

Application of Burr XII Mixture Distributions to Model Unemployment Duration in Pricing Unemployment Insurance Assuming USA Data

Richard Onyino Simwa, Martin Mutwiri Kithinji*, Joseph Anthony McOtteku Otieno

School of Mathematics, University of Nairobi, Nairobi, Kenya

Email address:

rsimwa@uonbi.ac.ke (R. O. Simwa), mytwiri@gmail.com (M. M. Kithinji), joseph.otieno@uonbi.ac.ke (J. A. M. Otieno)

*Corresponding author

To cite this article:

Richard Onyino Simwa, Martin Mutwiri Kithinji, Joseph Anthony McOtteku Otieno. Application of Burr XII Mixture Distributions to Model Unemployment Duration in Pricing Unemployment Insurance Assuming USA Data. *International Journal of Statistical Distributions and Applications*. Vol. 2, No. 3, 2016, pp. 27-34. doi: 10.11648/j.ijstd.20160203.11

Received: July 20, 2016; **Accepted:** September 24, 2016; **Published:** October 28, 2016

Abstract: The objective of this research is to consider varying unemployment duration in the pricing of unemployment insurance with application to USA data. The study assumes that unemployment duration follows Burr XII mixture distribution while the discount rate to use in the pricing of the scheme will be determined by fitting market data into the capital asset pricing model. The Burr XII mixture distribution has been used to model unemployment duration in order to allow for heterogeneity in the unemployment duration of the insured employees. The results yield a mean unemployment duration of approximately 16 weeks and premium contribution rate of 5.10% of the taxable wage base per month for a benefit of 45% of the taxable wage base per month payable on weekly basis during spells of unemployment.

Keywords: Burr XII Mixture Distribution, Unemployment Insurance, Capital Asset Pricing Model, Taxable Wage Base, Discounted Cash Flow, Mean Present Value, Premium Rate

1. Introduction

Pricing an insurance cover using discounted cash flow technique involves determining the timing, amount and probability of payment of cash flows together with the discount rate before applying the discounting formula. The timing of benefit payments in unemployment insurance pricing depends on the duration of unemployment. Chuang and Yu [6] used the Weibull distribution to model unemployment duration. The decision is based on the fact that Weibull distribution is commonly assumed in most literature for unemployment duration. However, McDonald and Butler [10] noted that Weibull distribution assumes homogeneity in unemployment duration data which may lead to erroneous results. This paper seeks to improve the model of Chuang and Yu [6] by incorporating the findings in the model by McDonald and Butler [10]. This will be achieved by randomizing the scale parameter of the Weibull to form a mixture of the Weibull distribution. This will in turn allow for heterogeneity in the data.

The benefits will be a fixed percentage of the taxable wage base and the main objective in this paper is to estimate the insurance premium contribution rate to be expressed as a percentage of the taxable income. Mortality will be ignored, so that the probability of a claim depends only on unemployment. The risk-free rate is used to discount the premium income while the risk-adjusted rate, is used to discount the benefits. Capital asset pricing model is used to determine the risk-adjusted rate from the market data.

The premium charged for the specified level of coverage will be determined using the equation of value after calculating mean present values of both the premium income and the benefit stream using the expected present value principle.

According to Malinvaud [9], unemployment insurance is a 'special case' of insurance contract in that in most cases, it is compulsory and is wholly operated by the government. In such a set-up, unemployment insurance is considered as a social program whose main goal is to provide unemployment benefits to partially replace lost earnings for previously

working individuals who become involuntarily unemployed that is, they are able, available and actively seeking for employment. The program specifications differ from country to country. None the less, a common factor in most countries is the way the contributions to the unemployment insurance fund are mobilized. Most of the unemployment insurance schemes charge a flat percentage of the worker's income earned between some minimum and maximum levels. This approach may not be consistent with reality since it does not account for the level of risk a worker is exposed to.

2. Literature Review

In this section, we review the relevant literature on unemployment insurance. In section 2.1, literature on the topic is discussed in general while in section 2.2, the discussion is specifically with respect to United States of America (USA) economy, since the application in this paper is on the USA economic data.

2.1. General Literature on Unemployment Insurance

Some of the problems highlighted by Malinvaud [9] were moral hazard, dis-utility and adverse effects while classifying the risk groups.

Beenstock [2] developed a competitive pricing model to address the above problems by diversifying the unemployment risk and assuming that the unemployment benefits are deterministic. According to the model, the unemployment insurance contract would automatically be enacted when a person starts working and the insured was required to pay premiums right from the onset of their employment. They would then receive unemployment benefits in the event that they become involuntarily unemployed until they secure another job if this occurs before the contract expires. To be able to determine the amount of premiums payable for the cover, Beenstock in [2] assumed that the insurer has identified various risk groups, just as is the case in car insurance, and considered each risk group as a stationary fund. Since the benefits are deterministic, then equating the discounted value of the benefits gives the amount of premiums payable.

Bronars [5] uses Capital Asset Pricing Model to determine the fair premiums in a theoretical model of a hypothetical regulated private market for unemployment insurance. This is an improvement of the existing work by Beenstock [2] where unemployment risk is undiversified and an appropriate risk-adjusted interest rate is specified for the unemployment insurance.

Chuang and Yu [6] extend the results by Bronars [5] by incorporating survival analysis models to estimate the unemployment duration and to calculate the fair premium rate for the unemployment insurance program. In their study they used data from the unemployment insurance program in Taiwan. In the development of the model, the Weibull distribution was used to estimate the average unemployment duration while the capital asset pricing model was used to determine the interest rate used to discount the benefits.

Bowers [3] probed issues surrounding unemployment duration ranging from methodological, measurement and results interpretation of existing statistics on unemployment duration and observes that most unemployment spells are of short durations of less than 10 weeks, although with some fluctuations especially during recessions. The author uses transition probabilities among the three states namely employed, unemployed and not in the labour force to estimate the duration of unemployment. According to their study, the short unemployment durations do not imply an active labour market so that in the event of a job loss, one is able to find his usual type of a job in a relatively short period. This is because a large portion of job changes occurs without any intervening spells of unemployment. More so, the ambiguity in labour force classification, particularly in differentiating between the unemployed and not in the labour force states, is problematic. This is because some of those who withdraw from the labour force experience a brief spell outside and soon enter the labour force as unemployed again.

The sorting model by Salant [12] assumed a constant individual hazard rate which was allowed to vary among different individuals. The constant hazard rate was accounted for by the exponential distribution while the variation among the individuals was accounted for by the gamma distribution. The resulting mixture model, that is Pareto, yielded a decreasing hazard rate for the whole cohort of unemployed individuals.

McDonald and Butler [10] reviewed several mixture distributions of generalized beta distributions. Statistical tests on the model by Salant [12] model in [11] and Burr XII together with the need to conform to job search theory revealed that the Burr XII distribution, a mixture of Weibull and the inverse generalised gamma distributions, is better than Pareto in estimating the spells of unemployment since then there is allowance for heterogeneity in unemployment data.

Cummins [7] points out the problem posed in research on insurance pricing due to parallelism in research on the three major paradigms on insurance. These are statistical modeling, financial modeling and economics. Although few attempts have been made to integrate research in the three areas, the technicality and high specialization exhibited in each have posed a great challenge in the exercise. Cummins [7] made an attempt of integrating the three by considering both statistical and financial models and how they are applied in insurance together with some of the errors made in application. Some of the statistical models, concepts and laws looked at include individual and collective risk models, central limit theorem, law of large numbers and the concept of homogeneity of risks. The financial models explored include the application of capital asset pricing model in determining underwriting rate of return, discrete time discounted cash flow models, option pricing models and sensitivity analysis of the assets and liabilities of the insurance firm. Although economic models were not considered in his integration, Cummins notes that financial models consider insurance variables in an economic setup

which in a way incorporates economic models.

This study seeks to improve Chuang and Yu [6] model by incorporating the findings in the model given by McDonald and Butler [10]. This will be achieved by mixing the Weibull distribution with the gamma distribution to allow for heterogeneity in unemployment duration data. A more appropriate interest rate for discounting and accumulating cash flows is also derived and applied.

2.2. Unemployment Insurance in the United States of America

The Unemployment insurance scheme in the USA is a federal-state partnership based upon federal law. The arrangement is anchored on a strong use of incentives to enhance efficiency. The Federal government ensures conformity and compliance of state programs through Federal Unemployment Tax Act (FUTA) and Social Security Act (SSA). States have enacted their own laws to regulate their individual schemes. The arrangement can be explained by considering separately, the method of financing the scheme, eligibility, unemployment benefits and the waiting period of the scheme.

1) Financing the scheme

The program is entirely funded by employer taxes, both federal and state, although the states of Alaska, Pennsylvania and New Jersey levy unemployment taxes on employees to supplement employer contributions. Unemployment taxes to the federal government have been at a rate of 6% per annum of the first \$7,000 wage base per employee following the decline from 6.2% per annum in July 2011. Tax credit is available up to a maximum of 5.4% of FUTA taxable wages. However, the maximum tax discount is offered to employers who pay their respective state unemployment taxes in full, on time and on all forms of income subject to FUTA tax.

The Unemployment taxes to the federal government are used to pay for administrative costs incurred in the running the Unemployment Insurance programs in all the states together with other associated programs, federal share of extended benefits and to pay for other third tier programs like loans to states with deficits in payment of benefits.

All states finance their Unemployment Insurance programs through taxes from subject employers on the wages of their covered employees. The taxes are deposited into the state's Unemployment Tax Fund (UTF) and are withdrawn by the state to pay the benefits or tax overpayment refunds. Contrary to the federal tax rate, most states use experience rate system to set the tax rate for each employer. However, new employers are given a standard rate before their experience rate is determined. States set their own tax base with some preferring to use the federal government's tax base.

2) Eligibility to the scheme

An application brought to the State unemployment agency is reviewed to determine if the applicant qualifies to receive the benefits. To qualify, one must have worked for the base period or have earned the required wages as provided for in the State's labour laws and the cause of the unemployment

must be out of control of the insured.

Upon commencement of the benefits, one must file weekly or biweekly claims and reports regarding any incomes from work or job offers refused as well as respond to any questions from the state labour office. Additionally, one must report to the Unemployment Insurance Claims office when required to do so.

3) Waiting period

Workers are required to file a claim with the Unemployment Insurance Agency of the state they worked for immediately they become unemployed. During claim, workers furnish the agency with the details of their immediate former employer to aid in authentication of the claim. According to the United States Department of Labour, it takes an average of two to three weeks after a claim is filed for one to receive the first benefits. However, some states take as low as one week to process a claim.

4) Unemployment Benefits

States pay a benefit of between 40%–50% of average monthly earnings in the past one year before unemployment, up to a state's maximum amount. Benefits are advanced on weekly basis up to a maximum of 26 weeks unless in the case of extended benefits during periods of high unemployment.

3. Distribution of Unemployment Duration

To model unemployment duration for the calculation of the insurance premium rates either singular or mixture distributions may be applicable. However, singular distributions assume among other things, homogeneity in the population of study. This assumption can be relaxed by randomizing one of the parameters of the singular distribution to form a mixture. Mixtures allow for heterogeneity hence providing better fits to heterogeneous data than singular distributions.

Let H be a distribution function depending on a parameter θ and g the probability density function of θ . Then

$$F(u) = \int_{-\infty}^{+\infty} H(u, \theta) g(\theta) d\theta \quad (1)$$

is a monotone function of u increasing from 0 to 1 and hence a distribution function. If H has a continuous density h , then the probability density f of u is given by

$$f(u) = \int_{-\infty}^{+\infty} h(u, \theta) g(\theta) d\theta \quad (2)$$

In cases where θ changes discretely, then

$$f(u) = \sum_n h(u, \theta_n) p_n \quad (3)$$

where p_n , the probability of $\theta = \theta_n$ is such that $p_n \geq 0$ and $\sum_n p_n = 1$

In equations (1), (2) and (3) the parameter θ is treated as a random variable and a new probability distribution is defined in the U, θ – plane which acts as our new sample space. Densities of the form in equation (1) and distributions of the form in equations (2) and (3) are generally referred to as

mixtures (See Feller [8]). The Burr XII mixture distribution and the log-logistic mixtures are selected and considered in this case. The general features of the distributions are given in section 3.1, followed by discussion of the applications of the models to unemployment data in section 3.2. Sections 3.3 and 3.4 are on maximum likelihood estimation of model parameters and goodness-of-fit tests on the models, respectively.

3.1. Burr XII, Pareto and Log-logistic Mixture Distributions

In our case, suppose $\theta \sim \text{gamma}(a, c)$ and $h(u, \theta)$ is Weibull($u; b, \theta$). Then it can be shown that

$$f(u) = \frac{ab\left(\frac{u}{s}\right)^b}{u\left[1+\left(\frac{u}{s}\right)^b\right]^{a+1}}; u > 0; a, b, s > 0 \quad (4)$$

with $c = s^b$ where a and b are shape parameters while s is the scale parameter. $f(u)$ is the Burr XII probability distribution.

When $a = 1$, we have a special case of the Burr XII distribution called the log-logistic distribution. On the other hand when $b = 1$, we have a special case of Burr XII distribution called the shifted Pareto distribution.

3.2. Burr XII Specification of Unemployment Duration

In calculating the amount of benefits, we need to estimate the duration of receipt of the claims. The best estimate of this would be the expected value of the unemployment duration. However, due to the incomplete spells of unemployment duration, it would be inappropriate to use a statistical average of the unemployment duration as the estimated duration of receipt of the claims.

In this regard, we will apply survival analysis on the unemployment duration to estimate the duration of receipt of the claims. Let $U \in (0, \infty)$ denote the duration of unemployment and $f(u) = P(U = u)$ denote the probability density function of U . The cumulative distribution function of U is thus expressed as

$$F(u) = Pr(U \leq u) = \int_0^u f(s)ds \quad (5)$$

$$L(a, b, s) = \frac{a^n b^n}{s^{nb}} (\prod_{i=1}^n u_i)^{b-1} \left(\prod_{i=1}^n \left(1 + \left(\frac{u_i}{s} \right)^b \right) \right)^{-(a+1)}; u > 0, i = 1, 2, \dots, n \quad (11)$$

The log of the likelihood function is given by

$$\ln L(a, b, s) = n \ln a + n \ln b - nb \ln s + (b-1) \left(\sum_{i=1}^n \ln(u_i) \right) - (a+1) \left(\sum_{i=1}^n \ln \left(1 + \frac{u_i^b}{s^b} \right) \right) \quad (12)$$

Maximization of the log of the likelihood function gives:

$$\frac{\partial \ln L(a, b, s)}{\partial a} = \frac{n}{a} - \sum_{i=1}^n \ln \left(1 + \left(\frac{u_i}{s} \right)^b \right) = 0 \quad (13)$$

$$\frac{\partial \ln L(a, b, s)}{\partial b} = \frac{n}{b} - n \ln s + \sum_{i=1}^n \ln(u_i) - (a+1) \sum_{i=1}^n \left(\frac{\left(\frac{u_i}{s} \right)^b}{1 + \left(\frac{u_i}{s} \right)^b} \right) \ln \left(\frac{u_i}{s} \right) = 0 \quad (14)$$

And

$F(u)$ is the probability that the unemployment duration is less than or equal to u . For an individual who has been unemployed for t periods, the conditional probability density that he will be reemployed back into the labour force at time u is defined to be the hazard function $h(u)$ and can be written as

$$h(u)du = Pr(u < U \leq u + du | U > u) = \frac{f(u)}{s(u)} \quad (6)$$

where $s(u)$ is the survival function defined as

$$s(u) = Pr(U > u) = \int_t^\infty f(s)ds; s(0) = 1; s(\infty) = 0 \quad (7)$$

As suggested in [6], we assume that the unemployment duration here follows Burr XII distribution when calculating the duration of benefit payment.

It can be shown that the survival and hazard functions corresponding to the Burr XII distribution are

$$s(u) = \left[1 + \left(\frac{u}{s} \right)^b \right]^{-a} \quad (8)$$

$$h(u) = \frac{ab}{u \left[\left(\frac{u}{s} \right)^b + 1 \right]} \quad (9)$$

The duration of receipt of the unemployment benefits is thus equal to

$$d = E(u) = asb \left(\frac{1}{b} + 1, a - \frac{1}{b} \right) \quad (10)$$

With the parameters determined from maximum likelihood estimation of the Burr XII distribution.

3.3. Maximum Likelihood Estimation of Parameters of Burr XII ($u; a, b, s$) Distribution

Let U_1, U_2, \dots, U_n be a random sample of n independent and identically distributed Burr XII ($u; a, b, s$) random variables. As described by Okasha and Matter [11], the likelihood function is given by

$$\frac{\partial \ln L(a,b,s)}{\partial s} = \frac{-nb}{s} + \frac{b(a+1)}{s} \left(\sum_{i=1}^n \left(\frac{(u_i/s)^b}{1+(u_i/s)^b} \right) \right) = 0 \quad (15)$$

Solving the above three equations together, leads to

$$\hat{a} = \frac{n}{\left(\sum_{i=1}^n \ln \left(1 + \left(\frac{u_i}{s} \right)^{\hat{b}} \right) \right)} \quad (16)$$

and parameters s and b can be estimated using numerical methods such as Newton-Raphson.

For Newton-Raphson, an initial guess, say b_0 is assumed to be the solution to equation (14) (note that equation (14) is differentiable.) The approximation is then improved using

$$b_1 = b_0 + \frac{f'(b_0)}{f''(b_0)} \quad (17)$$

The process is repeated iteratively as

$$b_{n+1} = b_n + \frac{f'(b_n)}{f''(b_n)} \quad (18)$$

until a sufficiently better value is reached according to the required level of accuracy.

A similar process is followed to estimate c .

3.4. Goodness of Fit Criterion of Burr XII, Log-logistic and Weibull Distributions

Goodness of fit criterion is used to test the fit of the data into various probability distributions to aid in the selection of the best fit distribution among competing models. These are methods that are used to infer whether a particular data set follows a specified statistical distribution. Suppose X_1, X_2, \dots, X_n are independent and identically distributed samples from an unknown distribution $G_n(x)$. If we wish to check whether at a particular level of significance this sample comes from a particular hypothesized distribution $G(x)$, we test the hypothesis:

H_0 : Data follows a specified distribution

H_a : Data does not follow the specified distribution

Our analysis will incorporate both Aikake's and Bayesian information criterion.

1. Aikake's Information Criterion (AIC):

AIC balances between the complexity of the model and the statistical goodness of fit of the model by imposing a penalty for increasing the number of parameters in the model. It is defined as

$$AIC = -2L(\hat{\theta}) + 2p \quad (19)$$

where $L(\hat{\theta})$ is the maximized log likelihood function and p is the number of parameters in the model.

The preferred model is the one corresponding to the lowest index.

2. Bayesian Information Criterion (BIC)

This is an improvement of the AIC in the sense that BIC

factors in the size of the sample data in determining the amount of penalty to impose on a model due to increased number of parameters. It is defined as

$$BIC = -2L(\hat{\theta}) + 2p \ln(n) \quad (20)$$

where $L(\hat{\theta})$ and p are as defined above and n is the sample size of the data.

As in AIC, the preferred model is the one corresponding to the lowest index.

4. Premium Rate calculation

Equivalence principle or equation of value approach is applied in determining the premium rate for the unemployment insurance. In this approach we equate the Mean Present Value of benefits, to be denoted by $MPV(B)$ to the MPV of Premiums paid to unemployed beneficiary, to be denoted by $MPV(P)$. Necessary assumptions are stated in section 4.1 while equation of value for the computation of the premium rate follows in section 4.2.

4.1. Assumptions

Several assumptions have to be made in order to apply this approach, these include the following.

1. No expenses loading on schemes.
2. The taxable wage base, S payable monthly, will be constant.
3. The entry age into the labour force is 18 years and the retirement age is 65 years in USA.
4. The risk-free rate is assumed to be the rate of return on investment of premiums.
5. A single unemployment spell for the insured.
6. Zero mortality during period of insurance coverage.

They can however be relaxed as the model becomes more complex for better results.

4.2. Equation of Value for the Premiums

According to the equivalence principle for premium calculation in insurance, a fair premium is one that equates the expected present value of the benefits to that of the premium income. For a state premium rate W of the taxable wage base S , the mean present value of the premium income, $MPV(P)$ is given by

$$MPV(P) = W * 12S * \sum_{k=0}^{47} (1 + r_f)^{-k} \quad (21)$$

where r_f is the risk-free rate.

We use the Benefit Event Valuation approach (see [1], [10]) to discount the contingent claims. The mean present value of a monthly benefit, $MPV(B)$, of 45% of the taxable wage base per month, payable weekly during spells of unemployment is given by

$$MPV(B) = \sum_{k=0}^{2444} (1 + r_b)^{-\left(\frac{k+m}{52}\right)} \times q_k^u \times \left\{ \frac{(0.45-W)}{4} S \left(\sum_{t=0}^d (1 + r_b)^{-t/52} \right) \right\} \quad (22)$$

where k is the number of weeks since becoming involuntarily unemployed, m is the waiting period after applying for the unemployment benefits, q_k^u is the probability of a claim in week k , 52 is the number of weeks in a year and r_b is the expected risk-adjusted rate of return.

From the Capital Asset Pricing Model (CAPM), r_b is given by

$$r_b = E(r'_b) = r_f + (r_m - r_f)\beta_u \quad (23)$$

where r'_b is the corresponding random rate of return and r_m is the expected market rate of return. β_u , denotes the correlation between unemployment rate and the market rate of return, and is given by:

$$\beta_u = \frac{\text{cov}(\text{unemployment}, \text{market})}{\text{var}(\text{market})} = \frac{\text{cov}(r_u, r_m)}{\text{var}(r_m)} \quad (24)$$

In this case r_u is the rate of unemployment. W is determined by equating the right hand side of equations (22) and (23).

5. Application and Results

5.1. Data

The study uses secondary data from the United States of America retrieved from the Federal Reserve Bank of St. Louis and Robert Shiller's online data (see the data links specified in [13] – [17]). It comprises of the mean monthly unemployment duration from January 1948 to February 2016 together with the number of first payments of unemployment insurance benefits and the corresponding covered jobs per month. This is as reported by the United States bureau of labour statistics but retrieved from Federal Reserve Bank of St. Louis. The unemployment duration in this case is the number of weeks a worker has been involuntarily unemployed. The other components of the data are the annual return on three-month treasury bills, retrieved from the Federal Reserve Bank of St. Louis; and Standards & Poors 500 Price Indices as reported by Robert Shiller in his online data.

5.2. Goodness of Fit of the Distributions on Unemployment Duration Data

In order to determine the appropriate unemployment duration, we need to establish which distribution best fits the data.

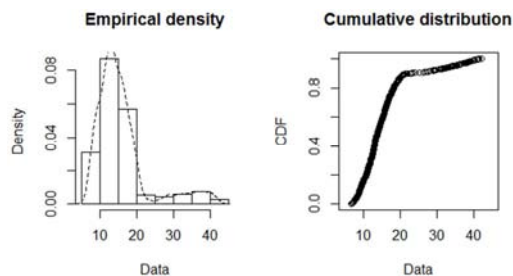


Figure 1. Empirical pdf and cdf.

An empirical plot of both the density and distribution functions of the raw unemployment duration as shown in Figure 1 indicates that the data follows one of the tailed distributions.

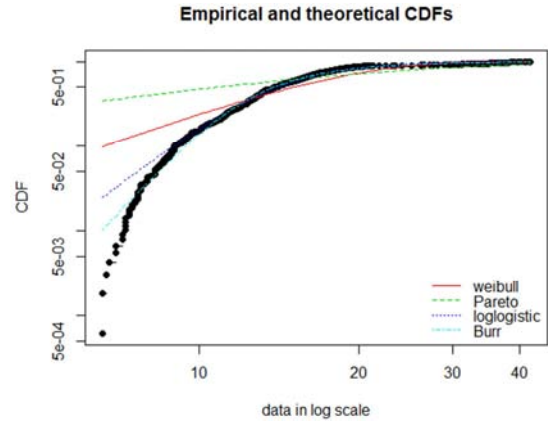


Figure 2. Empirical and theoretical cdfs.

A plot of both empirical and theoretical cumulative distribution functions of Weibull, Pareto, log-logistic and Burr XII of distributions shown in Figure 2 indicates that both Burr XII distribution and log-logistic provide better fits than the Weibull and Pareto.

Table 1. Goodness of fit criteria.

	Weibull	log-logistic	Pareto	Burr
Akaike's Information Criterion	5436.162	5099.712	6158.474	5057.250
Bayesian Information Criterion	5445.581	5109.130	6167.893	5071.378

From the goodness of fit indices outlined in Table 1 the Burr XII corresponds to the lowest indices under both AIC and BIC and is therefore the preferred model in estimating the duration of unemployment.

5.3. Estimation of Parameters

The parameters of the Burr distribution were then estimated using Maximum likelihood estimation in R programming. The results are as outlined in Table 2. The R codes used appear in Appendix 1.

Table 2. Estimation of parameters.

Parameter	Estimate	Standard Error
Shape 1 (a)	0.4955088	0.050208998
Shape 2 (b)	6.6921700	0.390146800
Rate (1/s)	0.0853068	0.002159988

The mean unemployment duration (d) is equal to the expected value of the Burr distribution, and from the calculations,

$$d = 15.7598 \text{ weeks}$$

r_f and r_m are average returns of treasury bills and the market respectively where individual entries of r_m are calculated as the annual rates of change of the S&P500 index

together with the associated dividends. From the analysis we have:

$$\begin{aligned} r_f &= 4.21\% \\ r_m &= 12.15\% \\ \text{cov}(r_m, r_u) &= -0.00013 \\ \text{var}(r_m) &= 0.025757 \end{aligned}$$

These estimates for covariance of the market return and unemployment rate; and the variance of the market return in Equation (24), the beta of unemployment, β_u is equal to -0.01187. Estimates for r_f , r_m and β_u yields, on using Equation (23), a risk-adjusted rate of return r_b equal to 4.11%. We assume that the waiting period, $m = 2$ weeks which is the average of what USA states take to process unemployment benefits.

q_k^u is estimated from the data and is assumed to be constant. It is computed as the average proportion of successful claims (first payments) to the total number of jobs insured at the time of claim. From analysis, $q_k^u = 0.007071728$, for all $k = 1, 2, \dots, 2444$ weeks.

Equating the right hand side of Equation (21) to the right hand side of Equation (22) and replacing for the estimates of the parameters obtained in section 5.3, yields a premium rate $W = 5.10\%$ of the taxable wage base per month.

6. Discussions, Conclusion and Limitations

6.1. Discussions

From equation (4), the scale parameter of the Weibull distribution was randomized by allowing it to follow the two parameter gamma distribution thereby accommodating heterogeneity in the data. This has led to a better fit of the unemployment data as illustrated in section 5.2. Burr XII Mixture distributions therefore, provide better fits for data relating to a population with varying unemployment duration. The mixture distribution incorporates characteristics of the various subjects that are otherwise ignored when using singular distributions, like the Weibull distribution.

The hazard function in equation (9), $h(u)$ is an increasing function of duration of unemployment duration, u . This coincides with the actual experience since we expect the probability of reemployment to decrease as the spell of unemployment lengthens.

Both the beta and covariance of market rate of return and unemployment rate are negative. This is expected because, when the market is doing well, that is, market rate of return is high; we would expect the rate of unemployment to be decreasing due to increased investments, hence a negative beta.

The premium rate obtained in the study, $W = 5.10\%$ of the taxable wage base per month, is within the range of what most states in USA recommend.

6.2. Conclusion

According to the study, states in USA should recommend premium contribution rates that are approximately 5.10% of the taxable wage base to ensure solvency of their

Unemployment Compensation Trust Fund.

6.3. Limitations of the Study

The assumption of zero expenses and zero mortality among insured employees may not hold in practice. The latter increases the mean present value of the benefits, income of unemployed beneficiaries, leading to a higher premium rate while the former reduces the mean present value of the benefits, resulting in a lower premium rate. However, the assumption of zero expenses is in line with US Unemployment insurance program since administrative expenses for the scheme are paid for by the federal government.

Also the model assumes to some extent, the USA economy and therefore some adjustments should be made before applying it to another economy.

The study was limited to only Burr XII distribution in modeling unemployment duration. Further research could consider application of other mixture distributions.

Acknowledgements

The authors are grateful to the African Mathematics Millennium Science Initiative (AMMSI) for the financial support towards this research. We also thank those who read and made several comments about the initial draft of this paper.

Appendix

Appendix 1: R code used in analysis

```
ud<- read.table("D:/studypacks/project/Unemployment
duration not seasonally adjusted.txt", header=T)
attach(ud)
plotdist(ud$VALUE, histo = TRUE, demp = TRUE)
ud1 <-ud$VALUE
ud.W<- fitdist(ud1, "weibull")
ud.P<- fitdist(ud1, "pareto", start = list(shape = 1, scale
= 500))
ud.ll<- fitdist(ud1, "llogis", start = list(shape = 1, scale =
500))
ud.B<- fitdist(ud1, "burr", start = list(shape1 = 0.1,
shape2 = 0.1, rate = 0.1))
cdfcomp(list(ud.W, ud.P, ud.ll, ud.B), xlogscale = TRUE,
ylogscale = TRUE, legendtext = c("weibull", "Pareto",
"loglogistic", "Burr"))
qqcomp(list(ud.W, ud.P, ud.ll, ud.B), xlogscale = TRUE,
ylogscale = TRUE, legendtext = c("weibull", "Pareto",
"loglogistic", "Burr"))
ppcomp(list(ud.W, ud.P, ud.ll, ud.B), xlogscale = TRUE,
ylogscale = TRUE, legendtext = c("weibull", "Pareto",
"loglogistic", "Burr"))
summary(ud.B)
gofstat(list(ud.W, ud.P, ud.ll, ud.B), fitnames =
c("weibull", "Pareto", "llogis", "Burr"))
#PARAMETERS AND EXPECTED VALUE
a<-0.4955088; b<-6.6921700;
```

$$r < -0.0853068$$

$$d < -a * (1/r) * (\text{factorial}(1/b) * \text{factorial}(a-1-(1/b))) / \text{factorial}(a)$$

References

- [1] Actuarial Education Company (2013), *ST4: "Pensions and other Benefits Specialist Technical"*, The Institute and Faculty of Actuaries.
- [2] Beenstock, M. (1985). Competitive Unemployment Insurance Pricing. *Geneva Papers on Risk and Insurance*, 23-31.
- [3] Bowers, N. (1980). Probing the issues of unemployment duration. *Monthly Labor Review*, 103(7), 23-32.
- [4] Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., & Nesbitt, C. J. (1997). *Actuarial Mathematics* (Illinois: The Society of Actuaries).
- [5] Bronars, S. G. (1985). Fair Pricing of Unemployment Insurance Premiums. *Journal of Business*, 27-47.
- [6] Chuang, H. L., & Yu, M. T. (2010). Pricing Unemployment Insurance—An Unemployment-Duration-Adjusted Approach. *Astin Bulletin*, 40(02), 519-545.
- [7] Cummins, J. D. (1991). Statistical and financial models of insurance pricing and the insurance firm. *The Journal of Risk and Insurance*, 58(2), 261-302.
- [8] Feller, W. (1968). *An introduction to probability theory and its applications: volume I* (Vol. 3). London-New York-Sydney-Toronto: John Wiley & Sons.
- [9] Malinvaud, E. (1985). Unemployment insurance. *Geneva Papers on Risk and Insurance*, 6-22.
- [10] McDonald, J. B., & Butler, R. J. (1987). Some generalized mixture distributions with an application to unemployment duration. *The Review of Economics and Statistics*, 232-240.
- [11] Okasha, M. K., & Matter (2015), M. Y. ON THE THREE-PARAMETER BURR TYPE XII DISTRIBUTION AND ITS APPLICATION TO HEAVY TAILED LIFETIME DATA. *Journal: JOURNAL OF ADVANCES IN MATHEMATICS*, 10(4).
- [12] Salant, S. W. (1977). Search theory and duration data: a theory of sorts. *The Quarterly Journal of Economics*, 39-57.
- [13] Board of Governors of the Federal Reserve System (US), 3-Month Treasury Constant Maturity Rate [DGS3MO], retrieved from FRED, Federal Reserve Bank of St. Louis <https://research.stlouisfed.org/fred2/seriesDGS3MO>, June 16, 2016.
- [14] Standards and Poors data: www.econ.yale.edu/~shiller/data/ie_data.xl
- [15] US. Bureau of Labor Statistics, Average (Mean) Duration of Unemployment [UEMPMEAN], retrieved from FRED, Federal Reserve Bank of St. Louis <https://research.stlouisfed.org/fred2/seriesUEMPMEAN>, June 16, 2016.
- [16] US. Employment and Training Administration, Initial Claims [ICNSA], retrieved from FRED, Federal Reserve Bank of St. Louis <https://research.stlouisfed.org/fred2/seriesICNSA>, June 16, 2016.
- [17] US. Employment and Training Administration, Covered Employment [COVEMP], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/COVEMP>, June 30, 2016.