Journal of Behavioral and Experimental Economics 72 (2018) 9–16

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

journal homepage: www.elsevier.com/locate/jbee

On using interval response data in experimental economics☆

James McDonald, Olga Stoddard*, Daniel Walton

Brigham Young University. Department of Economics, 130 FOB, Provo UT 84602, USA

ARTICLE INFO

Keywords:
Discount rates
Interval dependent variable
Distributional misspecifications

ABSTRACT

Many empirical applications in the experimental economics literature involve interval response data. Various methods have been considered to treat this type of data. One approach assumes that the data correspond to the interval midpoint and then utilizes ordinary least squares to estimate the model. Another approach is to use maximum likelihood estimation, assuming that the underlying variable of interest is normally distributed. In the case of distributional misspecification, these estimation approaches can yield inconsistent estimators. In this paper, we explore a method that can help reduce the misspecification problem by assuming a distribution that can model a wide variety of distributional characteristics, including possible heteroskedasticity. The method is applied to the problem of estimating the impact of various explanatory factors associated with individual discount rates in a field experiment. Our analysis suggests that the underlying distribution of discount rates exhibits skewness, but not heteroskedasticity, In this example, the findings based on a normal distribution are generally robust across distributions.

1. Introduction

Many empirical applications in the experimental economics literature involve interval response data. Examples include commonly used measures of risk aversion (see Harrison and Rutstrom, 2008; Charness et al., 2013, for an overview), second-price Vickrey auctions with interval bidding possibilities (Banerjee and Shogren, 2014), estimation of willingness-to-pay (WTP; Dominitz and Manski, 1997; Hanley et al., 2009, 2013) , and individual discount rates (Coller and Williams, 1999; Harrison et al., 2002). The typical critique against tasks that elicit point estimates in these contexts is "the payoff dominance" problem first raised by Harrison (1992). The Becker–DeGroot–Marschak (BDM) procedure, in particular, is known to have weaker incentives around the optimum. In addition, data that rely on single-response methods, such as the BDM, to elicit risk preferences or WTP are significantly noisier (Harrison, 1986).

Various methods have been considered to treat this type of data. One approach assumes that the data correspond to the interval midpoint and then utilizes ordinary least squares to estimate the model. Another approach is to use maximum likelihood estimation, assuming that the distribution of the underlying variable of interest is of a particular form, such as the normal. While these methods are widely used in the literature, they can yield inconsistent estimators and thus misleading results in cases of distributional misspecification or in the presence of heteroskedasticity.

In this paper, we consider the implications of using an estimator, which is based on a flexible distribution that can accommodate a wide range of skewness and kurtosis, hence having the potential to reduce the impact of distributional misspecification. In particular, we use maximum likelihood estimation of an interval response regression model that corresponds to the skewed generalized *t* distribution (SGT) and the generalized beta of the second kind (GB2). The SGT can model a wide range of distributional characteristics for real-valued skewed and leptokurtic data and includes many important distributions, such as the normal, Laplace, generalized error distribution, and skewed variations of these distributions as special and limiting cases. The GB2 is a flexible distribution for positive valued outcomes. These two flexible distribution functions serve as alternatives to the normal distribution often employed in interval regressions.

We apply this method to the problem of estimating the effects of various possible explanatory factors on individual discount rates in a field experiment described in Harrison et al. (2002), hereafter referred to as HLW. In this experiment, the authors elicit individual discount rates from subjects and test whether these rates vary (1) across households and (2) over time. HLW find that discount rates vary significantly with respect to several sociodemographic variables but not over a one- to three-year time horizon. This finding provides an important contribution to our understanding of the nature of individual discount rates, given their essential role in intertemporal welfare analyses.

☆ The authors would like to express their appreciation to Bryan Chia for the outstanding research assistance.
* Corresponding author.
E-mail address: olga.stoddard@byu.edu (O. Stoddard).

https://doi.org/10.1016/j.socec.2017.10.003
Received 18 October 2016; Received in revised form 20 October 2017; Accepted 22 October 2017

In this paper, we consider the implications of allowing for more general distributions in estimating the model. We observe that the underlying distribution of reported discount rates exhibits skewness, heteroskedasticity, or both. This is inconsistent with the assumption of normality and can impact parameter estimates. When applying more flexible distributions, which allow for a wide range of skewness and kurtosis values, such as SGT and GB2, we find that the nominal discount rates are significantly impacted by some sociodemographic factors. We compare and contrast our results with those obtained under the assumption of normality and find that the magnitudes and statistical significance of the coefficients are sensitive to the specification used, but they are generally consistent with the findings of HLW.

In particular, our results show that the GB2 family generally dominates the SGT as it provides a better fit with fewer parameters. Within the GB2 family, the 2-parameter and 3-parameter gamma (GA) and generalized gamma (GG) distributions are arguably the best choice, considering fit, parsimony, and easy interpretation. An added advantage of the GB2 family over SGT is that an assumption of "heteroscedasticity" (making σ a function of covariates) is unnecessary, considerably simplifying the interpretation of parameters. For both the GA and GG, we find support for the HLW conclusion that rates appear to be somewhat greater at a 6-month delay than for the longer delays, but constant across the longer delays. We also find that in addition to the discount rate that predictors found to be significant in HLW, our estimation of the GB2 model uncovers additional statistically significant covariates.

This paper contributes to a growing literature in experimental economics, which emphasizes various approaches to data analysis that are widely used by other research communities (Ashley et al., 2010; Frechette, 2012). While we discuss some well-known methods and their application to interval response data, we also highlight a new methodological framework and its advantages. We emphasize the important implications that the underlying theory has for econometric models and show how to check robustness of results to model specifications.

We focus this paper on the impact of accommodating diverse distributional characteristics of individual responses of monetary discount rates, rather than addressing the more complicated problem of joint estimation of the distribution and an underlying utility function as explored in Anderson et al. (2008). The methodological framework is outlined in Section 2. Section 3 provides an application of the methods to the problem of estimating individual discount rates, and Section 4 concludes.

2. Methodology

2.1. The model and likelihood function

The proposed model can be summarized as follows:

$$y_i^* = X_i\beta + \varepsilon_i \quad i = 1, 2, \ldots, n \tag{1}$$

where only the thresholds containing the latent variable y_i^* are observed, X_i is a *1xK* vector of explanatory variables with a corresponding *Kx1* coefficient vector β, and the ε_i are assumed to be independently and identically distributed random disturbances. The observed upper and lower thresholds of the latent variable y_i^* are denoted by U_i and L_i, respectively.

Stewart (1983) notes that inconsistent parameter estimates may result from using regular ordinary least squares (OLS), with the dependent variable being assigned to the value of the interval midpoint, and the open-ended groups being assigned values on an ad hoc basis. Stewart outlines different approaches to yield MLE (maximum likelihood estimation) under the assumption of normality and applies these methods to the problem of estimating an earnings equation. Stata's *intreg* command facilitates MLE of interval response data in the case of normally distributed errors and allows for the presence of heteroskedasticity.

We also apply a MLE approach to this estimation problem but allow for possibly non-normal distributions, which can accommodate skewness and kurtosis. We begin by noting that the conditional probability that y_i^* is in the interval (L_i, U_i) is given by

$$\Pr(L_i \le y_i^* \le U_i) = F(U_i; \beta, \theta|X_i) - F(L_i; \beta, \theta|X_i), \tag{2}$$

where F(.) denotes the cumulative conditional distribution of y_i^* and θ denotes a vector of distributional parameters. The corresponding log-likelihood function for interval regression models is given by

$$\ell(\beta, \theta) = \sum_i \ell n[F(U_i; \beta, \theta|X_i) - F(L_i; \beta, \theta|X_i)] \tag{3}$$

Interval regression programs allow not only for interval data but for censored data as well. For example, the Stata interval regression program, *intreg*, accommodates right censored $((-\infty, U_i])$ and left censored $([L_i, \infty))$ data by replacing the corresponding terms in (3) with $F(U_i; \beta, \theta|X_i)$ and $(1 - F(L_i;\beta, \theta|X_i))$, respectively.

Maximum likelihood estimation (MLE) will be used throughout this paper where Eq. (3) is maximized over the unknown parameters (β and θ).

2.2. Distributional assumptions

As noted in the introduction, the properties of the parameter estimates can be sensitive to the distributional assumptions. The most common implementation of the MLE approach to this type of data in the literature is based on the assumption of normally distributed errors. As mentioned earlier, Stata's interval regression command (*intreg*) assumes normally distributed errors and is a Tobit-like estimator for grouped data. However, these estimators can be inconsistent if the errors are not normally distributed or are associated with heteroskedasticity. Adaptive or semiparametric estimation of econometric models avoid specifying a particular probability density function but may be difficult to implement. Partially adaptive estimation relaxes the normality assumption by adopting a more flexible probability density function to approximate the actual error distribution. Caudill (2012) uses a mixture of normal distributions. Cook and McDonald (2013) use an inverse hyperbolic sine distribution to estimate censored regression models, finding that this specification improves estimator performance for the cases considered. We will use the skewed generalized *t* (SGT) and the generalized beta of the second kind (GB2), each of which allows a wide range of skewness and kurtosis. The SGT can model real-valued responses and includes the normal as a special case. The GB2 is a flexible model for applications in which the responses are positive, such as in the example considered in Section 3.

2.3. The skewed generalized t distribution

The SGT was introduced by Theodossiou (1998) and extends the generalized *t* (GT) (McDonald and Newey, 1988) and the skewed *t* (ST) (Hansen 1994) and allows for a wide range of skewness and kurtosis; for example, see Kerman and McDonald (2013). Other special cases of the SGT include the skewed generalized error distribution (SGED), skewed Laplace (SLaplace), generalized error distribution (GED), skewed normal (SNormal), *t*, skewed Cauchy (SCauchy), Laplace, Uniform, Normal, and Cauchy. The five-parameter SGT can be defined by the following density function:

$$SGT(y; \mu, \lambda, \sigma, p, q) = \frac{p}{2\sigma q^{1/p} B\left(\frac{1}{p}, q\right)\left(1 + \frac{|y-\mu|^p}{q\sigma^p(1+\lambda sign(y-\mu))^p}\right)^{q+1/p}} \tag{4}$$

where $-\infty < y < \infty$ and $B(.,.)$denotes the beta function.

The SGED is a limiting case of the SGT defined by

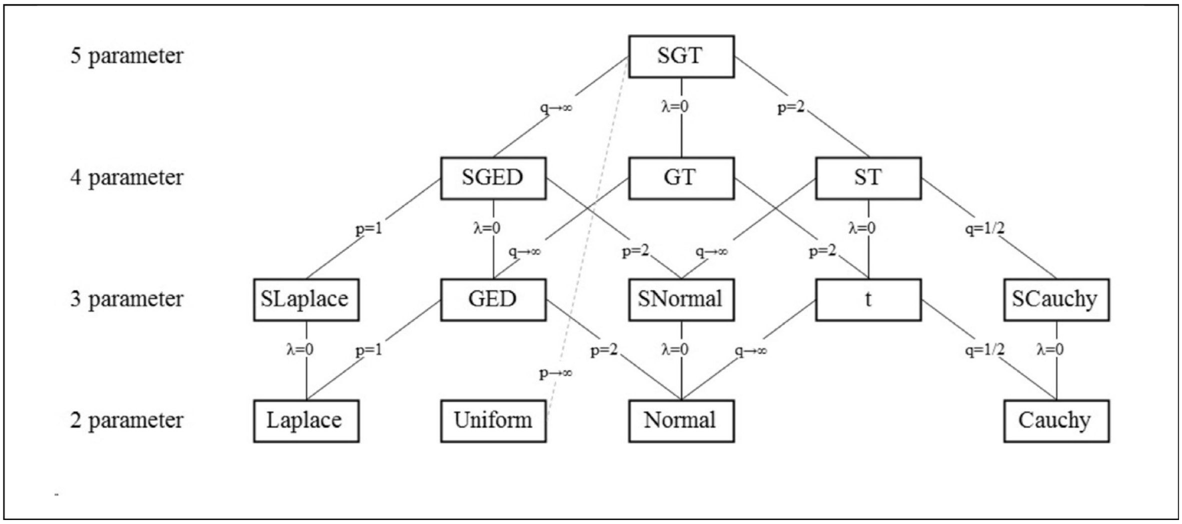

Fig. 1. SGT distribution tree (adapted from Hansen et al., 2010).

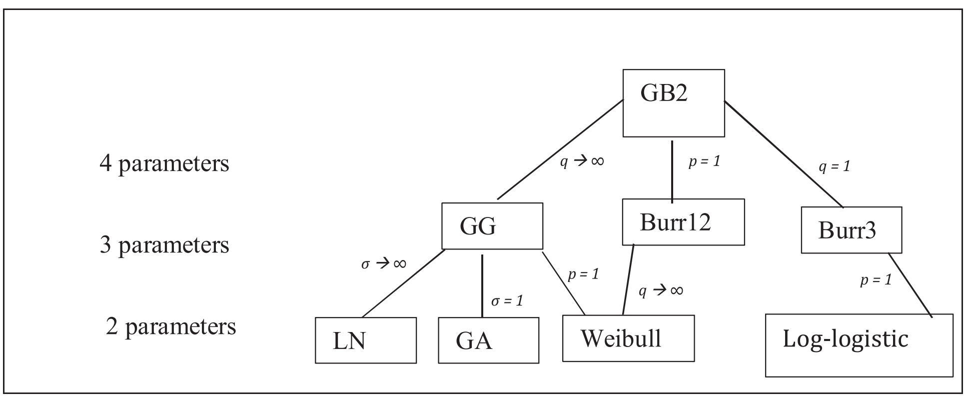

Fig. 2. GB2 distribution tree (McDonald, 1984).

$$SGED(y;\mu,\lambda,\sigma,p) = \lim_{q\to\infty} SGT(y;\mu,\lambda,\sigma,p,q) = \frac{pe^{-\left(\frac{|y-\mu|}{\sigma(1+\lambda sign(y-\mu))}\right)^{p}}}{2\sigma\Gamma\left(\frac{1}{p}\right)} \tag{5}$$

where $\mu = X\beta$ is a location parameter, σ is a positive scale or dispersion parameter, and $\Gamma(.)$ denotes the gamma function.[1] The parameter λ, $-1 < \lambda < 1$, measures skewness, with the probability of Y being greater than μ given by $\left(\frac{1+\lambda}{2}\right)$; hence, $\lambda = 0$ corresponds to a symmetric pdf (probability density function). The parameters p and q are positive and impact the shape of the pdf, and the product pq is referred to as the degrees of freedom parameter. The interrelationships between these pdfs are shown in Fig. 1.

The SGT family of distributions provides a generalization of many methods that have been used for analyzing models with interval response data. The expected value of the response corresponding to the SGT and its various special cases can be evaluated using

$$E_{SGT}(Y|X) = X\beta + 2\lambda\sigma\left\{\frac{q^{1/p}B(2/p,\, q-1/p)}{B(1/p,\, q)}\right\}$$

$$E_{SGED}(Y|X) = X\beta + 2\lambda\sigma\left\{\frac{q^{1/p}\Gamma(2/p)}{\Gamma(1/p)}\right\} \tag{6. a-b}$$

[1] Allowing σ to be a function of the explanatory variables can model heteroskedasticity.

2.4. The generalized beta of the second kind

The generalized beta of the second kind (GB2) is a well-known four-parameter distribution for positive-valued random variables that has been successfully used in applications such as the distribution of income and stock returns. It includes the generalized gamma (GG), gamma (GA), Weibull (W), lognormal (LN), Burr3, and Burr12, among others, as special or limiting cases. Fig. 2 illustrates the relationships between the members of the GB2 family.

The GB2 pdf is defined by

$$GB2(y;\delta,\sigma,p,q) = \frac{e^{p(\ln(y)-\delta)/\sigma}}{\sigma y B(p,q)(1+e^{(\ln(y)-\delta)/\sigma})^{p+q}} \tag{7}$$

where σ, p, and q denote positive distributional parameters (McDonald, 1984). Important special cases of the GB2 are the Burr3 and Burr12, corresponding to $q = 1$ and $p = 1$, respectively, and the generalized gamma as the following limiting case:

$$GG(y;\delta,\sigma,p) = \lim_{q\to\infty} GB2(y;q^{\sigma}\delta,\sigma,p,q) = \frac{e^{p(\ln(y)-\delta)/\sigma}e^{-e(\ln(y)-\delta)/\sigma}}{\sigma y\Gamma(p)}, \tag{8}$$

The GG includes the gamma (GA) and Weibull (W) as the following special cases:

$$GA(y;\delta,p) = G(y;\delta,\sigma=1,p) = \frac{e^{p(ln(y)-\delta)}e^{-e(ln(y)-\delta)}}{y\Gamma(p)} \tag{9}$$

and

$$W(y;\delta,\sigma) = GG(y;\delta,\sigma,p=1) = \frac{e^{(\ln(y)-\delta)/\sigma}e^{-e(\ln(y)-\delta)/\sigma}}{y\sigma} \quad (10)$$

Finally, the lognormal

$$LN(y;\delta,\sigma) = \frac{e^{-(\ln(y)-\delta)^2/2\sigma^2}}{y\sqrt{2\pi}\sigma} \quad (11)$$

is a well-known limiting case of the GG and GB2.

The regression specification in (1) corresponding to the GB2 family can be obtained by letting the parameter δ be a function of the explanatory variables, for example, $\delta = X_i\beta$. The econometrics package Stata estimates the exponential, gamma, Weibull, log-logistic, and generalized gamma regression specifications for individual observations but not for interval data or data with heteroskedasticity. As with the SGT, possible heteroskedasticity can be modeled by allowing the scale parameter, σ, to be a function of the explanatory variables, $\sigma(X)$.[2]

The expected value of the dependent latent variable corresponding to the GB2 and its special cases can be evaluated using the following results:

$$E(Y_{GB2}|X) = e^{X\beta}\frac{\Gamma(p+\sigma)\Gamma(q-\sigma)}{\Gamma(p)\Gamma(q)},\ \sigma<q \quad (12.a)$$

$$E(Y_{GG}|X) = e^{X\beta}\frac{\Gamma(p+\sigma)}{\Gamma(p)}, \quad (12.b)$$

$$E(Y_{LN}|X) = e^{X\beta+\sigma^2/2} \quad (12.c)$$

From these results, we see that the β_i in the GB2 family can be interpreted as estimating the percentage change in the y_i^* corresponding to a unit change in X_i. Thus, one might expect the SGT and GB2 regression coefficients to be roughly related to each other as follows:

$$(\beta_i)_{SGT} \approx E(y_i^*)(\beta_i)_{GB2} \quad (13)$$

3. Experimental design, model specification, statistical distribution, and statistical analysis

3.1. Experimental design

In this section we apply the previously described methods to the problem of estimating individual discount rates in experimental economics—a field experiment described in HLW. In this experiment a demographically representative sample of 268 Danish individuals aged between 19 and 75 years old[3] were invited to answer survey questions with real monetary rewards. Survey questions were designed by Coller and Williams (1999), who conducted similar experiments with university students in controlled laboratory settings. To elicit discount rates, individuals were asked whether they would prefer \$100 in one month or \$100 + x in one + y months, where x > \$0 and y = 6, 12, 24, or 36 months depending on the specific condition of the experiment. Each subject faced a sequence of ten questions, each with a different x. The point at which an individual switched from choosing the current income option to taking the delayed income option provided a bound on his or her discount rate. Participants were provided with the interest rates associated with the future payment option and knew that they would be paid for one randomly selected question.

Participants were randomized into one of five treatment groups that differed according to the possible time horizons for future income options: (1) 6 months, (2) 12 months, (3) 24 months, (4) 36 months, and (5) all four time horizons presented to an individual simultaneously. A total of 118 individuals participated in 15 single-horizon sessions,[4] and 150 individuals participated in the 12 multiple-horizon sessions.[5]

In addition to the main discount rate elicitation questions, researchers also asked a wide range of sociodemographic questions, such as participants' gender, age, household income, occupation, education level, retired and employment status, and marital status. Subjects were also asked about their access to financial accounts, such as checking account, credit card, or line of credit; annual interest rates, current balance on those accounts; and their perception of their own credit worthiness. These and other covariates are more formally defined in the first two columns of Table A.1 in Appendix A.[6]

The experiment was designed to test two specific hypotheses. The first hypothesis is that discount rates for a given time horizon do not differ with respect to an individual's sociodemographic characteristics. The second hypothesis is that discount rates for a given individual do not differ across time horizons. Harrison et al. (2002) found that discount rates among this sample of Danish individuals were relatively constant over the one-year to three-year time horizon but varied significantly across several sociodemographic characteristics. In particular, discount rates were significantly affected by the length of education, retirement status, unemployment, and the likelihood of obtaining a loan or being approved for a credit card.

3.2. Application of the SGT and GB2: model specification

In this section, we consider the implications of using more flexible distributions, namely the SGT and GB2 families. Parameter estimates were obtained by maximizing the log-likelihood Eq. (3) using various numerical optimization algorithms in MATLAB for selected distributions, with and without the assumption of homoskedasticity. Two regression scenarios were considered in some of these estimations: first, the full regression scenario (full) includes time-horizon variables and sociodemographic controls

$$DR_i = \beta T_i + \gamma X_i + \epsilon_i$$

where DR_i stands for the discount rate of individual i, T_i is a vector of the time-horizon indicators for the scenario that individual i received, and X_i are the various socio-demographic controls for the individual i. The second scenario (intercept specification) considered is the model with only a constant term, such that

$$DR_i = \beta_0 + \epsilon_i$$

The parameter estimates, standard errors, and log-likelihood (ℓ) values for each of these specifications are reported in the appendices (B, C, and D). Appendices B and C, respectively, include estimates for the full and intercept specifications for the SGT and GB2 families under the assumption of homoskedasticity, and Appendix D reports the corresponding results for the full SGT and GB2 heteroskedastic model specifications.[7]

3.3. Selecting an appropriate statistical distribution

The estimation results obtained from using fourteen statistical distributions are reported in the appendices. Their interrelationships are depicted in Figs. 1 and 2 with distributions higher on a given tree

[2] This specification accommodates variance of log(y) because the variance of log(y) is given by $\sigma^2[\Psi'(p)+\Psi'(q)]$ where $\Psi()$ is the digamma function.

[3] These individuals were selected based on their prior participation in the European Community Household Panel Survey (ECHP) administered by the Danish Social Research Institute (SFI) in collaboration with the Danish Ministry of Business and Industry.

[4] Within the single-horizon sessions, there were 26, 32, 31, and 29 subjects, respectively, in the 6-month, 12-month, 24-month, and 36-month treatments.

[5] This particular study is an excellent application because of the extremely low incidence of response inconsistency.

[6] All supplementary online materials referenced in the paper can be obtained at https://economics.byu.edu/Documents/Faculty/Olga%20Stoddard/APPENDIX.pdf.

[7] Formulas used to calculate clustered standard errors are reported in Appendix E.

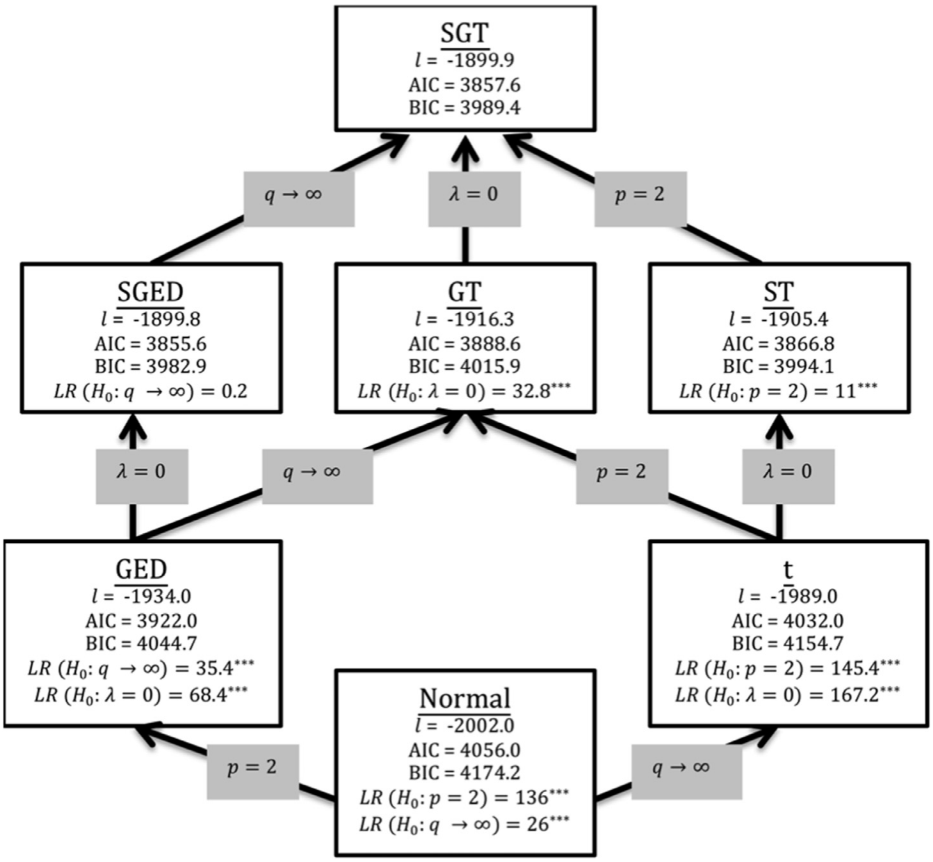

Fig. 3. Goodness of fit for the SGT distribution family.

having greater flexibility. For nested distributions on the same distribution tree, likelihood ratio tests are frequently used to select the best model. For example, the GED and Normal on the SGT distribution tree can be compared by taking twice the difference between their loglikelihood values found in Table B.1 (in Appendix B). This yields a loglikelihood value of 136 (LR = 2 (2002.0–1934.0) with the test statistic having an asymptotic Chi square distribution with one degree of freedom. Thus the GED provides a statistically significant improvement relative to the normal.

The likelihood ratio test is not valid for distributions on different trees or for nonnested models. For example, the SGT and GB2 can't be compared using a likelihood ratio test. To compare non-nested specifications, alternative criteria have been considered, such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Vuong (1989) or Clarke (2007) paired sign test.[8] These criteria reward goodness of fit as measured by ℓ (the optimized value of the log-likelihood function) and penalize a model's complexity as measured by the number of parameters (k). Thus, other things equal, a model with fewer parameters would be selected. The values of ℓ, AIC, BIC, and selected likelihood ratio values for the fitted distributions and different model specifications are reported in the last three rows of Tables B.1, B.2, C.1, C.2, D.1, and D.2 (in Appendices B, C, and D). For the full model with time horizon and sociodemographic variables, Figs. 3 and 4 summarize these goodness of fit indices for the SGT and GB2 families, respectively.

From Fig. 3 and using likelihood ratio tests, the SGT and SGED are seen to be observationally equivalent because of the large estimated value of the parameter q and yield statistically significant improved fits relative to their other special or limiting cases, including the normal. For the GB2 family, depicted in Fig. 4, the GB2 provides a statistical improvement relative to the Burr3 and Burr12, but not for the GG or gamma. Taking into account model complexity, as measured by the BIC or Vuong test and AIC, the SGED would dominate the SGT and the GG and gamma would dominate the GB2, SGT, and SGED.

Fig. 5 compares the fitted pdfs for the GG, SGED, and normal obtained from the reported interval responses.[9] Corresponding expected discount rates are reported in Table 1.[10] The normal is centered around its estimated mean of 28.3 and implies a positive probability of a negative discount rate, which is inconsistent with economic theory. The fitted pdf for the SGED is highly skewed and conforms to the expectation that discount rates will be positive. The SGED's large expected value reported in Table 1 is due to its thick right tail. In a sense, because we know that discount rates will be positive, using an SGED or SGT to model discount rates could be viewed as a type of model overspecification. Based on these comparisons, we will focus the rest of our analysis on the GG.

3.4. Economic analysis

We now consider the two hypotheses considered by HLW. For convenience, the estimated results corresponding to the normal, SGED, Lognormal, and GG specifications are presented in Table 2. It is important to recall from Eq. (13) that the "regression" coefficients in the SGT and GB2 families have different interpretations,

$$\frac{\partial E(y_{SGT}|X)}{\partial X_i} = (\beta_i)_{SGT} \text{and} \frac{\partial E(y_{GB2}|X)}{\partial X_i} = (\beta_i)_{GB2} E(y_{GB2}|X).$$

[8] The AIC and BIC are defined by $AIC = 2(k - \ell)$and$BIC = k\log(n) - 2\ell$ where k is the number of estimated parameters and ℓ is the optimized value of the log-likelihood function. A common form for the Vuong test is half of the difference of the BICs of the competing models.

[9] The estimates of the distributional parameters are given in Tables B.3 and C.3 (in the corresponding appendices). The normal is included as a benchmark because of its use in statistical software for interval regression models.

[10] The expected values are evaluated using equations (6 a–b) and (12 a–c).

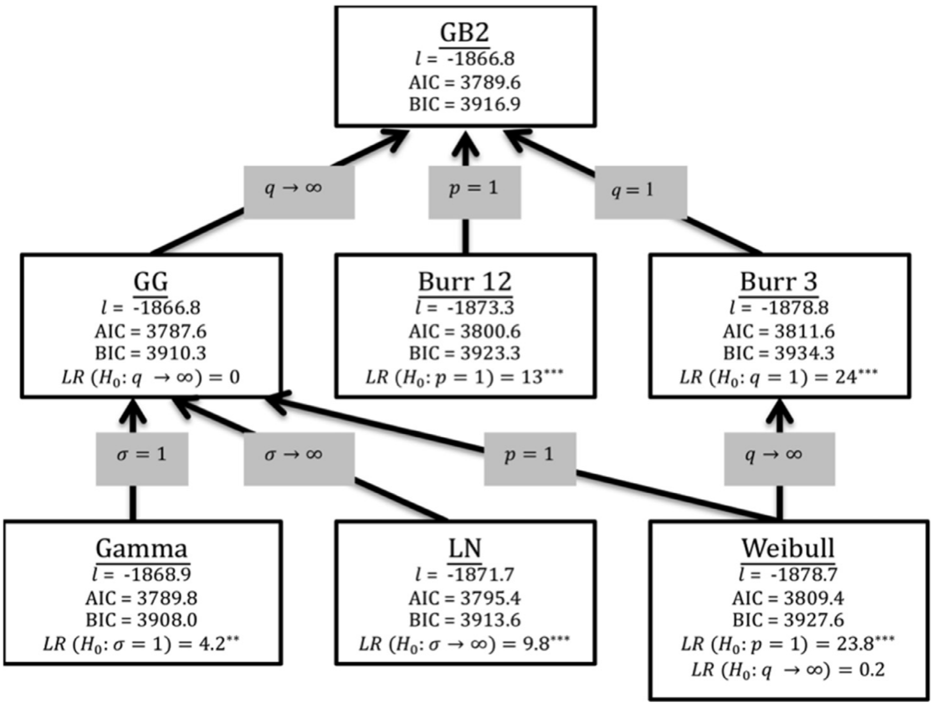

Fig. 4. Goodness of fit for the GB2 distribution family.

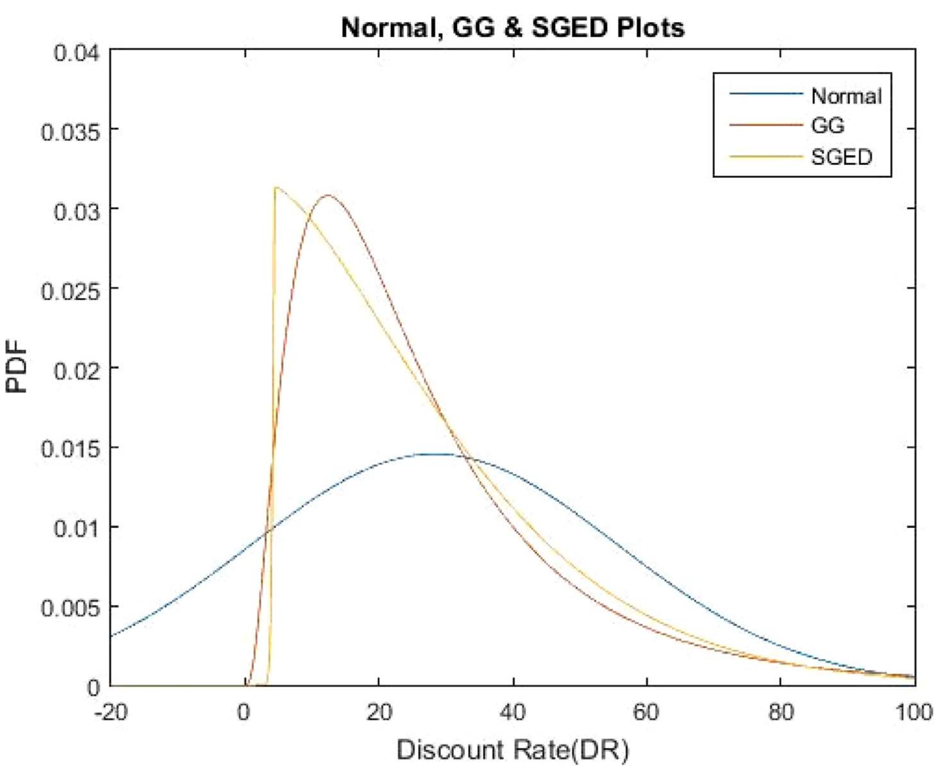

Fig. 5.

HLW found that varying the time horizon appeared to have little effect on discount rates for the 12-, 24-, and 36-month time horizons, though the 6-month time horizon was about 6 percentage points higher. Using a Wald test on the GG specification, we found that the equality of the coefficients for 12-, 24-, and 36-month time horizons could not be rejected; however, the coefficient for the 6-month horizon differed significantly.

The sociodemographic variables having a significant impact in the normal specification included the length of education and retirement, which were associated with lower and higher discounts of approximately 9% and 12%, respectively. Unemployment and having a poor chance of getting approval for a loan or credit card were significant at the 10% level and had associated estimates of −8% and 8%, respectively. These same variables are significant with a GG specification. However, income ("rich"), owner-occupied housing ("owner'), and whether the subject has a positive balance in a line of credit or credit card ("balance") were significant at 5-, 10-, and 15-percent levels, respectively. A comparison of the marginal impacts of "important" variables implied by both specifications is presented in Table 3, where the GG coefficients from Table 2 are multiplied by the corresponding

Table 1
Expected discount rates.

Distribution	Expected discount rate
Normal	28.3
SGED	51.1
GG	27.9

Table 2
Selected results for the full model (homoskedasticity), n = 696.

	Normal	SGED	Lognormal	GG
t6	34.8607***	31.6146***	3.3631***	3.3252***
	(7.8358)	(4.3030)	(0.2496)	(0.9007)
t12	28.9523***	31.2218***	3.1493***	3.1066***
	(7.7221)	(3.7192)	(0.2510)	(0.8963)
t24	27.4407***	29.8356***	3.0714***	3.0285***
	(7.9025)	(4.5096)	(0.2497)	(0.8963)
t36	27.8716***	29.4325***	3.0897***	3.0453***
	(8.2525)	(4.6663)	(0.2615)	(0.8974)
multiple	0.8359	−0.1392	−0.0173	−0.0337
	(2.4257)	(1.4425)	(0.0899)	(0.0889)
female	1.0149	0.3653	0.04637	0.0456
	(2.1972)	(1.156)	(0.0779)	(0.0750)
young	−1.0947	0.5725	0.00036	0.0105
	(3.3403)	(2.3009)	(0.1145)	(0.1132)
middle	0.1786	−2.3528**	−0.0956	−0.0520
	(2.6682)	(1.1595)	(0.0976)	(0.0941)
old	−0.45954	−0.7187	−0.0919	−0.0441
	(3.2395)	(2.4376)	(0.1166)	(0.1153)
middle1	−1.3060	−0.5305	−0.0627	−0.0593
	(2.9677)	(2.0558)	(0.0993)	(0.0932)
middle2	−3.2142	−1.8643	−0.1576	−0.1559
	(3.8167)	(2.0993)	(0.1393)	(0.1338)
rich	−5.3412	−1.4471	−0.2958**	−0.3134**
	(3.9006)	(2.8488)	(0.1377)	(0.1307)
skilled	0.74265	1.1748	0.07158	0.06769
	(2.4868)	(1.8887)	(0.0930)	(0.0887)
student	4.2049	−1.6263	0.0321	0.0481
	(4.2717)	(2.6647)	(0.14737)	(0.1365)
longedu	−9.2027***	−2.2147	−0.3975***	−0.3991***
	(2.5555)	(1.6893)	(0.0956)	(0.0926)
copen	−1.1308	−0.0859	−0.0009	0.00673
	(2.9125)	(1.9748)	(0.1083)	(0.1024)
town	3.1719	2.2304	0.1493*	0.1196
	(2.4045)	(1.4922)	(0.0896)	(0.0832)
owner	−3.7647	−2.0555	−0.1769*	−0.18608**
	(2.5200)	(1.4032)	(0.0922)	(0.0860)
retired	12.3783***	0.1513	0.3592**	0.4087***
	(4.2659)	(3.4017)	(0.1456)	(0.1289)
unemp	−7.7693**	1.1640	−0.3123**	−0.3503**
	(3.7921)	(2.8049)	(0.1438)	(0.1387)
single	−2.4016	−1.0694	−0.1229	−0.1007
	(2.6878)	(1.7728)	(0.09368)	(0.0883)
kids	0.2498	−0.1479	0.0412	0.0586
	(2.5272)	(1.0956)	(0.0912)	(0.0893)
gsize	0.0239	−0.1733	0.0018	0.0050
	(0.3125)	(0.1665)	(0.0109)	(0.0108)
balance	1.8294	2.0692	0.1451**	0.1353*
	(2.113)	(2.002)	(0.0728)	(0.0716)
Chances	7.6481**	0.7959	0.2746**	0.279**
	(3.3035)	(2.3305)	(0.1227)	(0.1169)
LogL	−2002.0	−1899.8	−1871.7	−1866.8
BIC	4067.1	3982.9		
	3806.5	3799.2		
AIC	4056	3855.6		
	3795.4	3787.6		

expected discount rate (27.9) and are in fairly close agreement with the results from the normal assumed by HLW.

Next, we test for the presence of possible heteroskedasticity in the different specifications. Likelihood ratio tests can be used to test the null hypothesis of homoskedasticity. Based on a comparison of the log-likelihood values for homoskedastic and heteroskedastic specifications reported in Table B.1 and Table D.1, respectively (see appendices), we reject the null hypothesis of homoskedasticity for the SGT family. For example, the LR value for the Normal is 98.8 (=2 (2002.0–1952.6)), with the test statistic being asymptotically distributed as a chi square with 25° of freedom. Testing for heteroskedasticity in the GG specification yields a statistically insignificant likelihood ratio value of 29.6(=2(1866.8–1852.0)). Hence, while we reject the assumption of homoskedasticity with the Normal specification, we do not reject it with GG. The same results hold when comparing the more general forms of the GB2 and SGT families with and without heteroskedasticity.

Table 3
Comparison of the marginal impact of sociodemographic variables, n = 696.

Variable	Normal	GG
rich	−5.34	−8.74
longedu	−9.20	−11.14
owner	−3.76	−0.5.19
retired	12.38	11.40
unemp	−7.77	−9.77
balance	1.83	3.78
chances	7.65	7.79

4. Summary and conclusions

Interval response data are used extensively in the experimental economics literature to estimate such important variables as discount rates, willingness to pay, and risk aversion. While various methods have been widely used in the prior literature to estimate models of this type, their properties can be sensitive to distributional assumptions and can yield inconsistent estimates.

In this paper, we present a methodology that accommodates diverse distributional characteristics. The method of estimation is based on the assumption of a flexible distribution, which allows for a wide range of data skewness and kurtosis values and has the potential to reduce the impact of distributional misspecification. In particular, we use maximum likelihood estimation of an interval response regression model that corresponds to the skewed generalized *t* distribution (SGT) and the generalized beta of the second kind (GB2). These methods are described and applied to the problem of estimating individual discount rates in a field experiment considered by HLW.

The results of this paper generally confirm those obtained using a normal specification (or HLW) that discount rates may be somewhat greater for a 6-month delay than for longer delays but are constant for longer delays. Additionally, both specifications find discount rates to be significantly impacted by the length of education, retirement status, unemployment, and the likelihood of obtaining a loan or being approved for a credit card and yield similar marginal effects. The GG specification also finds income and owner-occupied housing to be statistically significant.

In particular, our results show that the GB2 family generally dominates the SGT as it provides a better fit with fewer parameters. Within the GB2 family, the 2-parameter and 3-parameter gamma (GA) and generalized gamma (GG) distributions are arguably the best choice, considering fit, parsimony, and ease of interpretation. An added advantage of the GB2 family over SGT is that an assumption of "heteroscedasticity" (making σ a function of covariates) is unnecessary, considerably simplifying the interpretation of parameters. For both the GA and GG, we find support for the HLW conclusion that rates appear to be somewhat greater at a 6-month delay than for the longer delays, but constant across the longer delays. We also find that in addition to the discount rate predictors found to be significant in HLW, our estimation of the GB2 model uncovers additional statistically significant covariates.

While our results generally confirm the results of HLW, we anticipate that further applications of this methodology will have important implications to estimation of other interval response data, particularly in the case of heteroscedasticity. To make these methods more accessible, a STATA module has been written and is currently being tested.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.socec.2017.10.003.

References

Anderson, S., Harrison, G.W., Lau, M.I., Rutstrom, E.E., 2008. Eliciting risk and time preference. Econometrica 76 (3), 583–618.
Ashley, R., Ball, S., Eckel, C., 2010. Motives for giving: a reanalysis of two classic public goods experiments. Southern Econ. J. 77 (1), 15–26.
Banerjee, P., Shogren, J., 2014. Bidding behavior given point and interval values in a second-price auction. J. Econ. Behav. Organiz. 97, 126–137.
Caudill, S.B., 2012. A partially adaptive estimator for the censored regression model based on a mixture of normal distributions. Statist. Methods Appl. 21, 121–137.
Charness, G., Gneezy, U., Imas, A., 2013. Experimental methods: eliciting risk preferences. J. Econ. Behav. Organiz. 87, 43–51.
Clarke, K.A., 2007. A simple distribution-free test for non-nested model selection. Political Anal. 15 (3), 347–363.
Coller, M., Williams, M., 1999. Eliciting individual discount rates. Exp. Econ. 2, 107–127.
Cook, J., McDonald, J.B., 2013. Partially adaptive estimation of interval censored regression models. Comput. Econ. 42, 119–131.
Dominitz, J., Manski, C., 1997. Perceptions of economic insecurity: evidence from the survey of economic expectations. Public Opinion Q. 61, 261–287.
Frechette, G., 2012. Session-effects in the laboratory. Exp. Econ. 12, 485–498.
Hanley, N., Kristrom, B., Shogren, J., 2009. Coherent arbitrariness: on value uncertainty for environmental goods. Land Econ. 1, 41–50.
Hansen, B.E., 1994. Autoregressive conditional density estimation. Int. Econ. Rev. 35 (3), 705–730.
Hansen, C., McDonald, J.B., Newey, W.K., 2010. Instrumental variables estimation with flexible distributions. J. Econ. Bus. Statist. 28 (1), 13–25.
Hansen, J., Jacobsen, R., Lau, M., 2013. Willingness to pay for insurance in Denmark. J. Risk Insurance 10, 2–28.
Harrison, G., 1986. An experimental test for risk aversion. Econ. Lett. 21, 7–11.
Harrison, G., 1992. Theory and misbehavior of first-price auctions: reply. Am. Econ. Rev. 82 (5), 1426–1443.
Harrison, G., Lau, M., Williams, M., 2002. Estimating individual discount rates in Denmark: a field experiment. Am. Econ. Rev. 92 (5), 1606–1617.
Harrison, G., Rustrom, E., 2008. Risk aversion in the laboratory. Res. Exp. Econ. 12, 41–196.
Kerman, S., McDonald, J.B., 2013. Skewness-kurtosis bounds for the skewed generalized t and related distributions. Statist. Probab. Lett. 83, 2129–2134.
McDonald, J.B., 1984. Some generalized functions for the size distribution of income. Econometrica 52 (3), 647–665.
McDonald, J.B., Newey, W.K., 1988. Partially adaptive estimation of regression models via the generalized t distribution. Econometric Theory 4 (3), 428–457.
Stewart, M.B., 1983. On least squares estimation when the dependent variable is grouped. Rev. Econo. Stud. 50 (4), 737–753.
Theodossiou, P., 1998. Financial data and the skewed generalized t distribution. Manage. Sci. 44 (12), 1650–1661.
Vuong, Q., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 50, 1–25.