A Multivariate Kernel Density Estimator for the Joint Distribution of Bequest Recipients *

Jason DeBacker† Richard W. Evans‡ Parker Rogers§

April 2016

(version 16.04.a)

Abstract

We provide an algorithm and code for estimating a flexible joint distribution of bequest recipients in the United States by age and by net worth. Estimating this joint distribution is important in macroeconomic models with population demographics in which agents give bequest and the population is heterogeneous in terms of age and earnings ability. We use a multivariate kernel density estimator with a flexible bandwidth parameter. In addition we provide a method with accompanying code to produce the resulting discrete joint distribution over arbitrary age and lifetime income bins.

keywords: multivariate kernel density estimation, bequests, inheritances, intervivos transfers, Survey of Consumer finances.

JEL classification: C14, C63, D31, D91, E21

*We are grateful to the BYU Macroeconomics and Computational Laboratory and to the Open Source Policy Center at the American Enterprise Institute for research support on this project. All Python code and documentation for the computational model is available at https://github.com/rickecon/OGbequests. We also thank James McDonald for some helpful suggestions.

†Middle Tennessee State University, Department of Economics and Finance, BAS N306, Murfreesboro, TN 37132, (615) 898-2528,jason.debacker@mtsu.edu.

‡Brigham Young University, Department of Economics, 167 FOB, Provo, Utah 84602, (801) 422-8303, revans@byu.edu.

§Brigham Young University, Department of Economics, 121B FOB, Provo, Utah 84602, parker.rogers2@gmail.com.

1 Introduction

In models that address wealth accumulation and inequality, careful attention must be given to transfers between agents in the form of bequests at time of death and in the form of inter-vivos transfers. A large empirical literature exists on who receives these transfers as well as who gives these transfers, although it focuses mostly on how these transfers are given and received in terms of each possible dimension heterogeneity independently. Further, the theoretical literature on wealth transfers focuses mostly on modeling the giver of the transfers and their motives.

In this paper, we provide a contribution to both the empirical and theoretical literature on recipients of bequests. We use United States data to estimate a robust and flexible joint distribution of bequest recipients by age and net worth of recipient. This joint distribution represents a step forward in our empirical understanding of of bequest recipients. This joint distribution can also be effectively used in the theoretical literature to accurately model who receives transfers and test the persistence of wealth inequality.[1]

Wolff (2015) provides an expansive survey of the empirical and theoretical bequest-inheritance literature. The primary sources of U.S. data are the Survey of Consumer Finances (SCF), Panel Study of Income Dynamics (PSID), and Health and Retirement Study (HRS). The empirical portion of Altonji et al. (1997) focuses on the distribution of transfer recipients by income. McGarry and Schoeni (1995) address a coarse distribution of transfers by age by looking at transfers to children and transfers to parents. Schoeni (1997) and Wolff et al. (2007) separately estimates the distribution of transfers received by recipient income levels and then by recipient age using data from the PSID. All four of these studies find either that transfers received decline with income or that transfers received are a U-shaped function of income. In contrast, Cox (1987), Cox and Rank (1992), Zissimopoulos (2001), and Wolff (2015) find that transfer receipts are strongly positively correlated with income. We find this positive correlation as well. None of these studies look at the joint distribution

[1] The Python code for this multi-variate estimator of the joint distribution of bequest recipients is available at https://github.com/rickecon/OGbequests.

by income and age.

Most of the theoretical literature on bequests, inheritances, and transfers focuses on the motives of the giver.[2] Dynastic models—such as Gokhale et al. (2001) and Farhi and Werning (2010)—often assume that the wealth of one generation is divided evenly among surviving children. This does not account for proportion of bequests to grandchildren or non-family members, nor does it account for the distribution of bequests with potential disproportionality by income or ability of children or general recipients. Piketty and Saez (2013) do calibrate the distribution of bequest recipients and givers in their simple model....

The current conventional wisdom among Americans is that inheritances propagate wealth inequality in the United States Wolff (2015). In order to accurately model and predict how inheritances affect wealth inequality, we must better understand which individuals are receiving these wealth transfers. Recent literature suggests that the relationship between age and inheritance reception rates is of the form of an inverted U-shape. More specifically, the rates of inheritance reception among age groups were, 16.2% for the youngest age group, 26.9% for ages 55-64, and 8.3% for ages 75+. Moreover, there is strong positive correlation between income level, and the percentage of households receiving wealth transfers Wolff (2015).

2 Survey of Consumer Finances Data

The data we used to create our joint distribution was pulled from the Survey of Consumer Finances. The Survey of Consumer Finances (SCF) is a triennial survey conducted by The Federal Reserve that began in 1983. Each data set contains around 5,400 variables with over 30,000 observations. Respondents range from age 18 to age 96.

[2] Wolff (2015) divides these theoretical motives for giving into (i) altruism or others' wellbeing in one's utility function, (ii) exchange or giving bequests in exchange for some service, and (iii) insurance or giving bequests to others in bad times for the promise of returning the favor if bad times should arise in the future. See (i) Barro (1974), Becker (1974), Becker and Tomes (1979), Tomes (1981), (ii) Bernheim et al. (1985), Cox (1987), and (iii) Cox (1990), Cox and Jappelli (1990), and Kochar (1997).

Another possible data source that is rich in financial information is The Panel Survey of Income Dynamics (PSID). The PSID is a cross-sectional survey conducted by The University of Michigan that began in 1968 and follows 18,000 individuals. Data can be found on employment, income, wealth, expenditures, health, marriage, childbearing, child development, philanthropy, education, and numerous other topics. Insititute for Social Research (2015)

The SCF provides a comprehensive list of all variables included in the data set, called the codebook, where we found all of the useful variables for our distribution. The following is a list of codebook variables and an identifying symbol for theoretical purposes:

X8022 - ($s \in \{18, ..., 96\}$), age of respondent.

net worth - ($j \in (-139,000 \ , \ 890,000,000)$), an SCF summary variable that contains the self-perceived net worth of the respondent at the time of the survey.

X5805 - (t_1), the year that they received the first instance of inheritances (corresponding to X5804).

X5810 - (t_2), the year that they received the second instance of inheritances (corresponding to X5809)

X5815 - (t_3), the year that they received the third instance of inheritances (corresponding to X5814)

X5804 - (b_1), the approximate value of the first instance of inheritance when the respondent received them. We consider the indeces $i, \ s, \ j$ to denote the individual, their age, and their net worth respectively.

X5809 - (b_2) the second reported value of inheritance reception.

X5814 - (b_3) the third reported value of inheritance reception.

wgt - (ω_i), an SCF summary variable that contains the weights provided by the Federal Reserve to adjust for selection bias in the survey. Each weight corresponds to individual i

We define t_0 as the year that the survey being used was conducted.

In our particular distribution we assign our variable net worth the following values (this can be done according to different income percentiles),

$$j = \begin{cases} 0 & j < 15,000 \\ 1 & 15,000 \leq j < 25,000 \\ 2 & 25,000 \leq j < 50,000 \\ 3 & 50,000 \leq j < 75,000 \\ 4 & 75,000 \leq j < 100,000 \\ 5 & 100,000 \leq j < 250,000 \\ 6 & j \geq 250,000 \end{cases}$$

In order to increase our sample size, we considered data from SCF survey years 1989-2013. Since the SCF is a triennial survey, we first wanted to prevent any overlap of bequest reception between survey years. In order to do this, we used the variables t_1, t_2, and t_3 as described above. These variables give the time that each instance of bequest was received, and allow us control for the bequests received outside of the current, and previous two years of the survey year. Formally, we include the instances of bequests $(b_1,\ b_2,\ b_3)$ in our distribution if their corresponding $(t_1,\ t_2,\ t_3)$ satisfy the following condition,

$$(t_0, t_0, t_0) - (t_1, t_2, t_3) < 3$$

where t_0 is the year of the survey being used.

3 Estimating the Joint Distribution of Bequests

After each instance of reported bequests satisfies the condition,

$$(t_0, t_0, t_0) - (t_1, t_2, t_3) < 3$$

we define new bequest variables, using variables s and j, that are conditional upon the individual, their age, and net worth. We define them as follows,

$b_{1,i,s,j,}$, the first reported bequest reception (b_1) for individual i, with age s, and with net worth j.

$b_{2,i,s,j,}$, the second instance of bequest reception (b_2) for that same individual.

$b_{3,i,s,j,}$, the third instance of bequests (b_3) for that same individual.

we place these bequests into a matrix, representing our joint distribution, conditional on different age and net worth categories . Let $S,\ J$, be the total number of age groups and net worth groups respectively and let B be an $S \times J$, matrix representing our joint distribution of bequests conditional upon age, s, and net-worth category, j. Any arbitrary $b_{s,j}$ entry of B is defined by the following condition,

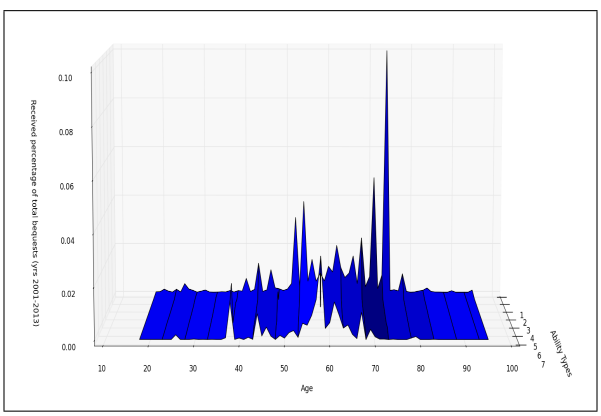
$$b_{s-18,j} = \int_i b_{1,i,s,j} + b_{2,i,s,j} + b_{3,i,s,j}$$

We then create a matrix of proportions that describes the portion of total bequests that each age-income group receive by doing the following,

$$b_{s-18,j} = \frac{\int_i b_{1,i,s,j} + b_{2,i,s,j} + b_{3,i,s,j}}{\int_i \int_s \int_j b_{1,i,s,j} + b_{2,i,s,j} + b_{3,i,s,j}}$$

Note that $\int_s \int_j b_{s-18,j} = 1$. The following graph is the result of this matrix of proportions,

Figure 1: Joint Distribution of Inheritances Before Weights Applied



Notice the spike in bequest reception at age 73 and net worth group 0. This is due to the selection bias in the survey. In order to reduce this effect, the SCF provides a Summary Variable called "wgt" (ω_i). This variable is a vector of weights corresponding to each individual provided to "compensate for unequal probabilities of selection in the original design and for unit nonresponse (failure to obtain an interview)" of Governers of the Federal Reserve System (2015). We applied these weights to our inheritances variables by multiplying the following,

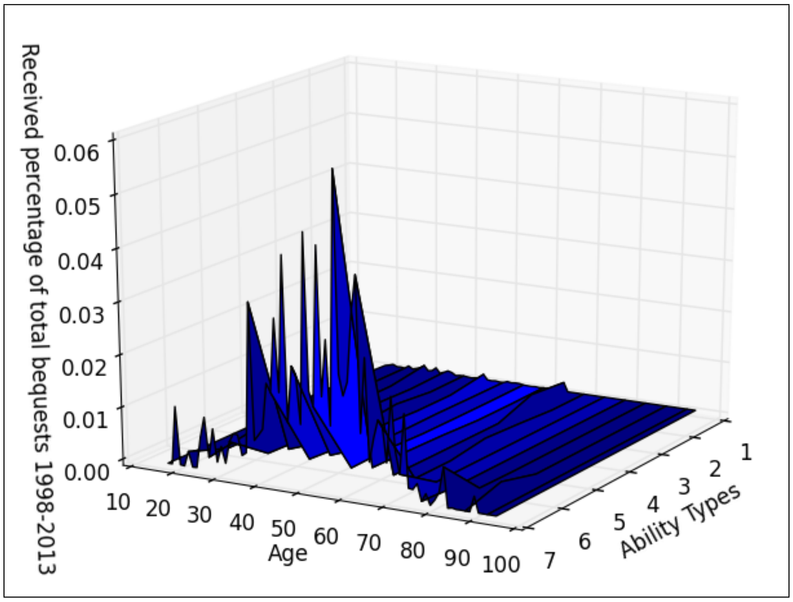
$$b_{s-18,j} = \frac{\int_i \omega_i * (b_{1,i,s,j} + b_{2,i,s,j} + b_{3,i,s,j})}{\int_i \int_s \int_j \omega_i * (b_{1,i,s,j} + b_{2,i,s,j} + b_{3,i,s,j})}$$

Lastly, we adjusted the nominal values of (b_1, b_2, b_3) to account for inflation between the used survey years of 1998-2013. We used the St. Louis Federal Reserve's CPI for all urban consumers: all items of St. Louis (2015) . We then set 2013 as the base year and calculated the percentage inflation factors for all the years down to 1998. The inflation factors, which are be multiplied by the bequest amounts that correspond to each year, are given as follows,

1998	1999	2000				
1.300280732	1.284934882	1.260857994				
2001	2002	2003	2004	2005	2006	2007
1.240039148	1.227912707	1.210171616	1.189103802	1.161807505	1.134803101	1.109966432
2008	2009	2010	2011	2012	2013	
1.076012397	1.078969961	1.063898833	1.034477726	1.014431538	1	

The graph of the resulting matrix after applying weights and inflation factors, which shows the joint distribution of bequests received over the years 1998-2013, is as follows,

Figure 2: Distribution of Inheritances 1998-2013



Notice the peak of inheritance reception at the ages of 55-65 years and net worth group 7 (top earners). This suggests that most inheritances given fall on middle-aged, wealthy individuals.

3.1 Multivariate Kernel Density Estimator (MVKDE)

After creating a matrix that describes the joint distribution of inheritances over age and net worth categories, we fit a functional form to this matrix in order to facilitate flexiblility among these categories. This functional form is useful because it would allow our joint distribution to accomodate different model inputs. Now, a model that is characterized by any number of heterogenous net worth and age cohorts can use our distribution simply by defining the desired amount of age and net worth groups. Since our matrix of proportions sums to one, we use a Multivariate Kernel Density Estimator to fit a probablity distribution function to our data in Figure 2.

Kernel Density Estimation is a nonparametric way to approximate a density function when the parametric form of the sampling density is unknown. We used a Multivariate Kernel Density Estimator to estimate the sample density as there were no clear parametric density forms that matched our data set. We treat both the univariate and the multivariate Kernel Density Estimators.

The Univariate Kernel Density Estimator is defined as follows:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i)$$

where $K_h(x) = K(x/h)/x$ is the kernel function of choice, h is the bandwidth, and n is the number of sample points Scott (2015).

Much like the histogram, a Kernel Density Estimator, creates functional values that correspond to a chosen kernel function at each sample point. Moreover, upon summing up each kernel function corresponding to its respective sample point, this sum of kernels is evenly scaled by dividing the functional values by the total number of sample points and the bandwidth selection.

In our case we use a joint distribution, with variables s, j. Thus, we use the multivariate case of kernel estimation, which is given by the following,

$$\hat{f}_{\mathbf{H}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i)$$

We choose $K_h(x)$ to be a Gaussian Kernel, described by the following Kernel Density Estimator for $\mathbf{x} \in \mathbb{R}^d$

$$\hat{f}(\mathbf{x}) = \frac{1}{n2\pi^{d/2}|\Sigma|^{1/2}}\sum_{i=1}^{n} \exp\big[-\frac{1}{2}(x - x_i)^T\Sigma^{-1}(x - x_i)\big]$$

David W. Scott (2015)

A noteworthy characteristic of the MVKDE function that we implement is the use of a bandwidth parameter. This parameter allows for the graph to be more or less smooth. Smoothness corresponds to higher values of the bandwidth in our MVKDE function, and represents a distribution that is less determined by the data. The lower the bandwidth selection, the rougher the graph is and the more jagged the distribution; however, lower bandwidths more accurately portray the behavior of the sample set.

In order to implement this kernel, we used a kernel function from an imported scipy module. Since kernel estimators work much like a histogram, we changed the nature of our data by converting the matrix of proportions into two individual vectors filled with frequencies that represent the number of times that different age and net worth groups occurred. To do this, we first sum our matrix of proportions over our age axis, for our length J net worth proportion vector (n_j), and then over our net worth axis for our length $S - 18$ age proportion vector (a_s), which is given by the following,

$$n_j = \int_s b_{s,j} \quad \text{and} \quad a_s = \int_j b_{s,j}$$

where $\int_j n_j = 1$ and $\int_s a_s = 1$. Note that these proportion vectors, can be thought of as vectors of probabilities. We then take 70,000 random draws from each probability vector, which results in new vectors filled with the frequencies that each age (in ascending order from 18-95), and net worth group (in ascending order from 1-7) were drawn from each probability vector. For example, if $a_0 = 1,876$, then age 18 was drawn 1,876 times. Note that

$$\int_j n_j = 70,000 \quad \text{and} \quad \int_s a_s = 70,000$$

We now create two new vectors, A_s and N_j, defined as follows,

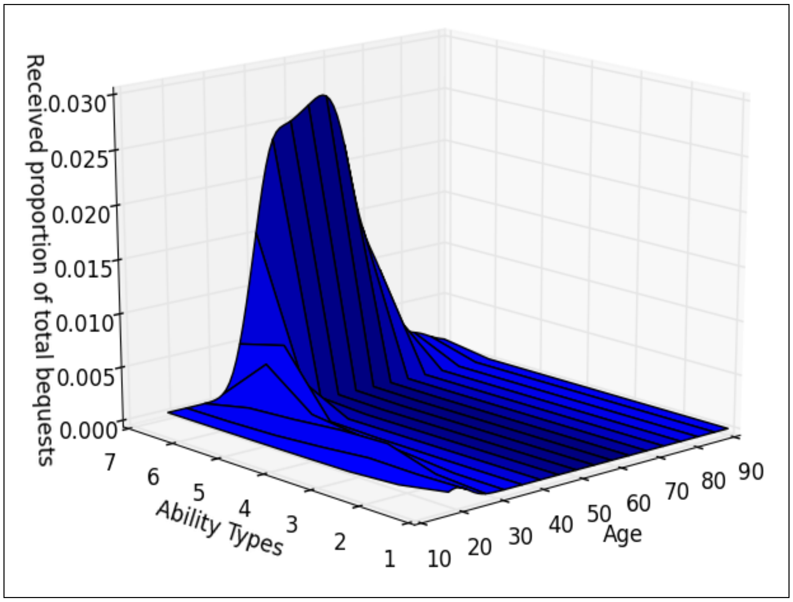
$$A_s = (s + 18, s + 18, s + 18, ..., s + 18) \quad \text{and} \quad N_j = (j + 1, j + 1, j + 1, ..., j + 1)$$

where A_s is a sequence of length a_s, s is the index of a_s, and the vector A is of length 70,000, and N_j is a sequence of length n_j, j is the index of n_j, and the vector N is of length 70,000. The vectors A and N now provide the necessary information for scipy's kernel function, and are vectors full of the ocurrences of each age or net worth group.

3.2 Flexible Joint Distribution of Bequests

Our flexible joint distribution of bequests, created with Multivariate Kernel Density Estimator, is illustrated in the following graph,

Figure 3: Flexible Joint Distribution of Inheritances 1998-2013



Note that the peak of bequest reception still lies in the age group 55-65 years and net worth group 7 as in Figure 2. Our estimated distribution is ultimately defined by a matrix of proportions whose dimensions are determined by the user. In this particular case, we specified 78 age groups (ages 18-95) and 7 net worth type groups. Any desired amount of age groups and net worth groups can be given, and the resulting matrix will match these amounts. This flexibility allows our distribution to accomodate different specifications in macroeconomic models.

4 Conclusion

Our estimation of the joint distribution of bequests allows for a data-driven allocation of inheritances among both age and net worth groups. This allows for more heterogeneity among agents within macroeconomic models, and can ultimately lead to more accurate results. By modeling our distribution using a MVKDE we are able to have flexible age and net worth categories, which can be used in many macroeconomic models.

In other models, the assumption that inheritances are distributed equally among different net worth groups, or that they are distributed only to individuals who belong to the same net worth group as those giving the inheritances is most assuredly not true. By using a more accurate distribution of inheritances we can better predict the effects that inheritances have on the distribution of wealth.

Even though we were able to find data in The Survey of Consumer Finances regarding individuals who receive bequests by both age and net worth, we were unable to find data that links these individuals with those who are giving bequests. It would be useful to know which individuals, by age and net worth categories, are giving bequests so that we can also accurately model this component of bequest distribution.

References

Altonji, Joseph G., Fumio Hayashi, and Laurence J. Kotlikoff, "Parental Altruism and Inter-Vivos: Theory and Evidence," *Journal of Political Economy*, December 1997, *105* (6), 1121–1166.

Barro, Robert J., "Are Government Bonds Net Wealth?," *Journal of Political Economy*, Nov./Dec. 1974, *82* (6), 1095–1117.

Becker, Gary S., "A Theory of Social Interactions," *Journal of Political Economy*, Nov./Dec. 1974, *82* (6), 1063–1093.

_ **and Nigel Tomes**, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, December 1979, *87* (6), 1153–1189.

Bernheim, B. Douglas, Andrei Shleifer, and Lawrence H. Summers, "The Strategic Bequest Motive," *Journal of Political Economy*, December 1985, *93* (6), 1045–1076.

Cox, Donald, "Motives for Private Income Transfers," *Journal of Political Economy*, June 1987, *95* (3), 508–546.

_ , "Intergenerational Transfers and Liquidity Constraints," *Quarterly Journal of Economics*, February 1990, *105* (1), 187–218.

_ **and Mark R. Rank**, "Inter-Vivos Transfers and Intergenerational Exchange," *Review of Economics and Statistics*, May 1992, *74* (2), 305–314.

_ **and Tullio Jappelli**, "Credit Rationing and Private Transfers: Evidence from Survey Data," *Review of Economics and Statistics*, August 1990, *72* (3), 445–454.

Farhi, Emmanuel and Iván Werning, "Progressive Estate Taxation," *Quarterly Journal of Economics*, May 2010, *125* (2), 635–673.

Gokhale, Jagadeesh, Laurence J. Kotlikoff, James Sefton, and Martin Weale, "Simulating the Transmission of Wealth Inequality via Bequests," *Journal of Public Economics*, January 2001, *79* (1), 93–128.

Kochar, Anjini, "An Empirical Investigation of Rationing Constraints in Rural Credit Markets in India," *Journal of Development Economics*, August 1997, *53* (2), 339–371.

McGarry, Kathleen and Robert F. Schoeni, "Transfer Behavior in the Health and Retirement Study: Measurement and the Redistribution of Resources within the Family," *Journal of Human Resources*, 1995, *30* (4 Supplement), S184–S226.

of Governers of the Federal Reserve System, Board, "Files: Survey of Consumer Finances Codebook," 2015.

of Michigin Insititute for Social Research, University, "PSID The Panel Study of Income Dynamics - PSID - is the longest running longitudinal household survey in the world.," 2015.

of St. Louis, Federal Reserve Bank, "FRED Economic Data Index for All Urban Consumers: All Items," 2015.

Piketty, Thomas and Emmanuel Saez, "A Theory of Optimal Inheritance Taxation," *Econometrica*, September 2013, *81* (5), 1851–1886.

Schoeni, Robert F., "Private Interhousehold Transfers of Money and Time: New Empirical Evidence," *Review of Income and Wealth*, December 1997, *43* (4), 423–448.

Scott, David W., *Kernel Density Estimators, in Multivariate Density Estimation: Theory, Practice, and Visualization*, John Wiley and Sons, Inc., 2015.

Scott, Stephan R. Sain David W., *9-Multidimensional Density Estimation*, John Wiley and Sons, Inc., 2015.

Tomes, Nigel, "The Family, Inheritance, and the Intergenerational Transmission of Inequality," *Journal of Political Economy*, October 1981, *89* (5), 928–958.

Wolff, Edward N., *Inheriting Wealth in America: Future Boom or Bust?*, Oxford University Press, 2015.

Wolff, Francois-Charles, Seymour Spilerman, and Claudine Attias-Donfut, "Transfers from Migrants to Their Children: Evidence that Altruism and Cultural Factors Matter," *Review of Income and Wealth*, December 2007, *53* (4), 619–644.

Zissimopoulos, Julie, "Resource Transfers to the Elderly: Do Adult Children Substitute Financial Transfers for Time Transfers?," RAND Labor and Population Program Working Paper Series 01-05, RAND 2001.