suffice. The advantages of pooling mortality risk are evident, even with relatively small numbers of participants.



We need not repeat the experiment for periods farther in the future, since the results are easily foreseen. The chance that Sue will be alive 30 years hence is roughly 34%. If she wants to be able to have $X to spend at the time if she lives that long, she will have to put aside an amount that will grow to $X in 30 years. But if she pools her longevity risk with others, only close to 1/3 as much will be needed (plus something for the insurance company). Of course this does not come without a downside. If she saves and invests the entire amount without purchasing an annuity, there is a roughly 2/3 chance that the proceeds will go to her estate. We will have more to say about this subsequently.

Note also that these graphs assume that the underlying mortality tables are correct, which almost certainly will not turn out to be precisely true. We will consider this issue at some length later in the chapter. But for now, we proceed on the assumption that the probabilities reflected in the client personal state matrix are appropriate.

Types of Fixed Annuities

Due to varied personal interests as well as competition among insurance providers, there are many types of fixed annuities. We will consider only the most popular.

A *single life* annuity guarantees income payments to one named person as long as he or she is alive. In contrast, a *joint and survivor* (or *joint life)* annuity guarantees income payments as long as one or both of two persons is alive. Less expensive joint policies may provide higher payments when both parties are alive and lower payments when only one is alive.

If annuity payments begin shortly after a policy is purchased (for example, a month thereafter) the policy is termed an *immediate* annuity. In contrast, a *deferred* annuity provides no payments until a scheduled date that can be years after purchase, with the amount paid dependent on the life or lives of the insured at that time.

Some annuities provide constant payments while the annuitants are alive. Others provide for annual increases of either fixed percentages (e.g. 1%, 2%, etc.) or amounts based on changes in some index of the cost of living. The latter are considerably rarer than the former since they are best backed by portfolios of inflation-adjusted government bonds (TIPS) held by an insurance company, and as shown in chapter 6, there are no such bonds with maturities greater than 30 years as well as some significant gaps in the range of maturities for shorter periods.

Annuity purchasers are sometimes shocked to see the difference in initial income between an annuity with fixed nominal payments and one with fixed real payments. To offset the impact of inflation averaging, say, 2% per year can require a great deal more income in future years, and it is often difficult for people to fully recognize the power of compounding, whether it is beneficial or, as in this case, adverse. For whatever reason, most purchasers of fixed annuities choose to receive payments that are either constant in nominal terms or increase by a predetermined percentage each year. While the payments on the latter are not fixed for all time, the amounts do not depend on any other variable and hence the policies are still classified as "fixed annuities".

Finally, some annuities guarantee that certain amounts will be paid for a *period certain* – a specified number of years – whether the insured are alive or not. In effect, a policy with a guarantee period of n years is like a combination of a portfolio similar to a ladder of riskless bonds with specified payments for n years plusa deferred annuity with payments beginning n years hence. Some insurance companies even offer policies that provide only specified payments for a period of years that are not contingent on anyone's life – in effect a customized bond ladder with a guarantee from the insurance company that payments will be made in full. There is reason to question the desirability of either of these options. One can obtain the same results by investing directly in bond portfolio, withdrawing cash as desired over the corresponding years. This avoids any additional costs that would be charged by an insurance company and provides the option of spending more than planned when and if needed. It would appear that only those with insufficient will power should purchase annuities with guaranteed payment periods. However, for completeness, we provide at least approximations for annuities with such features.

Fixed Annuity Estimates

The traditional way to obtain an estimate of the cost of a fixed annuity is to consult an insurance agent. He or she can obtain information on the potential annuitants' health and other indicators of potential mortality rates, help explain the properties of alternative annuity types, the risks and ratings of different insurance companies, then provide a relatively customized policy. Online alternatives do exist, however, although they may assume adverse selection, with more healthy applicants likely to purchase policies when no information relevant for mortality estimates is given.

Two prominent examples are online quotations provided by Fidelity Investments and Heuler Investment Services. The former can be obtained by anyone, while the latter are typically available for employees of firms with retirement plans that have contracted for the service, although Heuler's quotations are available for those with accounts at Vanguard Investments. The Fidelity quotations are “based on a number of guaranteed, fixed income annuities available through Fidelity” and are apparently not binding. The Vanguard site provides specific quotations from one or more insurance companies that are listed by name, with associated ratings from major rating agencies. One may apparently purchase such an annuity on the stated terms, which include a one-time transaction fee equal to 2% of the deposited amount.

Each of these online systems provides two ways to obtain quotes. In the first, the user provides the amount of money to be used to purchase an annuity and the relative amounts to be paid to each of the covered annuitants; then the program shows the absolute dollar amounts that will be paid. In the second approach, the user indicates the dollar amounts to be received in future years, then the program provides the total cost for the annuity. We will follow the first approach but it is a simple matter to deal with a case of the other type since all the elements can be scaled. Once you know the amounts to be received annually and the total amount invested for a particular set of inputs, the amounts can be multiplied by any desired positive constant to find the parameters for another annuity, assuming that it is priced in the same manner.

Fixed Annuity Data Structures

It is time to provide a data structure to can represent a fixed annuity and the incomes it can provide to retirees such as Bob and Sue Smith. As usual, we break the task into two parts – one to create a data structure with the relevant parameters, the second to process the information in that structure to produce desired outputs.

To reflect the main task at hand, we begin the name of the structure with "i" to indicate the goal: provision of *income (*and we will follow this convention for other income sources). Here is a program to create an *iFixedAnnuity* structure:

```
function iFixedAnnuity = iFixedAnnuity_create( );
   % guaranteed relative or absolute incomes for years 1,...
     iFixedAnnuity.guaranteedIncomes = [  ];
   %  incomes in first post-guarantee year for personal states 0,1,2,3 and 4
     iFixedAnnuity.pStateIncomes = [ 0 .5 .5 1 0 ];
   % graduation ratio of each post-guarantee income to prior post-guarantee income
     iFixedAnnuity.graduationRatio = 1.00;
   % type of incomes (real 'r' or nominal 'n');
     iFixedAnnuity.realOrNominal = 'r';
   % ratio of value to initial cost
     iFixedAnnuity.valueOverCost = 0.90;
   % cost
     iFixedAnnuity.cost = 100000;
end
```

The first element is a vector of guaranteed incomes for years 1, …. These will be paid in every scenario without regard to the personal state of the recipients. The default is an empty vector, indicating no guaranteed payments. If some payments are to be made regardless of the client's personal state, they should be included in this vector, with the first paid at the beginning of year 1 (the present) , the second at the beginning of year 2, and so on. If this vector has n values, the first regular annuity payment will be made at the beginning of year $n+1$.

To represent a deferred annuity with the first payment in year t, one would set the *guaranteedIncomes* element to a vector of $t-1$ zero values.

Our general conventions are that incomes are received and fees paid at the beginning of the year, immediately after the personal state of the client is known (as well as the cumulative market returns and inflation values up to the beginning of that year).

The second data element is the *pStateIncomes* vector. This must have exactly five values, corresponding to the incomes to be received in each of the possible client personal states (0, 1, 2, 3 or 4). All the values are relative amounts, with the actual dollar values to be determined when the data structure is processed. In most cases, the first and last values of this element will be zero, indicating that no payments are to be made in a year in which the last client has just died (personal state 4) or thereafter (personal state 0). For generality, we provide for values for these two cases nonetheless. As can be seen, the default settings indicate a joint and survivor annuity in which the income provided if only one of the two recipients is alive (personal states 1 and 2) equals half the amount paid if they are both living (personal state 3). The actual amounts will be determined when the data structure is processed.

The next element, *graduationRatio*, specifies the ratio of each annuity payment (after any guaranteed payments) to the first such payment. The default value indicates that the amounts paid be the same each year. If it were desired that payments increase by, say, 2% each year this element should be set to 1.02.

Thus far, the values could refer to either real or nominal dollars. The *realOrNominal* element indicates which is desired. If this is set to '*n*' (or '*N*'), all prior values are interpreted as nominal (i.e., not adjusted for inflation). If it is set to '*r*' (or '*R*') all values are considered real (inflation-adjusted). The default case specifies real values, which might better serve retiree's needs than the more common annuities with amounts specified in nominal terms.

The next element specifies the ratio of (a) value of the payments that an annuity can provide to (b) the one-time cost of purchasing it. This is sometimes termed the annuity's *money's worth*. The difference we consider the annuity *fee* and assume that it is paid at the time the annuity is purchased (at the beginning of year 1). The default value is 0.90, indicating that 10% of the annuity purchase cost is to be paid in fees. Thus if the cost of an annuity is $100, the present value of its (contingent) payments is $90. A number of academic studies have attempted to estimate the money's worth ratio of various annuities, with results ranging from 85% to over 95%. We will have more to say about this later in the chapter.

The final element is the total amount to be paid for the annuity at the beginning of year 1 to cover the present value of all promised incomes and fees. The default value is $100,000, although this can be easily changed after a data structure is created and before it is processed.

One final comment about our conventions. Since the annuity is purchased at the beginning of year 1 and income may also be provided in year 1, the net amount initially paid could equal the difference between the two values. In practice, income payments are generally made monthly, with the first received at least a month after the initial purchase cost is paid. Our convention using annual periods is thus an approximation of reality. That said, most adjustments to annuity payments for inflation or graduation changes are made at annual dates rather than monthly, so the use of annual intervals should not produce results that are egregiously oversimplified.

Processing a Fixed Annuity Data Structure

Once the data structure for an income source has been created and elements modified as needed, the structure should be processed, providing two key scenario matrices – one for incomes, the other for fees, and the values in these matrices added to the values in corresponding matrices in the client data structure.

While it may seem straightforward to create matrices of incomes and fees for a fixed annuity, based on a set of values of the elements included in *iFixedAnnuity_create( ),* care is required to insure that the results can be obtained quickly and accurately. Readers not fascinated by the judicious use of matrix operations may wish to skip this section; others may find it interesting and reassuring.

To perform the required calculations we need information from three sources: an iFixedAnnuity structure, a client structure and a market structure. When the function is run, it will update the client incomes and fees matrices as desired. We thus create a function of the form:

```
function client  = iFixedAnnuity_process( iFixedAnnuity, client, market );
  % creates fixed annuity incomes matrix and fees matrix
  % then adds values to client incomes and fees matrices
   ...
end
```

To start, we find the number of scenarios and years for the case at hand:

```
  % get number of scenarios and years
    [ nscen nyrs ] = size( client.pStatesM );
```

Next we create a table with five rows, one for each of the possible personal states (0,1,2,3,4). Each row of the table will contain the incomes to be paid in each of the *nyrs* years. For generality, we start each of the rows with the guaranteed incomes (if any). If anyone is alive (*pState = 1, 2 or 3*) the guaranteed payments are made as planned. But if the last recipient has just died (*pState = 4*), we assume that the remaining guaranteed payments are made as a lump sum. This is likely to be counterfactual, since annuity providers will generally make the guaranteed payments year-by-year, whether the annuitants are alive or not. Moreover, if they were to pay a lump sum to an estate, they might well discount the remaining payments to take interest rates into account. But we wish to hold to the assumption that no payments are made from any income source after the year following the death of the last recipient (that is, personal state 4). To do this, the required payments to the estate are computed by taking the cumulative sum, then flipping it from left to right with the handy Matlab function, *fliplr*. And of course, after the estate has been paid (*pState* = 0) we make no further payments.

After any guaranteed payments have been determined, we add the annuity incomes for each personal state, based on the initial annuity income value given in the *pStateIncomes* vector and a set of multipliers based on the *graduationRatio* variable.

The statements that accomplish all this follow:

```
% make matrix of incomes for states 0,1,2,3 and 4
  psIncomesM = [  ];
  for pState = 0 : 4
    % guaranteed incomes
      if pState == 0
        guarIncomes = zeros( 1, length(iFixedAnnuity.guaranteedIncomes) );
     end;
     if (pState > 0) & (pState < 4)
        guarIncomes = iFixedAnnuity.guaranteedIncomes;
    end;
    if pState == 4
      guarIncomes = fliplr( cumsum(iFixedAnnuity.guaranteedIncomes) );
   end;
   % annuity incomes
     nAnnYrs =  nyrs – length( iFixedAnnuity.guaranteedIncomes );
     gradRatios = iFixedAnnuity.graduationRatio .^ ( 0 : 1:nAnnYrs – 1 );
     annIncomes = iFixedAnnuity.pStateIncomes( pState+1 ) * gradRatios;
   % guaranteed and annuity incomes
     psIncomes = [ guarIncomes annIncomes ];
  % add to matrix
     psIncomesM = [ psIncomesM ; psIncomes ];
  end; % for pState = 0:4
```

The next set of calculations takes advantage of a very useful Matlab function. Assume, for example, that X is a rectangular (two-dimensional) matrix. Then the command ***ii = find(X >5)*** will produce a vector of the locations of all elements in X that exceed 5, treating X as a vector, with column 1 first, then column 2, and so on. And here is the really good part. A subsequent command such as ***X(ii) = 0*** will set each of the selected elements to zero. And the matrix will still be rectangular with its original dimensions.

The following statements take advantage of this capability.

```
% create matrix of relative incomes for all scenarios
  incomesM = zeros( nscen, nyrs );
  for pState = 0:4
    % make matrix of incomes for personal state
      mat = ones( nscen, 1 ) * psIncomesM(pState+1, : );
    % find cells in client personal state matrix for this state
      ii = find( client.pStatesM == pState );
    % put selected incomes in incomes Matrix
      incomesM( ii ) = mat( ii );
  end;
```

First, we create an incomesM matrix with zero values for every scenario and year. Then we process each of the five possible personal states. We begin by creating a matrix ***mat*** with the entire vector of personal incomes for the state in question in every row. Then we find the locations of the all the cells in the client personal state matrix for which the client personal state equals the state being analyzed. Finally, for each such cell, we place the entry in the new matrix ***mat*** in the same location in the incomes matrix for the entire fixed annuity. Short, sweet and very, very fast!

Parenthetically, this approach could be applied to a more complex fixed annuity, such as one with different graduation ratios for alternative personal states. One would only need to produce a different matrix of incomes by year for each personal state (***psIncomesM***), then execute the set of commands shown above.

The next task deals with inflation. Our convention is to state all incomes and other values in real (*inflation-adjusted*) terms. If a fixed annuity provides such incomes, no adjustments are needed. But if the terms are stated as nominal values, we need to convert each of the incomes from a nominal to a real value. In our approach, this turns out to be very simple indeed. Recall that the market data structure includes matrix ***market.cumCsM*** with cumulative changes in the cost of living for every scenario and year. For example, if the value in this matrix for scenario (row) ***i*** and year (column) ***j*** is 1.10, this indicates that it will cost $1.10 to buy goods at the beginning of year ***j*** that cost $1.00 at the present. If a nominal income of $***X*** is to be provided at the beginning of year **j**, its real value in today's dollars will thus be ***$X/1.10.*** Since both incomes and cumulative values of inflation are stated in terms of values available at the beginning of each year, we can simply divide every element in the ***psIncomesM*** by the corresponding element in the ***market.cumCsM*** matrix and voila – we have a matrix of real incomes:

```
% if values are nominal, change to real
  if lower( iFixedAnnuity.realOrNominal ) == 'n'
     incomesM = incomesM ./ market.cumCsM;
  end; % if lower(iFixedAnnuity.realOrNominal) == 'n'
```

Now to determining present values.

First, we compute the present value of all the real incomes in our matrix by multiplying each entry times the present value of a (real) dollar in that scenario and year, summing all the results by row and then summing the resulting values across the columns:

```
% compute present value of all incomes
 pvIncomes = sum( sum( incomesM.*market.pvsM ) );
```

Next we need to create a scenario matrix for fees charged by the annuity provider, then add the calculated value to every entry in the first column. The calculations are straightforward. The data structure provides the annuity cost and the ratio of its value over cost. The product of these two amounts will equal the annuity value. The fee will thus be the difference between the annuity's cost and its value. This amount will, in effect, be paid at the beginning of year 1 regardless of the scenario. The program statements are:

```
% create fee matrix
   feesM = zeros( nscen, nyrs );
% compute value of annuity purchased
   annVal = iFixedAnnuity.valueOverCost * iFixedAnnuity.cost;
% add fee to matrix in column 1
   feesM( :,1 ) = iFixedAnnuity.cost - annVal;
```

Next we need to scale all the incomes so that their present value equals the amount invested minus the fee. The process is straightforward:

```
% scale incomes so pv = amount invested - fee
   factor = annVal / pvIncomes;
   incomesM = incomesM * factor;
```

We finish by adding the fixed annuity incomes matrix to the current client incomes matrix, adding the fixed annuity fees matrix to the current client fees matrix, and then subtracting the annuity cost from the client budget:

```
% add incomes and fees to client matrices
   client.incomesM = client.incomesM + incomesM;
   client.feesM = client.feesM + feesM;
% subtract cost from client budget
   client.budget = client.budget - iFixedAnnuity.cost;
```

As we will see, most retirement plans involve more than one source of income. It is thus desirable to process each of them in turn, with the incomes and fees from each source added to the corresponding client matrices and the cost subtracted from the client budget. Adding the fee and income information to the client matrices has the added advantage of not retaining the large matrices created when income sources are processed, hence saving valuable memory space. This does not preclude the study of the characteristics of a single source of income in a separate analysis, if desired.

The Case Program

Here is the entire Smith case program as it now stands.

```
% SmithCase.m

% clear all previous variables and close any figures
  clear all;
  close all;

% create a new client data structure
  client = client_create( );
% change client data elements as needed
%    ...
% process the client data structure
  client = client_process( client );

% create a new market data structure
  market = market_create( );
% change market data elements as needed
%   ...
% process the client data structure
  market = market_process( market, client );

% Create a fixed annuity
  iFixedAnnuity = iFixedAnnuity_create( );
% change fixed annuity data elements as needed
%   ...
% process fixed annuity and update client matrices
  client = iFixedAnnuity_process( iFixedAnnuity, client, market );
```

Short, sweet and remarkably efficient. On the author's venerable macbook, the entire process took less that 2 seconds. (This is not a misprint!). And any needed changes to data elements could have been made without a significant effect on the run time. Thank you Matlab.

Annuity Prices and Guarantees

Our procedure for estimating the terms for an annuity (the income amounts obtained for a given amount invested) depends not only on the annuitant's age and sex and the terms of the annuity but also on the mortality tables utilized, the assumed capital market returns, and the annuity fee ratio. It is instructive to compare the terms offered by online estimators with those produced by our program to investigate the magnitudes of differences and possible sources thereof. Here are the results of one such experiment.

In late October 2015, an estimate was obtained using the Fidelity Guaranteed Income Estimator for Bob and Sue Smith, then aged 67 and 65 respectively, for a immediate joint and survivor annuity with nominal payments and 50% paid to the surviving beneficiary. The amount invested was set at $100,000. The result was an estimated monthly income of $545 as long as both are alive and $272 if only one is alive.

When these terms were processed using the Smith Case with the iFixedAnnuity.realOrNominal value set to 'n' , the result was a set of incomes that provided roughly $5,760 per year when both were alive. Dividing by 12 gives $ 480 per month (the amounts differed slightly from run to run due to variations in the present values, as described in Chapter 8).

Clearly, $480 per month is very different from $545 per month, so something in the real world differs from one or more of our assumptions. A likely candidate is the riskless real rate of interest, which we have set to 1%. Changing this a few times yielded the conclusion that the estimated joint income would be close to $545 per month if *market.rf* = 1.0225, reflecting a riskless real rate of 2.25%. This, combined with our assumption that expected inflation is 2.0% (*market.eC* = 1.02), implies a total nominal return of 4.25%. Thus if an insurance company could invest in bonds with an expected nominal return of 4.25% or higher, it could cover the cost of the annuity after taking a fee of 10% of the amount invested. And if the investments turned out to return more (or less) than 4.25%, its earnings (fee) would be greater (or smaller).

One might think that providers invest the amounts received from annuity sales in U.S. government bonds with maturities matched to the likely payments required, using standard government bonds for nominal liabilities and TIPS for real liabilities. But this is not the case, as perusal of the financial statements of annuity providers indicates. A study, by the National Association of Insurance Commissioners and the Center for Insurance Policy Research, found that at the end of 2010, U.S. Treasury bonds constituted only 7.5% of life insurers' bond portfolios. Corporate bonds made up 57.1% of the portfolios and mortgage-related securities 20.3%. Other investments included real estate, mortgages, and even some derivative securities. Now, as then, annuity providers hold reserves that are of lower investment quality and promise higher returns than government bonds. Perhaps not coincidentally, on the day when we found our implied nominal return of 4.25% for an annuity quote, the yield on an index of 20-year A-rated corporate bonds was 4.23%.

There is, of course, no free lunch in efficient capital markets. Higher returns tend to go with greater risks. This raises an important question. What might happen if returns on the securities in an annuity providers' portfolio prove to be well below expectations? Could the holders of its annuities be at risk? The answer is possibly, but not probably. There are two main reasons.

First, publicly held insurance companies have issued common stock, so their total assets should exceed liabilities associated with outstanding policies. And privately held issuers undoubtedly have some type of equity capital. That said, this type of buffer is typically quite small. For example, based on numbers from Google Finance, the total market value of the stocks of four major public annuity issuers (Lincoln National, Met Life, Prudential and Voya) on October 28, 2015 was only 5.7% of the total value of their liabilities at the end of the preceding year. Public companies that issue annuities and life insurance are clearly very highly levered. And many annuities are issued by private companies (including 4 of the top ten writers in 2013, according to the 2015 Insurance Fact Book) that may also have small equity cushions.

A second reason why annuities may not be overly risky concerns *state guaranty associations*. If an insurance company experiences “severe financial difficulties” it may be taken over by the life insurance department of the state in which it is based. A policy holder's payments will then be made by the guaranty association in his or her state of residence. However, a caveat is in order. According to information from the National Organization of Life & Health Insurance Guaranty Associations, in most states coverage is limited to $250,000 in present value of annuity benefits. Fortunately, the amounts that had to be paid in such cases have been relatively small. As of 2014, cumulative net costs of annuity payments by the associations over many years reached $3.31 billion, with the greatest amounts paid to residents of California, New York, Pennsylvania and Florida (in that order).

Historically, at least, the risk of not obtaining annuity payments has small in the United States. This may or may not be the case in the future. At the very least, it seems wise to diversify across annuity companies if more than $250,000 is to be invested and to pay attention to the insurance company ratings made by companies such as Moody's, S&P or A.M. Best.

This leaves open the question of how to represent annuities using our programs. It seems undesirable to increase the riskless rate of return. One possibility would be to leave the value-over-cost ratio at the default value of 90% and regard the resulting incomes as those from a a completely riskless annuity. Another approach would be to increase the value-over-cost ratio (perhaps to 1.0 if required) to obtain annuity terms similar to those offered by issuers with excellent grades from rating agencies. We will opt for the former. Fortunately, the choice will not affect some of the qualitative results in subsequent chapters concerning characteristics such as cost efficiency, implied marginal utilities and the like.

Tontine Annuities

Uncertainty about the returns on investments held in insurance company reserves is one source of concern about the guarantees made to the holder of an annuity. But there is another. What if the mortality (longevity) tables used to price annuities turn out to be wrong? Consider, for example, the proverbial "cure for cancer" or some other unanticipated medical breakthrough that extends life. In such an event, annuity issuers might be unable to make all their promised payments, and guaranty associations might be unable to cover any shortfall. Of course there is the other type of possibility: some sort of plague (possibly Methicillin-Resistant Staphylococcus Aureus (MRSA)) might cause people to unexpectedly die earlier, to the benefit of any remaining holders of the shares of annuity companies. We refer to these as sources of (actuarial) "table risk". Such risk clearly exists. But who should bear it, and how?

One answer to the question would be to incorporate in an annuity product a characteristic of a centuries-old gambling instrument called a *Tontine,* resulting in what Moshe Milevsky, a long-time student of the subject, called a *Tontine Annuity* and described at length in his book *King William's Tontine.*

The picture below (taken from Milevsky's book) is of Lorenzo Tonti, after whom the tontine is named.

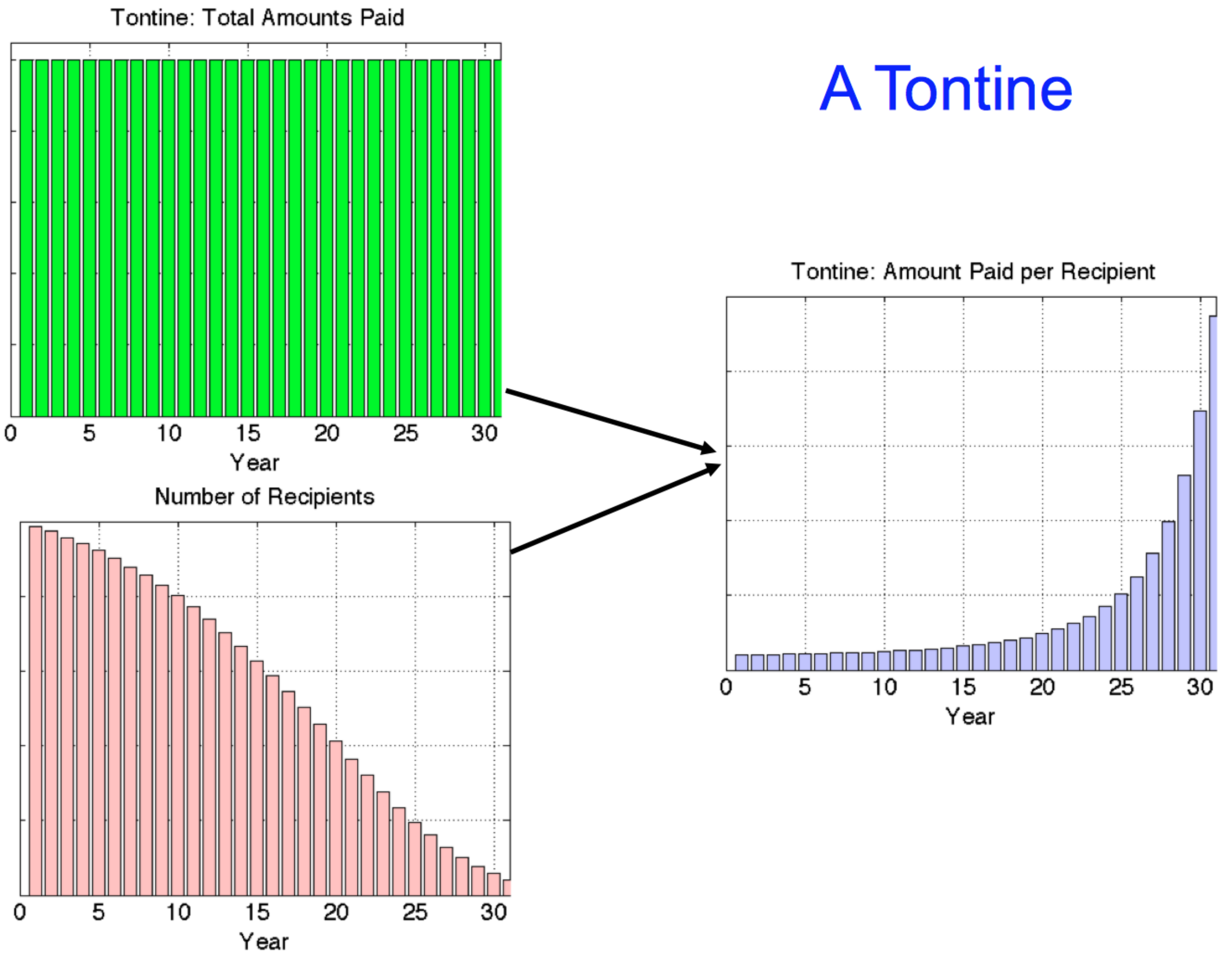
LORENZO TONTI

Tonti's colorful life was documented well by R.M. Jennings and A.P. Trout in their 1982 book T*he Tontine: From the Reign of Louis XIV to the French Revolutionary Era*. From 1649 through 1660, he was *Donneur d'avis* ("giver of advice") to Cardinal Mazarin, who directed much of French policy as Chief Minister of Finance under Louis XIV. In 1652, the king issued royal orders endorsing Tonti's scheme, but the next year the Parlement de Paris failed to approve it. Thereafter, Tonti fell out of favor, ending up in the notorious Bastille prison where he languished from 1668 through 1676. No records have been found indicating the charges against him, but they must have been substantial, since for some of this period his two sons were also held in the Bastille for good measure. After his release from prison, Tonti lived in obscurity, dying in 1684 without seeing his name attached to any financial product.

The first implementation of Tonti's ideas was in Holland in 1670. In 1689, France finally followed suit using the self-explanatory title *Life Annuities with Increased Interest from the Deceased to the Profit of the Survivors*. In 1693, King William III of England passed The *Million Act*, the tontine featured in Milevsky's book, the full title of which is *King William's Tontine, Why the Retirement Annuity of the Future Should Resemble its Past*. And finally in 1696, France issued a set of securities using the name *Tontine*. Thereafter many such securities were created by both private and public issuers.

The idea was simple, although the details were often complex. The issuer would provide equal amounts of total income every year, to be divided among a number of holders of shares. Each share would carry the name of a *nominee.* As long as that nominee was alive, the holder of the associated share would receive a proportion of the total amount paid out. Once the nominee died, the share would be worthless. The issuer would keep paying the same total amount every year until the last nominee died, then stop entirely. In early tontines, the income came from a special issue of government bonds that paid interest until the last nominee died, then expired.

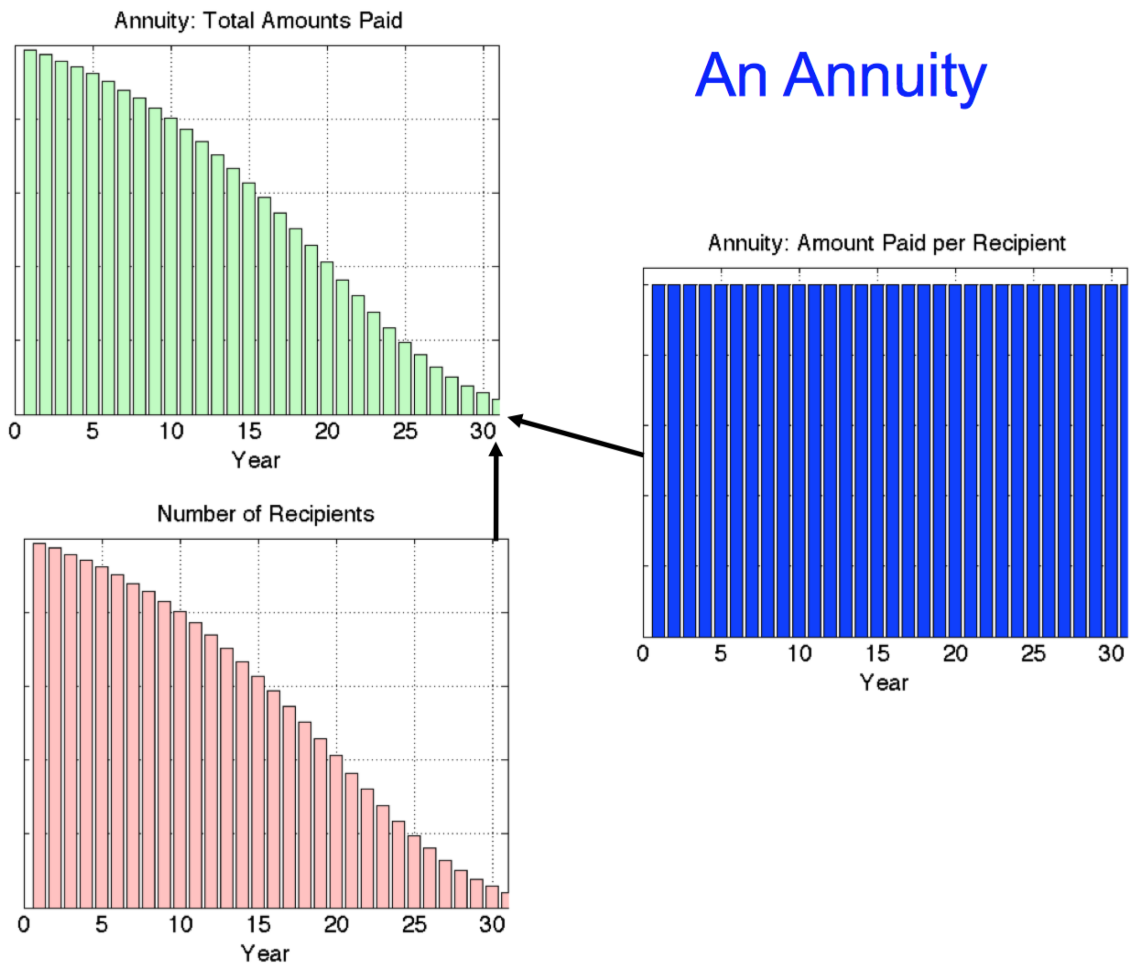
Here is a graphical description of a simple Tontine.



Note that as the number of recipients (holders of shares whose nominees are still living) declines over time, the total amount paid out remains constant but the amount received per share increases. Quite clearly, this is not an instrument designed to help people cope with their personal longevity risks. In fact, the nominees were often strangers or other family members selected for their health and chances of safe and comfortable lifestyles. Milevsky's book lists the names of the nominees for King William's tontine, including one Elizabeth Sharpe who was 11 years old at the time of issue in 1693 and still living in 1730, according to a subsequent list of survivors. The original document shows that she was the daughter of Thomas Sharpe of St. Sepulchers (London), a book binder – a possible relative of the author.

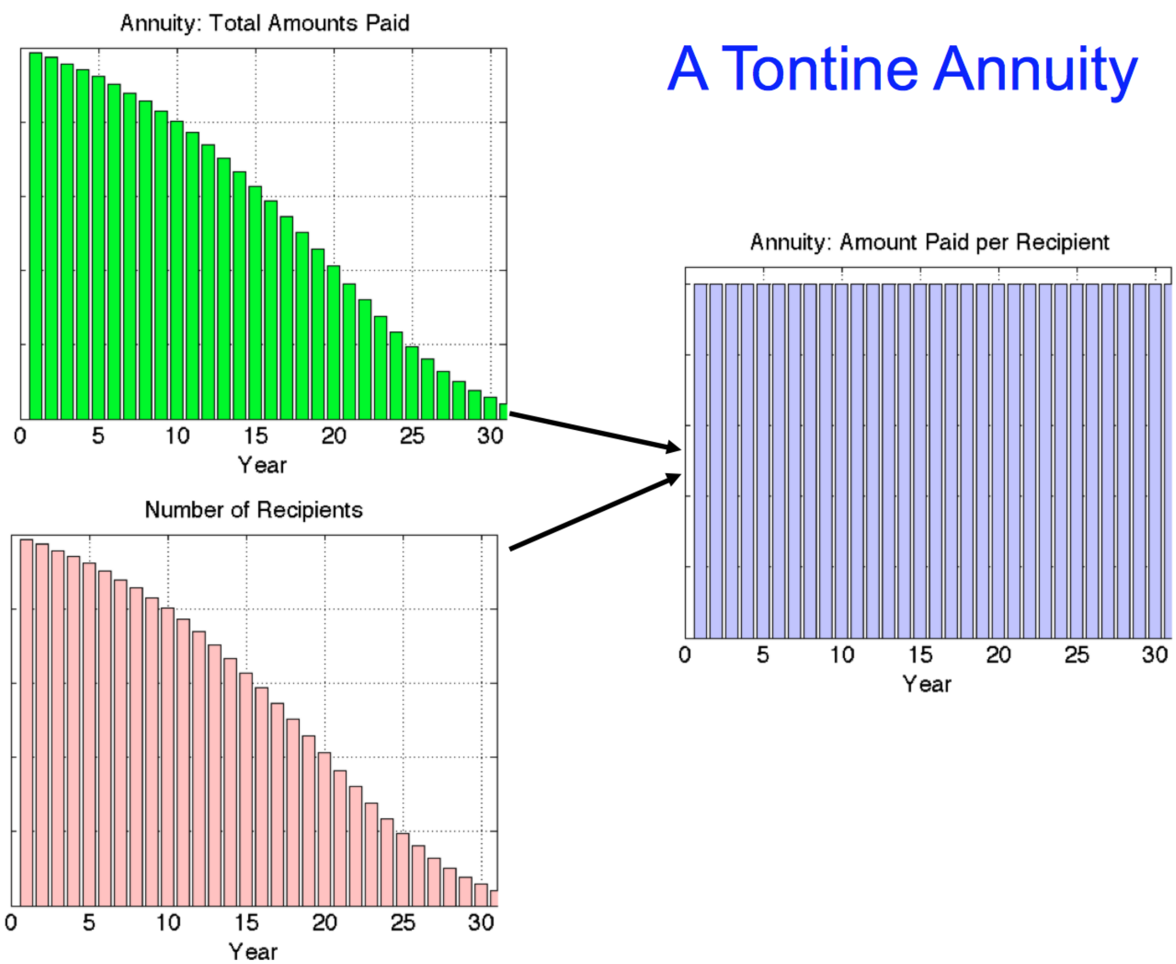
The key aspect of the original Tontine is the fact that the amount paid by the issuer each year is fixed, with the only source of uncertainty being the number of years that payments must be made. Up to that time, any uncertainty about the mortality of nominees affects only the amount received by each of the holders whose nominees are still living. In effect most of the actuarial table risk is borne by the holders of the tontine shares, as indicated by the arrows in the figure above.

Contrast this with the situation for a standard fixed annuity contract, shown below for a particular cohort of annuitants.



In this case, the amount to be paid each living recipient is fixed by contract. The insurance company estimates the (diminishing) number of recipients that will be alive each year, and the results determine the estimated amounts that will need to be paid in each future year. Any differences between predicted and actual mortality rates result in changes in the amounts that must be paid out each year, as the arrows indicate. If the promised payments are to be made to beneficiaries, the insurance company bears the actuarial table risk and must find a way to provide extra funds in the event of an adverse mortality experience.

Enter the idea of a Tontine Annuity, shown in the diagram below. Here the insurer guarantees the total amounts to be paid in each year, shown in the diagram in the upper left. These are designed so that if the number of recipients alive in each year corresponds to the estimates in current mortality tables, the amount received by each annuitant (per dollar of coverage) will be constant from year to year. But if the number of recipients differs from that forecast, the variation will be reflected entirely in the annual amounts paid, as shown by the arrows. In effect, all the mortality table risk is borne by the annuitants. Hence the name, since the contract combines elements of a Tontine with those of a traditional annuity.



There are, of course, a myriad of practical matters that would need to be addressed if such an idea were to be implemented. First, it might be considered a form of gambling and hence be illegal in some jurisdictions. Second, one would have to decide on the exact cohort of annuitants that would share mortality uncertainty. Perhaps all those of a certain age and sex purchasing annuities in a given calendar year might be included in a cohort, but this could require some sort of screening based on health, lifestyles, etc.. Or a larger group, could be included with complex formulas for sharing unexpected mortality experience.

Interestingly, the 1689 French Tontine was divided into *tranches*, each of which was restricted to nominees in a certain age range. The French government followed this procedure in subsequent tontines, as did other issuers. But even this was not completely sufficient. Jennings and Trout describe a group in Geneva that set up a precursor to today's hedge funds, profitably picking nominees for shares of the French Tontine of 1759 based on their own assessments of likely lifespans, then issuing their own shares in a portfolio of tontine shares with sixty different nominees. According to Jennings and Trout, "The Genevan Scheme worked because the speculators took advantage of statistics, as well as advanced medicine and the increase in human longevity resulting therefrom". It relied on "... a sophisticated system of record keeping – notably a store of genealogical data that facilitated research into the histories of families, and in particular, their records of longevity." Attempts to profit from the mispricing of complex financial assets go back a long way.

In a sense, there are existing examples of tontine-like annuities. Milevsky suggests that TIAA-CREF (Teachers Insurance and Annuity Association and College Retirement Equity Fund), which provides investment vehicles and annuities for educators and others, has some tontine-like characteristics. For example, payments received by the author from his TIAA-CREF annuity can vary, based in part on divergences between realized and predicted mortality rates. Here is a 2015 quotation from the TIAA-CREF web site:

> *Payments may also include additional amounts, which TIAA's Board of Trustees may declare each year. While not guaranteed, additional amounts have been applied to income payments every year since 1958. Additional amounts, when declared, remain in effect for the "declaration year" which begins each March 1, for accumulating annuities and January 1, for payout annuities. Additional amounts are not guaranteed for future years.*

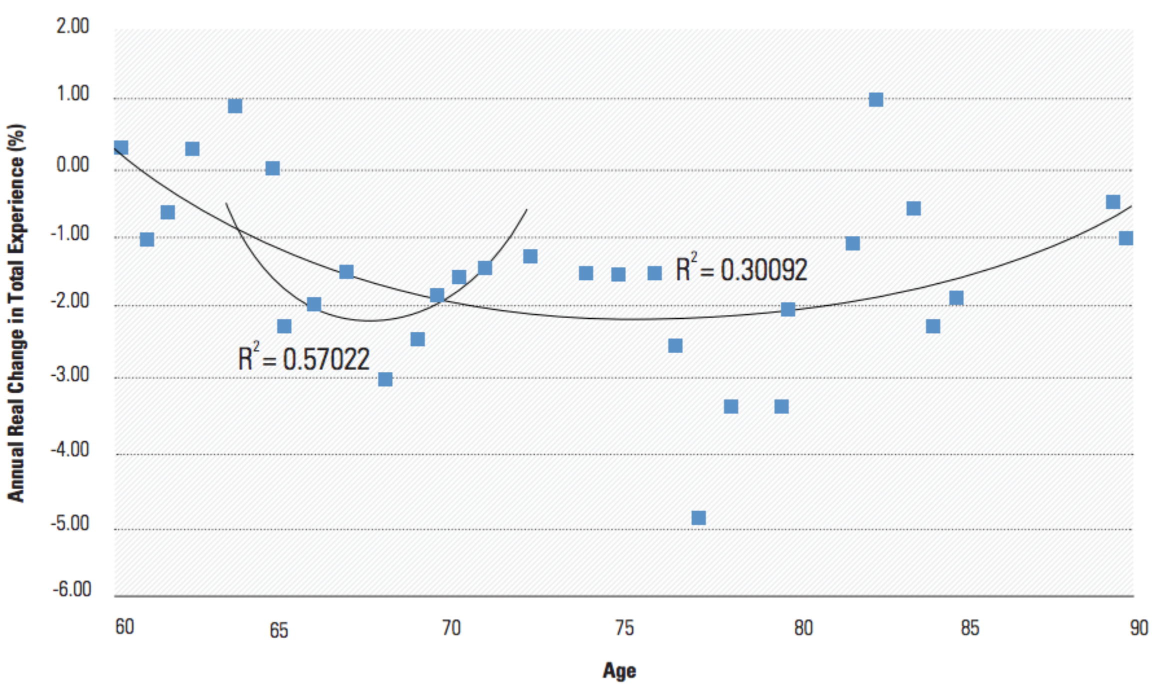
One can also argue that some government-provided annuities have Tontine-like features, with the possibility that benefits may be changed in response to unanticipated changes in the number of those entitled to annuity benefits.

Standard annuities allow for the sharing of individual mortality risk. But if they are to be completely riskless, someone has to bear the risk of major divergences between experienced and predicted mortality rates. Current institutional arrangements provide one way to handle this, perhaps imperfectly. Tontine-like approaches could provide another.

Changes in Real Expenditures during Retirement

Whatever the type of annuity (traditional or tontine), a decision must be made concerning the best affordable pattern of consumption during retirement. A common assumption holds that it is desirable for real consumption to remain relatively constant from year to year. This could be provided by a real annuity with a graduation ratio of 1.0. Alternatively, one could choose a nominal annuity with a graduation ratio of 1+expected inflation (e.g. 1.02 if the expected value of inflation is 2% per year), recognizing that the actual level of real consumption would vary depending on the actual course of inflation from year to year.

While these approaches are often advocated, some empirical evidence suggests that retirees have a tendency to *decrease* their real spending from year to year. In a 2013 study (*Estimating the True Cost of Retirement*), David Blanchett used data on the expenditures of 591 retirees (obtained from the RAND Health and Retirement Study) to find the average annual real change in expenditure for retirees of each age from 60 through 90. His results are shown in the figure below, reproduced from the paper:



Note first that a large majority of the points plot below a 0.00% real change in spending (here, titled *experience*), representing decreases in expenditures. The two curves were obtained by fitting second-order curves via regression analysis – one to all the data points, the other to only those from ages 65 to 75.

There is substantial variation in the points and the fits of the regression equations are relatively poor (with only 30% of the variance of the points for the full sample explained by the longer-term equation). And it appears visually as if a linear regression equation fit to the points for ages over 65 might be relatively flat. But the results suggest that retirees tend to spend less each year in real terms, averaging a decrease of roughly 2%.

Blanchett also broke his sample into four subsets based on spending ratio and net worth. At almost all ages, those in three of the groups decreased real spending from year to year. .Only people with high net worth and low spending rates tended to spend more in real terms as they aged.

There is no reason why these patterns should be optimal for everyone. But they may help explain the popularity of annuities with constant annual nominal payments. We shall analyze these and other types of annuities in the next chapter, which introduces a number of tools for evaluating these and other strategies for providing retirement income.