Investigating the distribution of the value of travel time savings

Mogens Fosgerau
Danish Transport Research Institute
mf@dtf.dk

23 November 2004

Abstract

The distribution of the value of travel time savings (VTTS) is investigated employing various non-parametric techniques on a large, high quality data set. When background variables are not included in the model it is found that the right 13% tail of the distribution is not observed and hence the mean VTTS cannot be evaluated. This conclusion changes when background variables are introduced into a semiparametric model. A partially constrained Johnson S_B distribution allowing evaluation of the mean VTTS is accepted against the nonparametric alternative and is preferred among 16 candidates for parametric VTTS distributions. The resulting mean VTTS is plausible but three times larger than the mean VTTS evaluated from a simple logit model and half as big as that arising from a model assuming a lognormal distribution for the VTTS. Such findings indicate the importance of properly accounting for the distribution when estimating the mean VTTS. The present findings may be used to guide the choice of mixing distribution in a mixed logit model.

Keywords: value of travel time savings, VTTS, distribution, nonparametric, semiparametric, Klein-Spady, Zheng, Johnson S_B, lognormal

1 Introduction

1.1 Motivation

The value of travel time savings (VTTS) is arguably the single most important number in transport economics. Travel time savings usually constitute a very large share of total benefits in cost benefit analyses of infrastructure projects (Hensher, 2001a, Mackie et al., 2001) and cost benefit analyses are in turn a main part of the information provided to decision makers on new projects. It is not only the average VTTS that is important but also its distribution, e.g., when forecasting market share for a tolled road (Hensher & Goodwin, 2004).

The VTTS is usually inferred from experimental data using the logit model (Gunn, 2000). Recently the mixed logit has become the model of choice, since it allows for considerable improvements over the logit model in both realism and ability to describe the data (Train, 2003). The mixed logit model works by allowing certain parameters in the logit model to vary randomly in the population according to some parametric distribution. The (hyper)parameters for this mixing distribution can then be estimated.

There remains, however, the problem of deciding which mixing distribution to specify; some common choices are normal, lognormal, beta, Johnson's S_B or triangular (Hess et al., 2004). The choice of mixing distribution can have considerable impact on results (Hensher, 2001b, also Heckman & Singer, 1984), but little evidence exists to guide this choice. This paper similarly finds that the choice of mixing distribution can have a very strong effect on the resulting estimate of the mean VTTS and further that this effect is mainly due to the behavior of the tails of the mixing distributions outside the range of data. Since the range of data is always bounded, the tail behavior of distributions may in many cases not be constrained by data.

Thus the aim of this paper is a comprehensive study of the distribution of the VTTS, applying nonparametric and semiparametric methods to a large, high quality data set. These methods avoid strong prior distributional assumptions.

There has been little application of nonparametric and semiparametric estimators in the transport literature. Hensher & Greene (2003) stress the importance of the issue of selecting parameter distributions in mixed logit modeling and suggest applying a kernel density estimator to parameter estimates after applying a jackknife procedure to a multinomial logit model. This method allows one to visually inspect the distribution of parameters, however, without confidence bands on the estimated densities. Their findings further suggest that a wide range for the variables in a stated choice design is preferable, something which the results here also indicate.

1.2 Nonparametric and semiparametric regression

Introductions to nonparametric and semiparametric econometrics are given, e.g., in Yatchew (2003), Pagan & Ullah (1999) and Härdle (1990). Consider the regression model $y = f(x) + \varepsilon$, where the point of interest is to determine the function f. The classical OLS regression assumes that the function f is linear in parameters and estimates these. Nonparametric kernel smoothers avoid such parametric assumptions by instead averaging the y's in the neighborhood of each x. The average of the y's then converges to f(x) under weak assumptions. This is a data hungry procedure, especially as the number of dimensions in x grows. Therefore semiparametric methods have been developed as a hybrid between parametric and nonparametric regression where just some of the relationship is modeled nonparametrically.

The results presented here use a two-step estimation procedure suggested by Lewbel, Linton & McFadden (2002).[1] In the first step the Klein & Spady (1993) estimator is used to estimate parameters in a linear index binary choice model without assuming a distribution for the error terms. In the second step the distribution of the error term is estimated. Details are given in the section below on semiparametric methodology.

1.3. Layout

The paper is organized as follows. Section 2 sets out the methodology. Section 3 presents a recent large dataset collected in a Danish value of travel time study, which is used in section 4 to investigate the stochastic distribution of the value of time without using covariates. Section 5 introduces background variables to explain the distribution and estimate the mean VTTS and section 6 concludes.

2 Methodology

2.1 Transformation of the data to contingent valuation format

Our data come from a stated preference exercise where respondents are presented with binary choice situations (Burge et al., 2004). The data are transformed into a format similar to contingent

[1] Koning and Ridder (2003) employ a similar idea.

valuation data, where we observe y=1 if a random latent w is smaller than a bid v set by experimental design.

Alternatives 1 and 2 are characterized by travel time t_i and cost c_i only; they are otherwise the same. The conventional model for this situation is the binary (mixed) logit model (Train, 2003). Here we merely specify conditional indirect utility functions $\alpha_t t_i + \alpha_c c_i$, where parameters α_t, α_c <0 are random and independent across observations.[2] It is not necessary to specify errors corresponding to the logit kernel for the present purpose (McFadden & Train, 2000), since alternatives are identical except for time and cost, observations are independent and parameters are random.

Alternative 1 is chosen if $\alpha_t t_1 + \alpha_c c_1 > \alpha_t t_2 + \alpha_c c_2$. It is customary to estimate separate parameters for time and cost. Here we shall use a form that directly emphasizes our object of interest: the value of time. Rearrange alternatives such that $t_1 < t_2$. We observe

$$y = 1\{\alpha_t t_1 + \alpha_c c_1 < \alpha_t t_2 + \alpha_c c_2\} = 1\{\alpha_t/\alpha_c < -(c_1 - c_2)/(t_1 - t_2)\},$$

where 1 is the indicator function. That is, we observe y=1 when the respondent is not willing to pay to have the faster alternative. Let $w = \alpha_t/\alpha_c$ be the random, unobserved VTTS and let $v = -(c_1-c_2)/(t_1 - t_2)$ be the bid VTTS presented in the choice experiment. We thus observe $y = 1\{w<v\}$, such that y is 1 if the value of time is smaller than the bid and the respondent is not willing to pay the stated difference to save the stated time difference.

[2] The panel data nature of data where respondents make repeated choices is thus ignored.

We discard observations where there is a dominant alternative, since such observations do not provide any information about the value of travel time, which cannot be negative (Hess et al., 2004). Then v > 0 and y = 1{w<v} = 1{log(w) < log(v)}.

2.2 Nonparametric estimation of the VTTS distribution

Note that P(y = 1) = P(w < v) = $F_w(v)$, where F_w is the c.d.f. of w. We can write y = $F_w(v) + \eta$, where $E(\eta)=0$ and estimate F_w by a nonparametric regression of y on v given weak smoothness conditions on F_w.

The mean of w can be estimated from the estimated F_w. Evidently, this requires that the range of v extends over the support of F_w. If not, then only a part of the distribution is observed. In this case the moments of F_w cannot be estimated without further assumptions.

The nonparametric regressions in this paper are all performed using the Nadaraya-Watson estimator, with a first-order normal density kernel and bandwidth selected by eyeballing (e.g. Pagan & Ullah, 1999).

Computation of confidence intervals is complicated slightly since η is heteroscedastic. The asymptotic 95% pointwise confidence interval at v is computed using

$$\hat{F}_u\left(v\left(\hat{\beta}\right)\right) \pm 1.96\sqrt{\frac{b_K \hat{\sigma}^2(v)}{\hat{p}(v)\lambda n}} \tag{1}$$

where $\hat{F}_u$ is the estimated regression function, $b_K = \frac{1}{2\sqrt{\pi}}$, $\hat{p}$ is a nonparametric estimate of the density of v, λ is the bandwidth, and n is the number of observations. The variance of η is estimated at each point (Pagan & Ullah, 1999, p.104); we use $\hat{\sigma}^2(v) = \hat{F}_u(v)\left(1 - \hat{F}_u(v)\right)$ since y can only take the values 0 and 1. The pointwise confidence intervals will contain 95% of the true points of the distribution function in repeated samples as n tends to infinity.

Uniform 95% confidence intervals are computed using the formula given in Yatchew (2003), where 1.96 in formula (1) above is replaced by

$$\frac{-\log(-\log(0.95)/2)}{\sqrt{2\log(1/\lambda)}}+\sqrt{2\log(1/\lambda)}+\frac{1}{\sqrt{8\log(1/\lambda)}}\log\left(\frac{1}{8\pi^2}\right).$$

The uniform confidence bands will contain the entire true distribution in 95% of repeated samples as n tends to infinity. Section 2.3 below illustrates the ability of this technique to recover various distributions.

2.3 Examples: Recovery of various distributions

This section illustrates the ability of nonparametric regression to identify cumulative distribution functions using only observations of binary choices. We have constructed 2000 observations where the value of time w follows some known distribution and bids v are taken from a normal distribution with mean ½ and standard deviation ¼. We observe v and y=1(w<v) and regress y on v using a normal density kernel with bandwidth 0.04, selected by eyeballing.

Four test distributions for w are selected to represent a variety of shapes as shown in Figure 1. The distributions are lognormal, a "double sawtooth" distribution constructed from combining two triangular densities, a normal distribution and a beta distribution. The parameters for the distributions are chosen such that almost all the mass lies within the unit interval, so that the requirements for the estimator in section 2.2 are satisfied fairly closely.

<Figure 1 about here>

Figure 2 shows the true distribution functions together with 95% pointwise confidence bands and 95% uniform confidence bands for each estimated distribution.

<Figure 2 about here>

The estimated uniform confidence bands are narrow and contain the true density in all cases, except for the lowest values in the lognormal distribution, and mostly the true distribution functions lie inside the estimated pointwise confidence intervals.

2.4 Nonparametric estimation of the VTTS distribution using covariates

Consider now the case when background variables are introduced as covariates. Assuming that log(w) depends on background variables x through a single linear index, we parameterize $\log(w) = \beta x + u$, where x is a vector of observed variables and u is an error, independent of x and with unknown c.d.f. named F_u. Then

$$P(y=1) = F_u(\log(v)-\beta x) = F_u(v(\beta)),$$

where $v(\beta)=\log(v)-\beta x$.

We have now $y = F_u(v(\beta)) + \eta$. If observations of $v(\beta)$ were available, we could just perform a non-parametric regression of y on $v(\beta)$ to estimate F_u. Lewbel et al. (2002) propose instead to regress y on estimated values $v(\hat{\beta})$, which yields a consistent estimate $\hat{F}_u$ of F_u, since $\hat{\beta}$ converges faster than $\hat{F}_u(v(\beta))$. The necessary assumptions are that $v(\beta)$ has compact support; the unknown distribution of w has a twice continuously differentiable, strictly monotonic, conditional c.d.f. and density with compact support; the test variable v has a continuous distribution with compact support extending over that of w; w and v are independent conditionally on $v(\beta)$ so that u and v are independent; u has compact support that contains zero; and $v(\beta)$ has support that extends over the support of u.

The nonparametric estimator described above is also applied for the case with covariates. The mean VTTS can be assessed since $E(w) = E(\exp(\hat{\beta}x + u)) = E(\exp(\hat{\beta}x))E(\exp(u))$ by independence of x and u.

2.5 The Klein & Spady estimator

The index parameters β are estimated using the Klein & Spady (1993) estimator. The (not logged) likelihood of an observation is

$$L^* = yP(y=1) + (1-y)P(y=0) = y\, F_u(v(\beta)) + (1-y)(1- F_u(v(\beta))).$$

Given F_u it would be possible to estimate β by maximum likelihood. For example, assuming a normal or a logistical distribution for u gives rise to a probit or logit model. We do, however, only want to impose minimal prior assumptions on F_u.

Klein & Spady (1993) propose to replace $F_u(v(\beta))$ by a nonparametric estimate that depends on β. We have $y = F_u(v(\beta)) + \eta$, such that given β we can perform a nonparametric regression of y on the index $v(\beta)$. We again use the Nadaraya-Watson estimator with a normal density kernel to find an estimate of $F_u(v(\beta))$ as $\tilde{F}_u(v(\beta))$. $\tilde{F}_u(v(\beta))$ has a closed form expression as a weighted average of y around $v(\beta)$. Then we can approximate the likelihood function for y by

$$L^* \approx y\,\tilde{F}_u(v(\beta)) + (1-y)\,(1- \tilde{F}_u(v(\beta)))$$

and maximize with respect to β in order to arrive at a semiparametric estimate $\hat{\beta}$ of β. This estimator is consistent and asymptotically normal under weak conditions (see also Pagan & Ullah, 1999). Standard errors are computed from the estimated Hessian (Klein & Spady, 1993). Finally, the $v(\hat{\beta})$ are computed. We shall refer to these as Klein-Spady residuals.

2.6 Testing parametric distributions

Given some parametric distribution for the VTTS we would like to test this against the nonparametric alternative. The Zheng (1996) test statistic is available for this. Letting the parametric model be given by $y = F(x;\theta) + \eta$, then the null hypothesis to be tested is that $P(E(y|x) = F(x;\theta_0)) = 1$ for some θ_0, while the alternative hypothesis is that $P(E(y|x) = F(x;\theta)) < 1$ for all θ. The alternative encompasses all possible departures from the null. Define residuals $e = y - F(x;\hat{\theta})$, where $\hat{\theta}$ is an estimate of θ and let K be the normal density kernel and again λ be the bandwidth. Then the Zheng test statistic is computed as

$$T = \frac{\sum_i \sum_{j \neq i} K\left(\frac{x_i - x_j}{\lambda}\right) e_i e_j}{\sum_i \sum_{j \neq i} 2K^2\left(\frac{x_i - x_j}{\lambda}\right) e_i^2 e_j^2}.$$

This is distributed as N(0,1) under the null hypothesis. Under the alternative hypothesis $T / n\sqrt{\lambda}$ converges in probability to a constant, and thus T converges to infinity when n grows faster than $\lambda^{-1/2}$.

Eight different distributions are tested, specified as indicated in Table 1, where F denotes cumulative distribution functions and f denotes densities. Logged versions of the distributions are obtained by applying them to log(x) instead of x, which yields a total of sixteen parametric distributions to be tested. The Beta and the Johnson S_B distribution occur in two versions, according as the upper bound is constrained to be the maximum of the independent variable.

3 Data

The data origin from a recent Danish value of time study undertaken by the Danish consultancy TetraPlan in joint venture with Rand Europe and Gallup for the Danish Ministry of Transport. Stated preference interviews have been conducted both via the Internet and computer aided personal interviews. Business travel is excluded. The stated preference design is discussed in Burge et al. (2004). Here only data from one experiment are used. From the data, a sample of 2197 interviews is selected of car drivers choosing between car trips distinguished by cost and time only. Each respondent has made 9 consecutive choices. One of these is always a dominated choice, which is excluded from the analysis. Respondents that do not choose the dominant alternative are discarded. Observations with errors such as unrealistic speeds, very long journey times etc. are rejected such that 17020 observations are available for estimation.

Time is described in the experiment as free flow travel time and additional time due to congestion in order to make the resulting VTTS estimates applicable to the output from an assignment model. The two time components vary proportionally such that the relative share of time under free flow conditions stays the same for each respondent through the choices made by each.

Table 2 presents some summary statistics for the data.[3] The bids in v are of particular interest. They are chosen in the design to cover the expected range for the mean value of time, not the range of individual values of time. Eight v's are chosen randomly from a number of specified ranges, in order to incorporate a range of low and high values, with most values in the region of the range of the ex ante expectation of mean VTTS. The range of VTTS bids in v is fairly large, from 3 to 201

[3] The currency is Danish Kroner: 7.5 DKK=1 EUR.

DKK/hour, restricting our ability to infer anything about the distribution of the VTTS above this interval. The maximum is imposed by the design (Burge & Rohr, 2004).

If we were not willing to parameterize the location of log(w) as βx, then v would have to be drawn from a continuous distribution in order to identify F_u (Lewbel et al., 2002, Appendix 1). Parametrising with β as we do means that only v(β) must be continuous, which is easier to achieve, at least approximately. It is, however, still relevant to know the distribution of the v's. The cumulative distribution of v in Figure 3 shows a nice dispersion of the values of v with most bids below 100 DKK/hour but also a significant number above.

<Figure 3 about here >

4 The VTTS distribution without covariates

First consider nonparametric regression of y on v. Before the regression v is transformed to logs; this affects the regression through the bandwidth such that in effect the bandwidth is larger for large v where data are sparser.[4] In the estimation, data have been rescaled to lie within the unit interval and a bandwidth of 0.03 is employed, selected by eyeballing. Afterwards, results are transformed back to the original range. The regression provides an estimate of the distribution of the VTTS in w, shown in Figure 4.[5]

---

[4] Using a design adaptive estimator is an alternative.

[5] Results are generated using Ox 3.30, (see Doornik, 2002). The code is available from the author on request.

A number of observations can be made from this regression. First, there is definitely a positive slope, which means that as the bid increases, more respondents decline to save time. Second, confidence bands are fairly tight, which means that choice probabilities can be assessed with a reasonable degree of accuracy and also the corresponding quantiles of the VTTS distribution. Third, there exists a monotone function within the confidence bands, which is consistent with the estimated function being a distribution function. Fourth, the distribution can be assumed to tend to zero at zero VTTS. We shall adopt this assumption henceforth. Fifth, the distribution does, however, not tend to one within the observed range. At the largest bid presented, there is a significant proportion of respondents who are willing to pay more in order to save time. The point estimate of $F_w(201)$ is 0.867 [0.834;0.899]. This in turn means that the mean VTTS cannot be assessed from this distribution.[6]

< Figure 4>

It is of interest to test various parametric distributions against the nonparametric distribution. Sixteen different distributions are estimated by maximum likelihood. Parameter estimates and loglikelihoods are shown in Table 3. The distributions are shown in Figure 5 and Figure 6 with the nonparametric pointwise confidence bands indicated for comparison. The best fitting distributions achieve loglikelihoods above –10750. Of these, the parameter estimates of the Beta, S_B and the $\log S_B$ distributions are extreme and standard errors could not be computed for the Beta and the $\log S_B$; otherwise parameter estimates are very significant. The Gamma distribution, having only two parameters, is preferred over the other distributions on a likelihood ratio test.

---

[6] Discarding observations of respondents who choose always the cheapest or the fastest alternative does not change this conclusion.

Each of the sixteen distributions is compared to the nonparametric alternative using the Zheng (1996) test; the test statistics shown in Table 4 converge in probability to the standard normal under the null hypothesis indicating a clear rejection of all sixteen parametric distributions. Presumably, this is due to the wriggles evident on the nonparametric estimate of the distribution.

The mean VTTS has been computed for each of the 16 parametric distributions. Values below zero have been truncated, with the interpretation that the distributions concerned have a point mass at zero. The results are shown in Table 5. An extremely large degree of variation is found when using different distributions. At the low end the normal distribution has a mean VTTS of 51.4 DKK/hour. The typical mean VTTS arising from these distributions is about 60 DKK/hour. However, the log-normal and in particular the loggamma have very long right tails as can be seen from Figure 6, leading to very high mean VTTS's, outside the range of data. The distributions with loglikelihoods better than –10750 lead to estimated mean VTTS ranging from 59 to 96 DKK/hour, a fairly wide range.

In summary, the nonparametric regression shows that the right 13% of the VTTS distribution is not observed. Thus the mean VTTS cannot be evaluated without some additional assumption about the right tail outside the range for which data exist. Founding such an assumption seems to be hard. Sixteen common parametric models have all been rejected. Choosing anyway those with the best loglikelihood lead to a wide range of estimated mean VTTS's.

5 The VTTS distribution including covariates

5.1 A semiparametric model

In this section the model is expanded by the inclusion of various covariates in a semiparametric model combining some parameterization with an additive nonparametric error. As indicated in section 2 this is achieved by specifying a model where log(w) is split into a parametric linear index plus an independent error. Descriptive statistics for the covariates in the index are provided in Table 6. The variables are mostly self-explanatory. The personal income is given in bands with 1 representing the interval from 0 to 100,000 DKK per year up to 11 representing the interval above 1000,000 DKK per year. Trip duration is defined as the average of the travel time in minutes in the two alternatives presented to the respondent, the log of the travel time difference between the two alternatives and the share of congestion time, denoted by s, which is the ratio of additional congestion time over free flow time to total time. Since the log function is nonlinear, a first order Taylor expansion can be used to arrive at the interpretation of the coefficient of s as a markup for congested time over the value of free flow time.

Panel (a) of Table 7 shows the parameter estimates and summary statistics for the Klein-Spady estimator. All parameters, except the first order term for age, are significant at 5% and most highly so. The model is extremely significant in a chi-square test against a model with zero parameters.

The VTTS of females is on average 25% (4%)[7] lower than males. The coefficient of personal income can be directly interpreted as an income elasticity of 0.68 (0.05). This is in line with, e.g., a review by Wardman (2001) who finds a typical household income elasticity of 0.6 on cross-

[7] Standard deviations in parentheses.

sectional data and indication that higher values are found when individual incomes are used. Note that the data involve pretax income. The Danish tax system is quite progressive: a back-of-an-envelope calculation indicates that the elasticity with respect to after tax income would be higher and close to 1.

The VTTS increases with the duration of the trip, here represented as the log of the average travel time in the two alternatives presented, such that the parameter becomes a VTTS elasticity of trip duration of 0.17 (0.03). The VTTS increases also with the size of the time saving with an implied elasticity of 0.36 (0.04). This is consistent with the finding of Hultkrantz & Mortazavi (2001), who discuss this effect as a result of a perceptual threshold, which may reflect either a real social cost or a decision rule employed in the course of completing a questionnaire. The congestion share is significant with a parameter that indicates that congested time is valued 52% (14%) above free flow time. The VTTS decreases with age. The dummy for commuting is positive while the dummy for travel to/from education may be either zero or equal to the commuting dummy.

5.2 Analysis of the Klein-Spady residuals

The nonparametric regression of y on the Klein-Spady residuals $v(\hat{\beta})$ is shown in Figure 7. This is an estimate of the distribution of the error u. A number of observations are possible. First, the wriggles have disappeared and the distribution seems much smoother than before. Second, the confidence bands are tight over a long range. Third, the distribution can be assumed to be bounded within the support of $v(\hat{\beta})$, since 0 and 1 are within the confidence bands at the ends of the distribution. Making this assumption makes it possible to compute the mean VTTS.

We again test a range of parametric distributions against the nonparametric alternative. Parameter estimates and loglikelihoods can be found in Table 8. Table 9 shows the Zheng test values for the sixteen distributions and Table 10 shows the corresponding mean VTTS's. The latter are computed by E(exp(u+βx)) = E(exp(u))E(exp(βx)), where E(exp(βx)) is replaced by its sample mean of 6.1907. The mean of exp(u) is calculated analytically where possible and simulated otherwise with truncation at zero where necessary. Graphs of the estimated distributions are shown in Figure 8 and Figure 9 with nonparametric confidence intervals indicated for comparison.

There are now a number of distributions that can be accepted against the nonparametric alternative for the distribution of the error term. These distributions also achieve the best loglikelihoods all better than –10143. We shall discuss each in turn.

The Lognormal is simple and is commonly used, parameter estimates are of reasonable size, although the estimated standard deviation of 2.0 seems fairly high and leads to a long tail. It is not bounded within the data and the estimated mean VTTS is the highest among the accepted distributions with a value of 183.6 DKK/hour. The S_B, logS_B and logBeta distributions are not bounded within the data, they lead to extreme parameter estimates and standard deviations are very high or sometimes not possible to calculate. Finally, the S_B1 distribution is specified such that its upper bound coincides with the upper bound of the Klein-Spady residuals. This leads to a lower mean VTTS of 105.1 DKK/hour. Except for the parameter for the lower bound, the standard deviations are small. It is accepted against the nonparametric alternative but rejected against the S_B distribution.

Requiring that of a distribution that it should be accepted against the nonparametric alternative, not have a long tail outside the range of data, and not have unreasonable parameters leads to the choice of the S_B1 distribution as the preferred parametric distribution. In applications focusing on prediction rather than on evaluating the mean VTTS, the lognormal might be preferred instead since it achieves a better fit.

5.3 Two parametric models

The index used to calculate the Klein-Spady residuals in the above analysis was based on the semi-parametric regression making no assumption on the distribution of u. Having now identified two candidates for the distribution of u, the lognormal and the Johnson S_B, we now estimate parametric models using these distributions. The results are shown in Table 7, panels (b) and (c) along with the semiparametric estimates. The parameter estimates change very little, reinforcing our conclusion that the two parametric distributions provide good approximations to the nonparametric distribution. The loglikelihoods improve relative to the parametric models on the Klein-Spady index, when the parameters in the index are estimated simultaneously with the parameters of the distributions. With the index from the S_B1 model we finally compute the mean VTTS as above, which yields a value of 89.2 DKK/hour. This would be our estimate of the mean VTTS.[8]

For completeness we finish by also estimating a logit model with specification chosen to include the index used above. We specify the indirect utility difference as $\Delta t \exp(\beta x) + \Delta c$ multiplied by a scale parameter, where x now includes a constant. The resulting estimated mean VTTS turns out to be 28.3 DKK/hour, less than a third of the value obtained with the S_B1 model.

---

[8] The value applies to the sample and is not directly suitable for use in economic evaluation. The official Danish VTTS figures for use in economic evaluation of transport projects are based on an average after tax hourly wage. For commuting by car the official value is about 60 DKK/hour for free flow time and a markup of 50% is applied for time under congested conditions (Trafikministeriet, 2003).

6 Conclusions

This paper has demonstrated the application of various nonparametric and semiparametric methods to the estimation of the distribution of the value of travel time savings and further to estimate the mean of the distribution. It is shown possible to estimate the VTTS distribution quite precisely with narrow confidence bands. However, when using no covariates except for the trade-off between time and cost, results indicate that the right tail of the VTTS distribution is not observed and hence the mean cannot be calculated without further assumptions. This observation may have implications for attempts to identify similar means from, e.g., the mixed logit model where applications have not verified that the mixing distribution is adequate and that the data allow observation of its tails. The observation further emphasizes the importance of the recommendation by Hensher & Greene (2003) to include a wide range for the variables in a stated choice design. It is possible to speculate that the missing tail is the reason behind the wide variability in the mean value of travel time found using different model specifications (e.g., Hensher, 2001b). The solution may then not be found through elaboration of the model specification but is rather to be found in the possibilities allowed for by the data.

Including a range of covariates in a semiparametric model resulted in plausible parameter estimates for a VTTS index allowing ready interpretation. Furthermore, it proved possible to accept that the distribution of the unknown error is bounded within the range provided by the estimated index. This in turn enabled estimation of the mean VTTS. The main candidate for a parametric distribution among sixteen distributions tested turned out to be the Johnson S_B with a fixed upper bound corresponding to the supremum of the index. In applications where prediction, e.g. of patronage of a tolled road, and not estimation of the mean is the issue, also the lognormal turned out to be a suitable choice. The estimated mean VTTS from the partially constrained Johnson S_B model turned out as 89.2 DKK/hour. In contrast the estimated mean VTTS from a logit specification was only 28.3 DKK/hour. The consequences of improving the econometric methodology as suggested here are thus dramatic.

The work presented here could be extended in several directions. First, it seems worthwhile to test the findings on other datasets. Perhaps the most pressing methodological issue is to deal with serial correlation, which does not seem possible with the present methodology. We have treated choices made by the same respondent as independent, which ignores both serial correlation and the fact that respondents may make mistakes when making choices. Accounting for mistakes may conceivably change the conclusions arrived at in the present paper and should be a subject for further research. Another line of research that seems worth pursuing is to extend the methodology to tackle several time components in a more direct way than how additional congestion time was included into the model presented here.

7 References

- Burge,P., Rohr, C. Vuk, G., Bates, J. (2004) "Review of international experience in VOT study design", Proceedings of the European Transport Conference.
- Burge, P., Rohr, C. (2004) "DATIV: SP Design: Proposed approach for pilot survey", Tetra-Plan in cooperation with RAND Europe and Gallup A/S.
- Doornik, J.A. (2002), "Object-Oriented Matrix Programming Using Ox", 3rd ed. London: Timberlake Consultants Press and Oxford: www.nuff.ox.ac.uk/Users/Doornik.
- Gunn, H. F. (2000) "An Introduction to the Valuation of Travel-Time Savings and Losses", Handbook of Transport Modelling, Chapter 26, Eds. D.A. Hensher & K. J. Button, Elsevier Science Ltd.
- Heckman, J. and Singer, B. (1984) "A method for minimizing the impact of distributional assumptions in econometric models for duration data", Econometrica, Vol. 52 No. 2, 271-320.
- Hensher, D.A. (2001a) "Measurement of the Valuation of Travel Time Savings", Journal of Transport Economics and Policy, Volume 35, Part 1, January 2001, 71-98.
- Hensher, D.A. (2001b) "The sensitivity of the valuation of travel time savings to the specification of unobserved effects", Transportation Research Part E 37, 129-142.
- Hensher, D.A. and Goodwin, P. (2004) "Using values of travel time savings for toll roads: avoiding some common errors", Transport Policy 11, 171-181.

- Hensher, D. and Greene, W. H. (2003) “The Mixed Logit Model: The State of Practice”, Transportation 30(2), 133-176.

- Hess, S., Bierlaire, M. and Polak, J.W. (2004) “Estimation of value of travel-time savings using Mixed Logit models”, Forthcoming Transportation Research Part A.

- Hultkrantz, L. and Mortazavi, R. (2001) “Anomalies in the Value of Travel-Time Changes”, Journal of Transport Economics and Policy, Volume 35, Part 2, May 2001, 285-300.

- Härdle, W. (1990), “Applied Nonparametric Regression”, Econometric Society Monograph Series, 19, Cambridge University Press.

- Klein, R. and R. Spady (1993), “An Efficient Semiparametric Estimator for Binary Response Models”, Econometrica, 61, 387-422.

- Koning, R.H. and G. Ridder (2003) “Discrete choice and stochastic utility maximization”, Econometrics Journal, volume 6, pp. 1–27.

- Lewbel, A., O. Linton and D. McFadden (2002), “Estimating Features of a Distribution from Binomial Data”, mimeo.

- Mackie, P.J., Jara-Diáz, S. and Fowkes, A.S. (2001) “The value of travel time savings in evaluation”, Transportation Research Part E 37: 91-106.

- McFadden, D. and Train, K. (2000) “Mixed MNL models of discrete response”, Journal of Applied Econometrics 15, 447-470.

- Pagan, A. and A. Ullah (1999), “Nonparametric Econometrics”, Cambridge: Cambridge University Press.

- Train, K. (2003), “Discrete Choice Methods with Simulation”, Cambridge University Press.
- Trafikministeriet (2003), ”Nøgletal”, www.trafikministeriet.dk.
- Wardman, M. (2001) ”Inter-temporal variations in the value of time”, ITS Working Paper 566, ITS Leeds, UK.
- Yatchew, A. (2003) “Semiparametric Regression for the Applied Econometrician”, Themes in Modern Econometrics, Cambridge University Press.
- Zheng, J.X. (1996) “A consistent test of functional form via nonparametric estimation techniques”, Journal of Econometrics 75, 263-289.

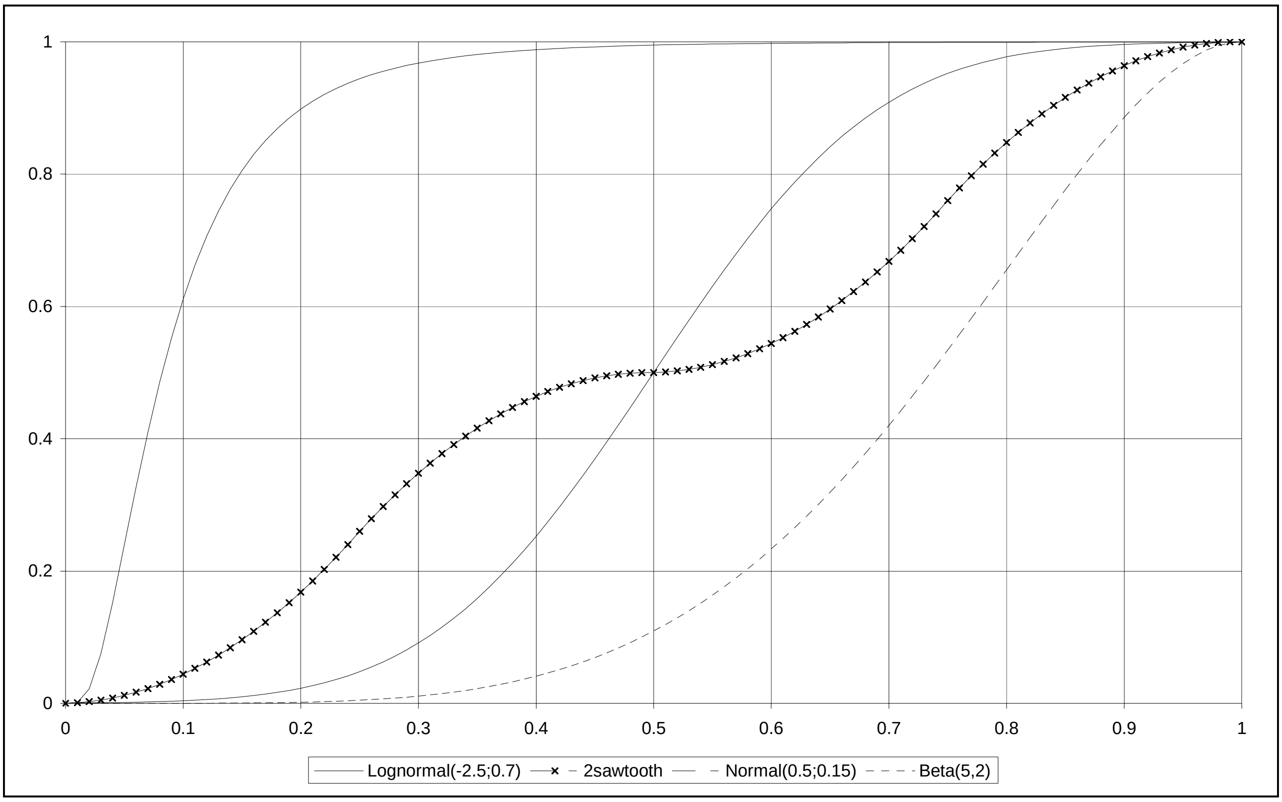

Figure 1 Four test distributions

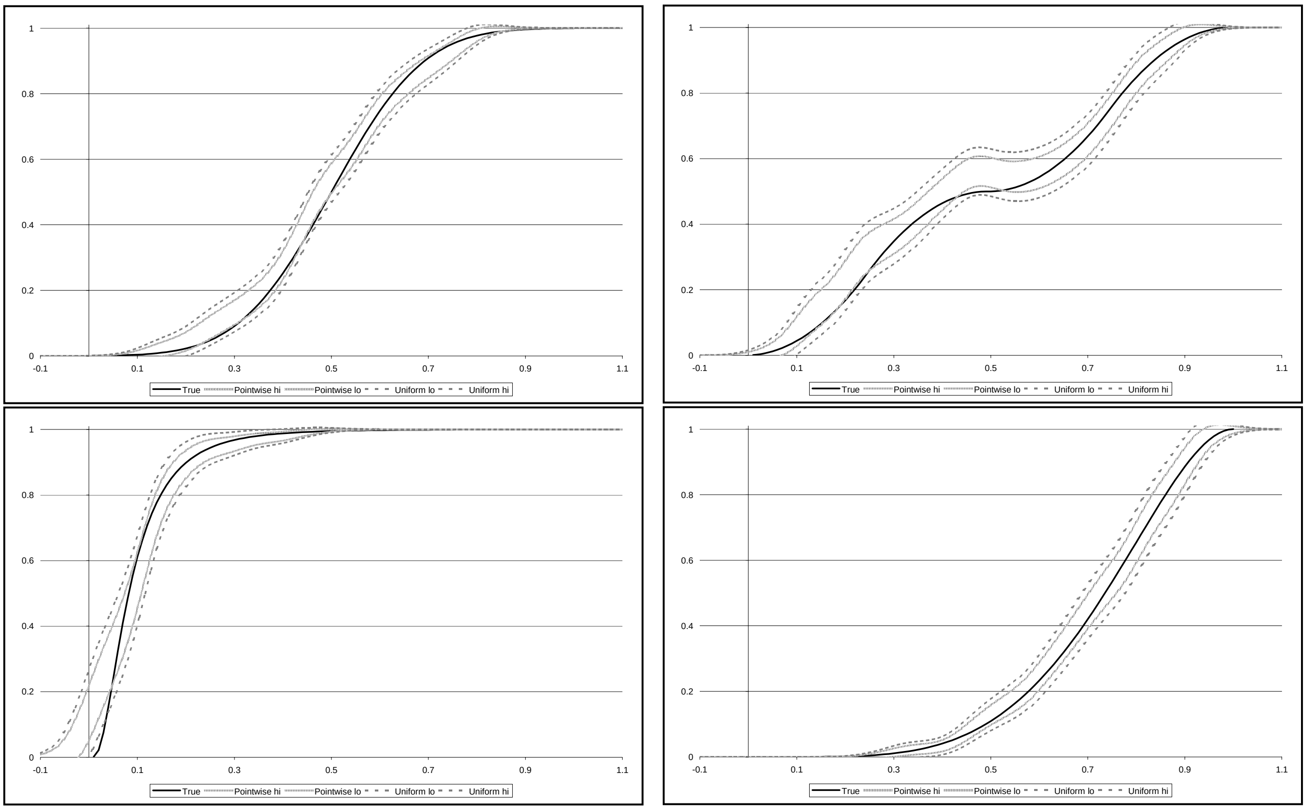

Figure 2 Estimated confidence bands around true distributions

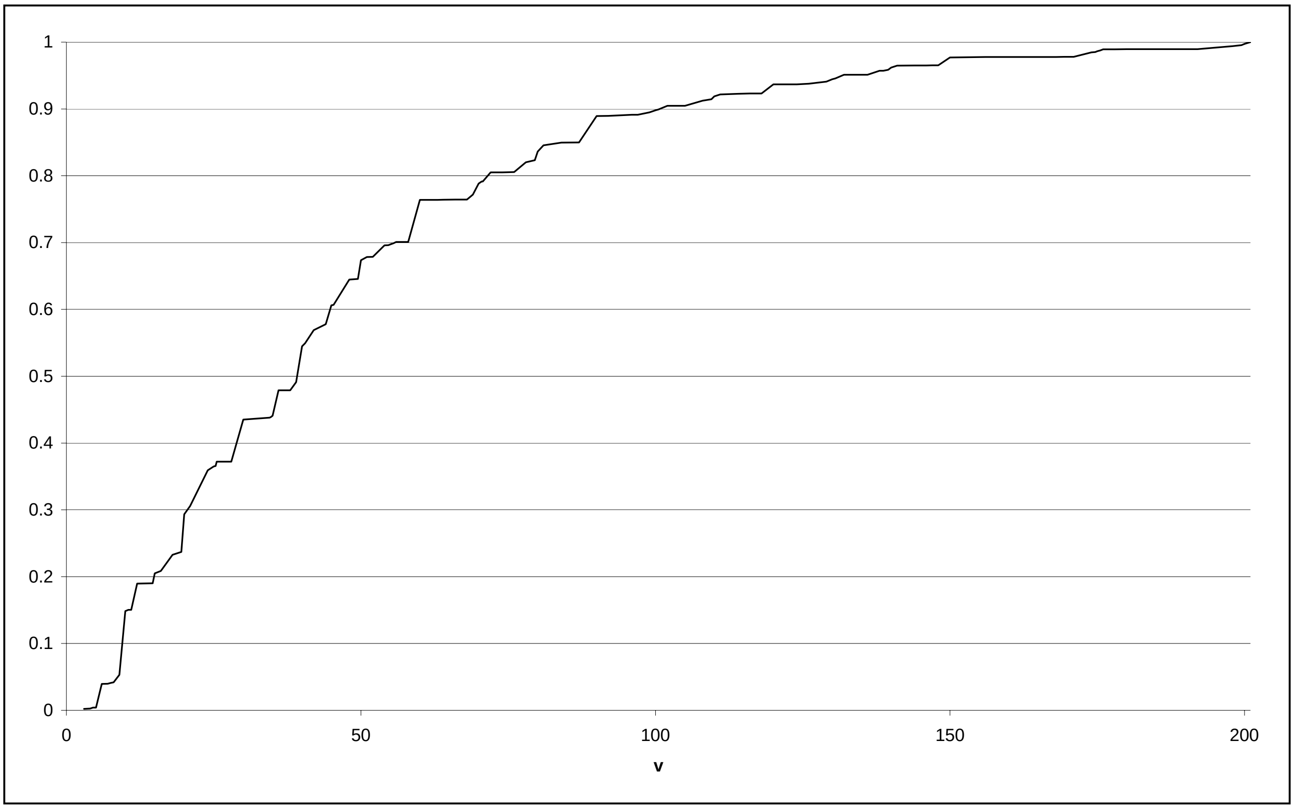

Figure 3 Cumulative distribution of bids v

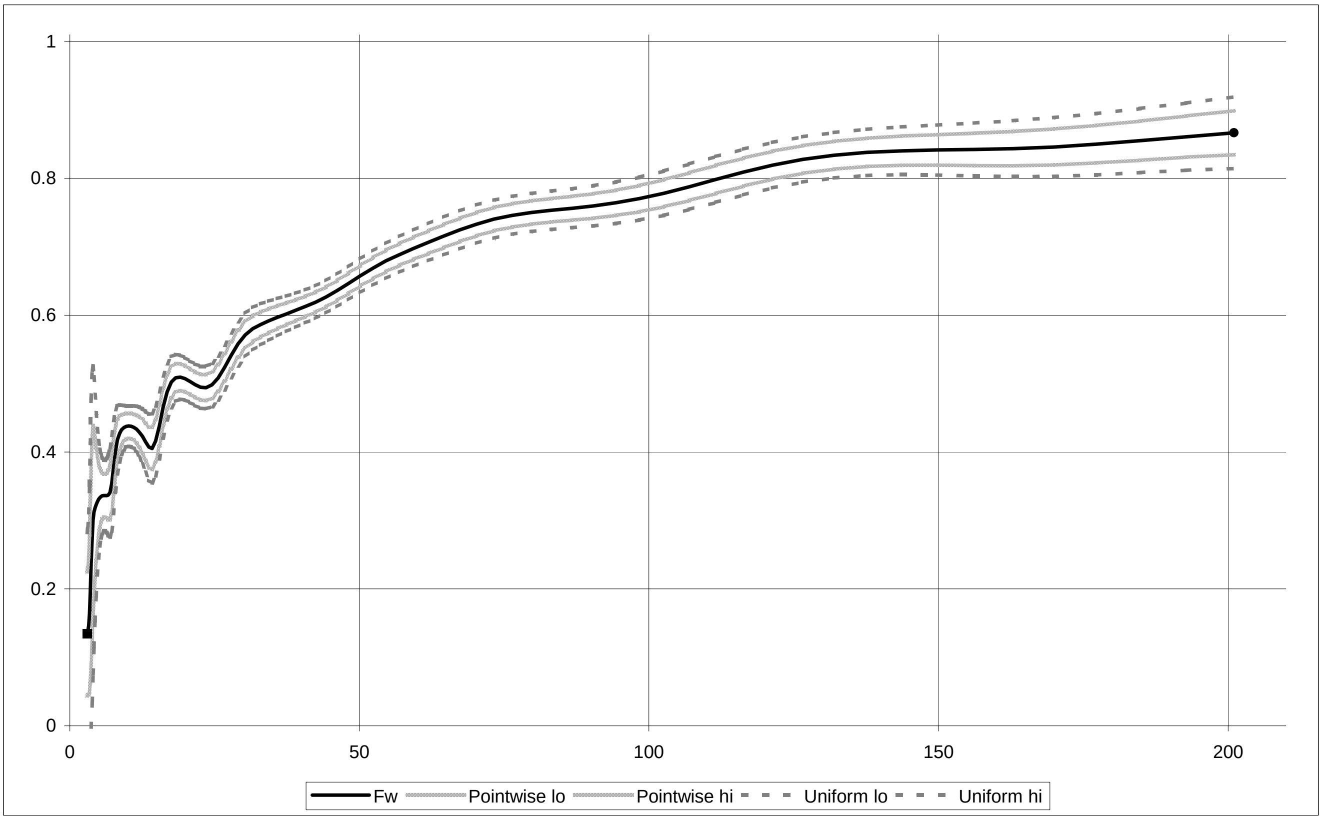

Figure 4 Nonparametric regression of choices y against bids v

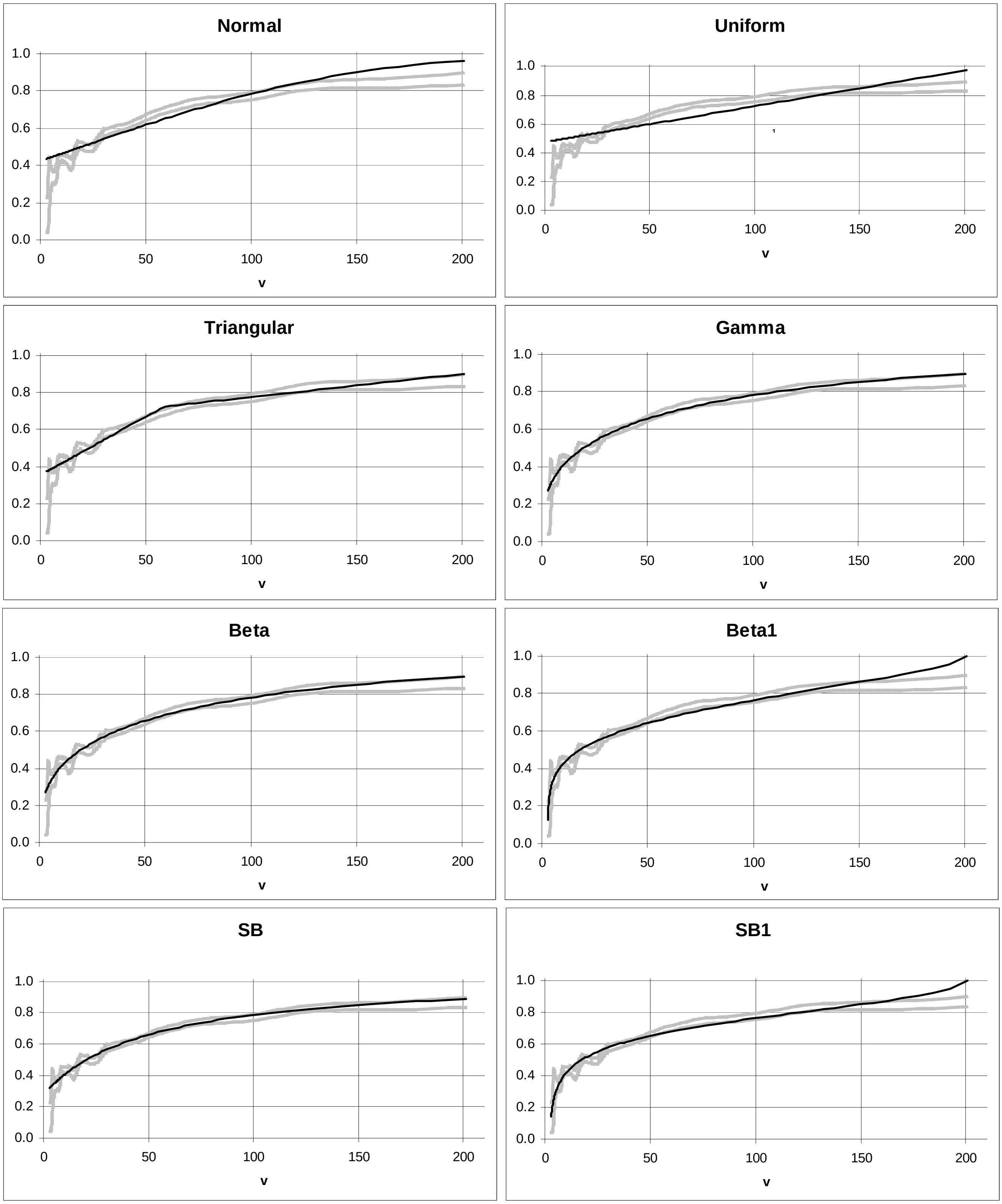

Figure 5 Eight parametric distributions estimated on bids v

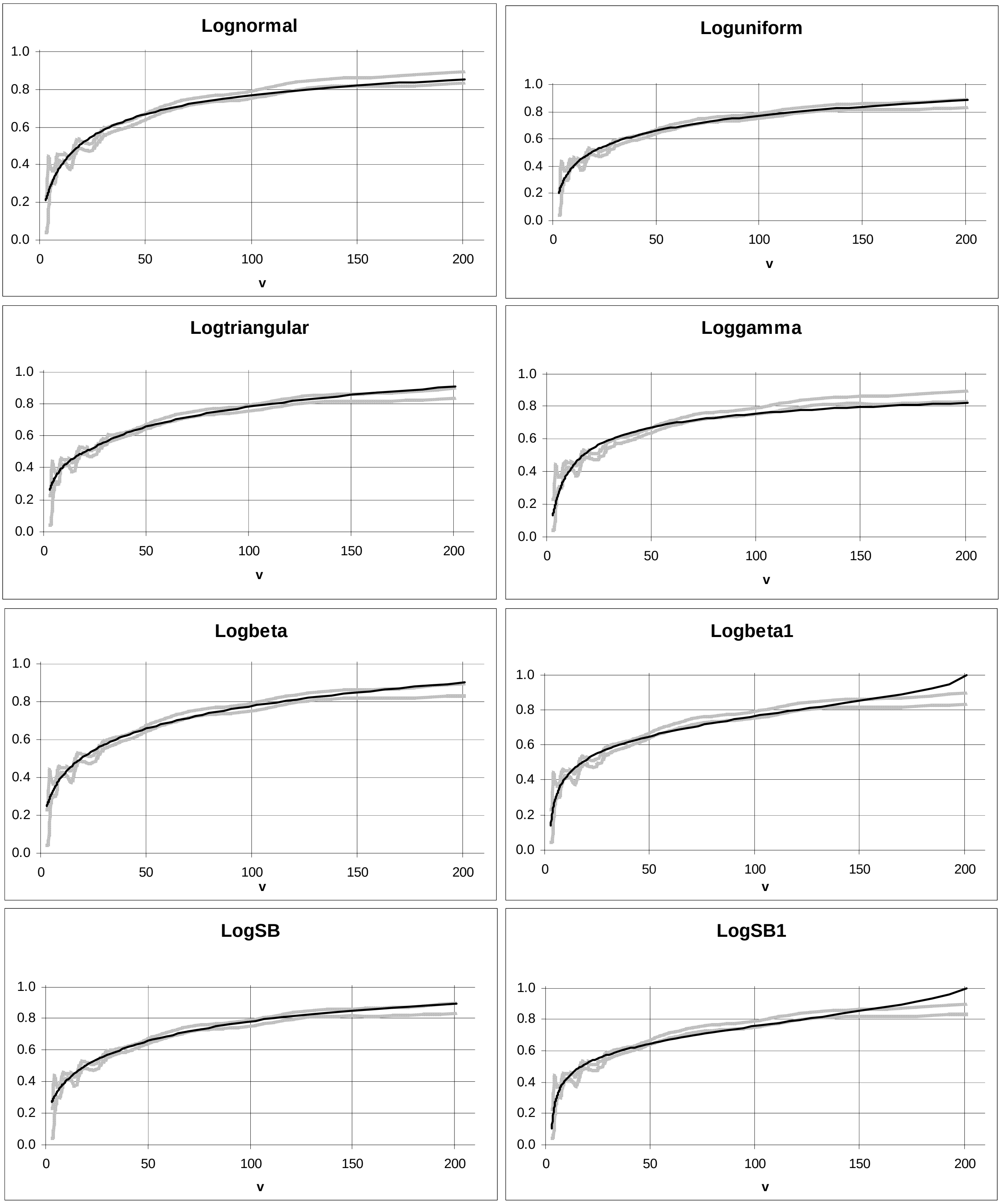

Figure 6 Eight logged parametric distributions estimated on bids v

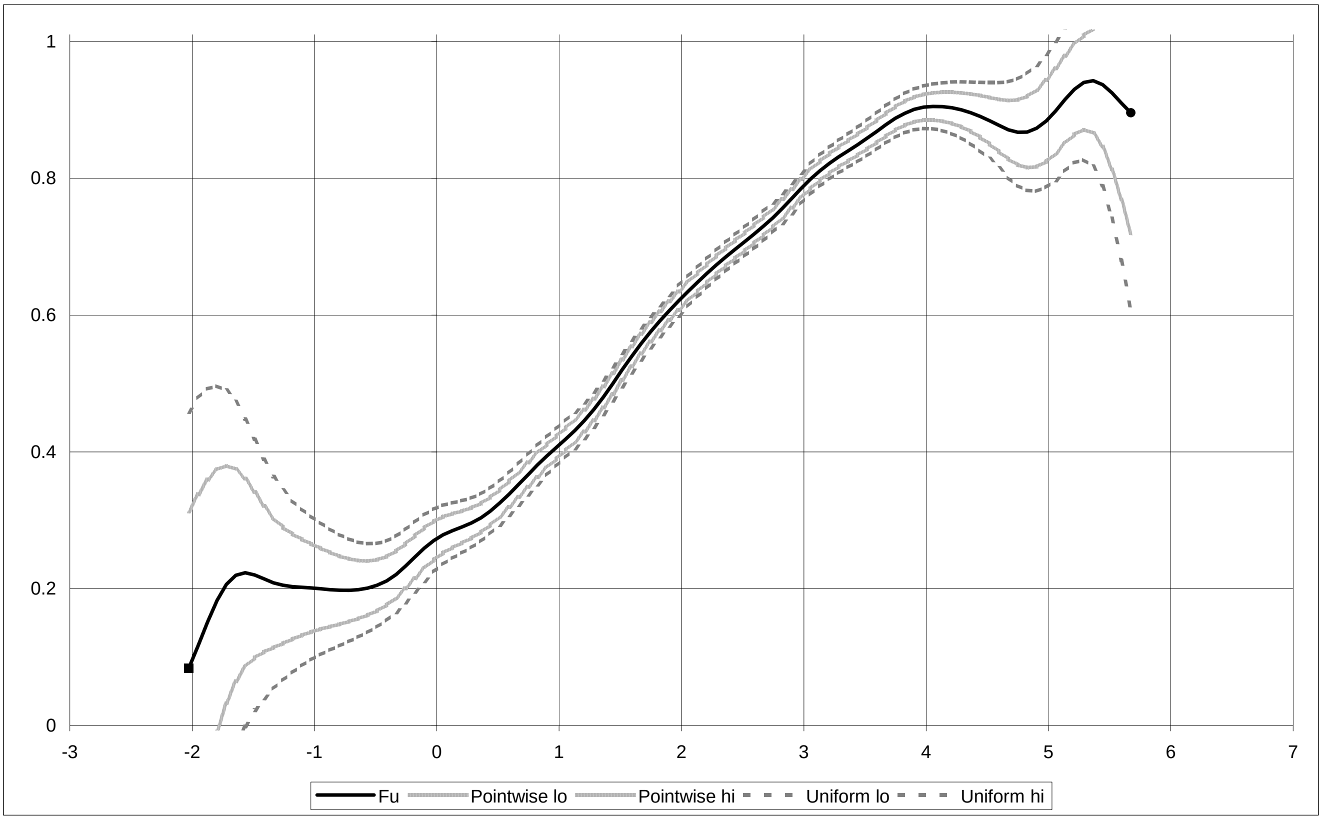

Figure 7 Nonparametric regression of choices y on Klein-Spady residuals

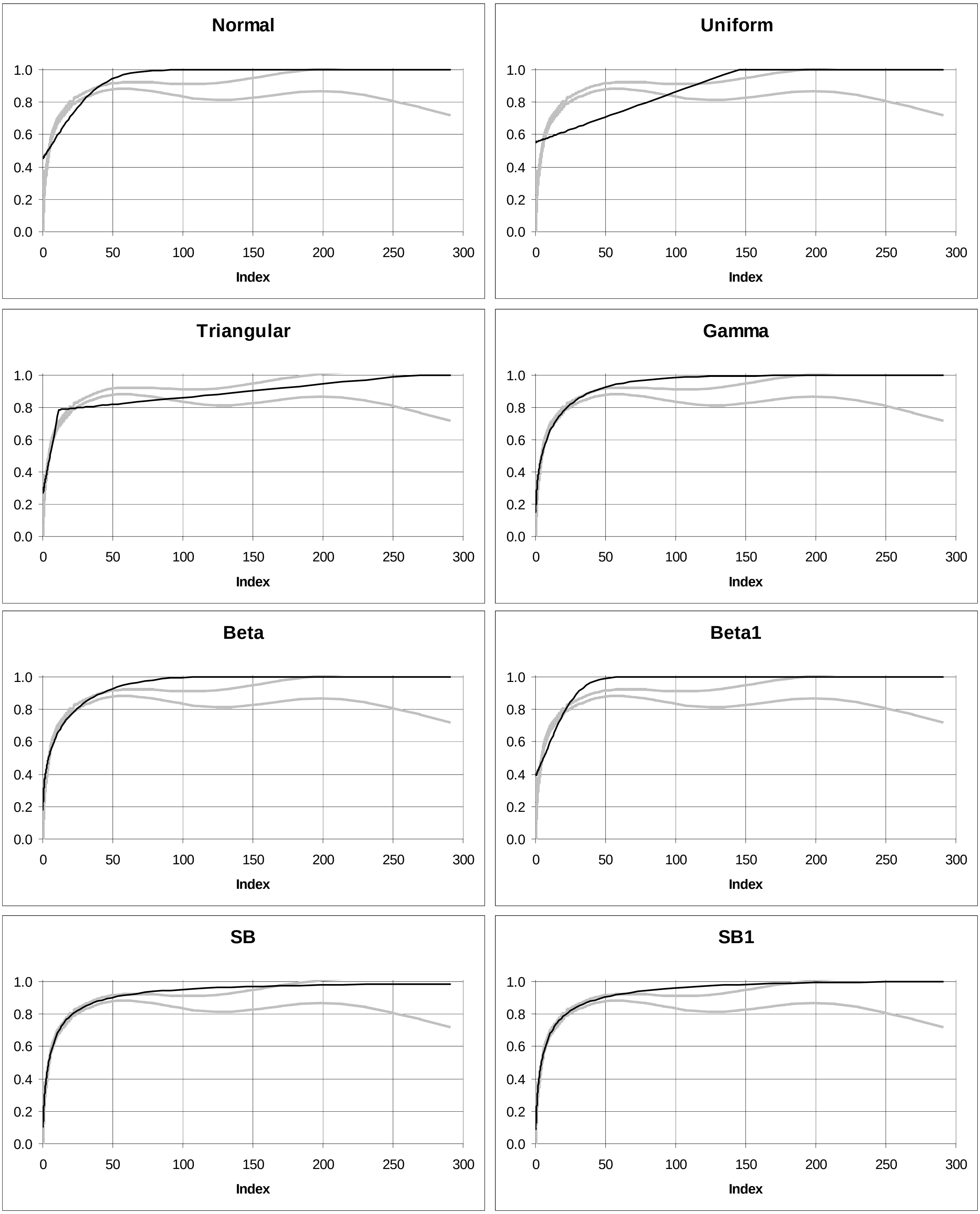

Figure 8 Eight parametric distributions estimated on Klein-Spady residuals. Index = $\exp(v(\hat{\beta}))$

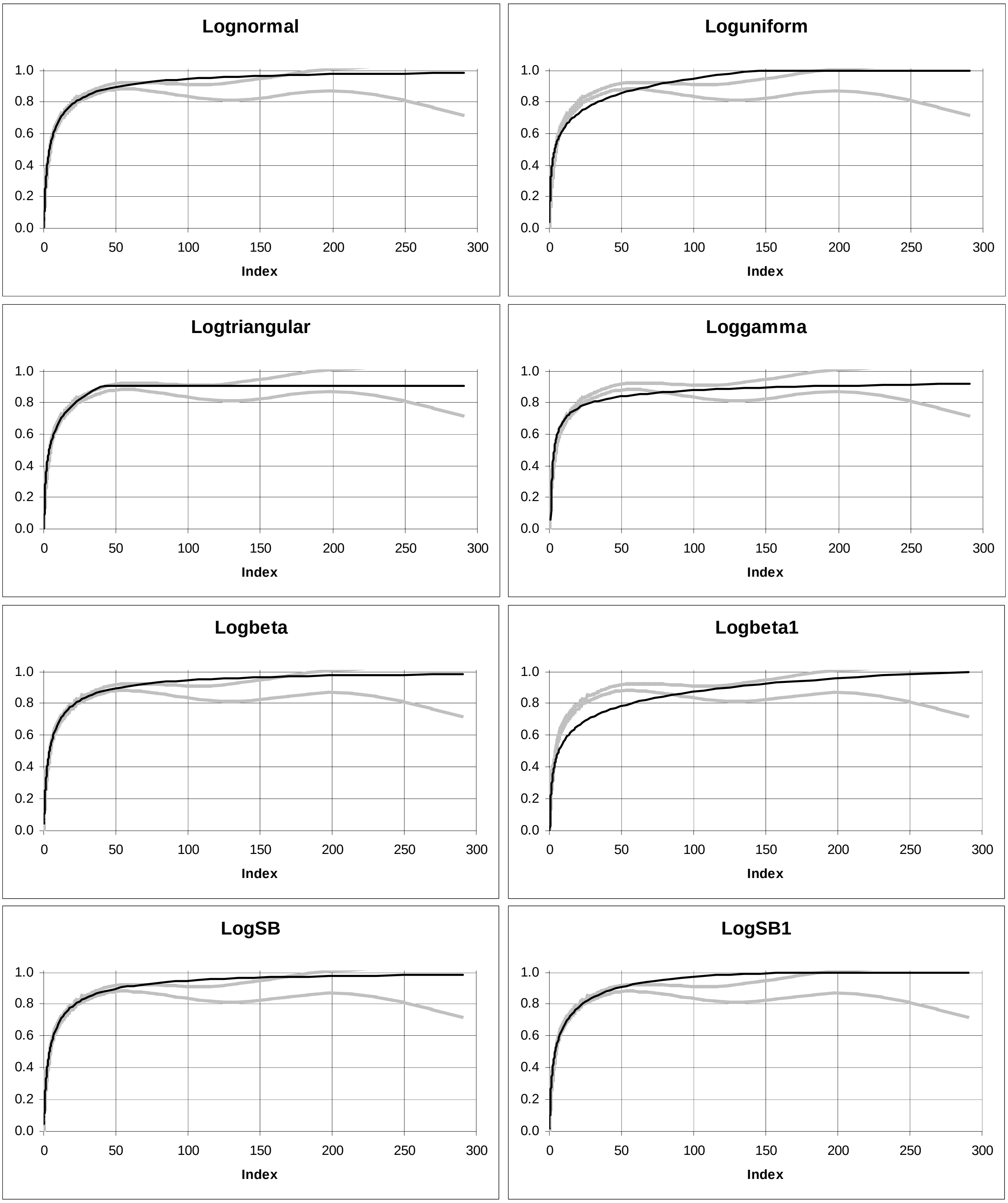

Figure 9 Eight logged parametric distributions estimated on Klein-Spady residuals. Index = $\exp(v(\hat{\beta}))$

Table 1 Specification of parametric distributions

Distribution	Expression
Normal	$F(x) = \Phi\left(\frac{x - p_0}{p_1}\right)$
Gamma	$F(x) = F_{Gamma}(x)$ with $f_{Gamma}(x) = \frac{p_1^{p_0}}{\Gamma(p_0)} x^{p_0 - 1} e^{-p_1 x}$, $x \geq 0$
Uniform	$F(x) = \frac{x - p_0}{p_1 - p_0}$, $p_0 \leq x \leq p_1$
Triangular	$F(x) = \begin{cases} p_3(x - p_0), p_0 \leq x \leq p_1 \\ p_3(p_1 - p_0) + \frac{1 - p_3(p_1 - p_0)}{p_2 - p_1}(x - p_1), p_1 \leq x \leq p_2 \end{cases}$
S_B	$F(x) = \Phi(p_2 + p_3 \log(\frac{x - p_0}{p_1 - x}))$
S_B1	as S_B with $p_1 = \max(x)$
Beta	$F(x) = F_{Beta}\left(\frac{x - p_2}{p_3}\right)$ with $f_{Beta}(z) = \frac{z^{p_0 - 1}(1 - z)^{p_1 - 1}}{B(p_0, p_1)}$, $p_2 \leq x \leq p_2 + p_3$
Beta1	as Beta with $p_3 = \max(x) - p_2$

Table 2. Summary statistics

	Unit	Mean	Std.dev.	Min	Max
y		0.4009	0.4901	0	1
v	DKK/hour	47.98	39.45	3	201

Table 3. Parameter estimates, 16 parametric models estimated on v

	p_0		p_1		p_2		p_3		
Distribution	Estimate	Std.err.	Estimate	Std.err.	Estimate	Std.err.	Estimate	Std.err.	Loglikelihood
Normal	19.61	1.237	102.3	3.037					-10821.7
Lognormal	2.926	0.02733	2.280	0.06230					-10760.6
Beta	0.3237	.	309.0	.	0.1948	.	67634	.	-10748.7
Logbeta	1.999	1.259	1.144	0.2976	-3.054	3.030	8.995	3.212	-10750.4
Gamma	0.3264	0.01165	0.004605	3.180E-4					-10748.7
Loggamma	2.053	0.1063	0.6060	0.03379					-10784.4
Uniform	-190.1	4.992	211.0	2.401					-10907.6
Loguniform	-0.1570	0.09401	5.998	0.06056					-10753.5
Triangular	-56.30	4.134	58.55	2.678	279.3	19.17	0.006291	3.045E-4	-10753.5
Logtriangular	-0.9969	0.4041	3.181	0.2453	5.817	0.07092	0.1250	0.01404	-10746.7
S_B	-14.00	4.564	103980	1460.6	5.288	0.4318	0.6602	0.0652	-10748.6
LogS_B	-1236	.	7.501	0.7162	-9.735	1.560	1.735	0.3324	-10748.8
S_B1	2.635	0.3445			0.7175	0.01718	0.2866	0.009756	-10873.8
LogS_B1	1.010	0.05873			0.1083	0.01623	0.3570	0.01176	-10888
Beta1	0.2090	7.402E-3			0.6119	0.03072	2.978	0.04173	-10876.3
Logbeta1	0.4550	0.04060			0.5160	0.02567	0.9563	0.1543	-10869.9

Table 4. Zheng (1996) test of parametric distributions against nonparametric alternative - regression of y against v

	Not log	Log
Normal	23.75	12.69
Gamma	5.71	21.20
Uniform	62.42	9.30
Triangular	4.57	4.20
S_B	5.13	5.93
S_B1	9.94	11.41
Beta	5.81	6.97
Beta1	8.91	9.49

Table 5. Mean VTTS from parametric distributions estimated on bids v, truncated below at zero

	Not log	Log
Normal	19.6	250.8
Gamma	70.9	3.4E+08
Uniform	55.5	65.2
Triangular	57.7	59.4
S_B	95.9	69.5
S_B1	54.5	54.2
Beta	71.0	61.6
Beta1	53.4	54.1

Table 6. Descriptive statistics for covariates

	Min	Max	Mean
Female dummy	0	1	0.42139
Log of personal income	0	2.3979	1.0362
Income NA dummy	0	1	0.096298
Log of trip duration	1.5041	6.2344	3.4308
Log of time difference	1.0986	4.0943	1.8528
Congestion share	0	0.68	0.090671
Age	16	89	49.791
Age sq./1000	0.256	7.921	2.7023
Commuting dummy	0	1	0.21345
Education dummy	0	1	0.079671

Table 7. Parameter estimates: semiparametric regression (a), lognormal model (b), and Johnson S_B (c)

	(a)			(b)			(c)		
	Variable	Std.err.	t-value	Variable	Std.err.	t-value	Variable	Std.err.	t-value
Female	-0.25157	0.043629	-5.8	-0.2636	0.04343	6.1	-0.25250	0.042871	-5.9
Log of income	0.6845	0.053891	12.7	0.65959	0.05039	13.1	0.64217	0.050112	12.8
Income NA	0.83342	0.086837	9.6	0.7642	0.08835	8.6	0.74895	0.086407	8.7
Log of trip duration	0.16551	0.033092	5.0	0.16378	0.03337	4.9	0.17663	0.033473	5.3
Log of time difference	0.36419	0.035263	10.3	0.35919	0.03552	10.1	0.34754	0.035640	9.8
Congestion share	0.51776	0.13998	3.7	0.49784	0.15147	3.3	0.49907	0.15093	3.3
Age	0.010192	0.0096087	1.1	0.00509	0.00982	0.5	0.004529	0.009592	0.5
Age sq./1000	-0.35745	0.097048	-3.7	-0.3164	0.09938	-3.2	-0.30153	0.096789	-3.1
Commuting	0.18432	0.050364	3.7	0.21296	0.05289	4.0	0.20431	0.052441	3.9
Education	0.144	0.083903	1.7	0.12153	0.08294	1.5	0.11861	0.081552	1.5
τ_1	1.4031	0.022547	62.2	1.5964	0.23971	6.7	-0.40209	0.21379	-1.9
τ_2	1.9934	0.040991	48.6	1.993	0.0518	38.5	2.2169	0.068523	32.4
τ_3							0.54627	0.024033	22.7
Loglikelihood		-10117.0			-10136.4			-10132.7	
Loglikelihood at zero parameters		-10723.3			-10760.6			-10873.0	
Number of parameters		10			10+2			10+3	
Number of observations		17020			17020			17020	

Note: In panel (a) and (b), τ_1 and τ_2 are mean and standard deviation in the normal distribution. In panel (a) these parameters are estimated from the Klein-Spady residuals and can be compared to the corresponding parameters from the lognormal model. In the Johnson S_B model in panel (c) the parameters correspond to p_0, p_2 and p_3 in Table 1.

Table 8. Parameter estimates, 16 parametric models estimated on Klein-Spady residuals

	p_0		p_1		p_2		p_3		
Distribution	Estimate	Std.err.	Estimate	Std.err.	Estimate	Std.err.	Estimate	Std.err.	Loglikelihood
Normal	3.571	0.3392	29.06	0.8048					-10628.8
Lognormal	1.403	0.02256	1.994	0.04100					-10137.4
Beta	0.3292	.	1148	.	0.1957	.	50330	.	-10186.3
Logbeta	604.6	.	614.3	.	-67.65	.	139.2	.	-10137.4
Gamma	0.3530	0.01026	0.02507	0.001341					-10192.6
Loggamma	0.7174	0.02701	0.3490	0.01719					-12830.2
Uniform	-178.4	3.070	144.0	0.4484					-11091.5
Loguniform	-2.319	0.06315	4.975	0.008561					-10263.2
Triangular	-5.873	0.3629	11.49	0.2572	263.8	10.22	0.04528	0.001419	-10217.6
Logtriangular	-1.612	0.01836	3.706	0.02955	29280	.	0.1698	.	-10148.9
S_B	-0.3000	0.1415	1261300	.	6.746	0.2166	0.5378	0.01898	-10133.8
LogS_B	-87.98	128.3	52.93	229.3	-8.999	40.02	16.34	51.76	-10137.3
S_B1	-0.1591	0.1135			2.095	0.05120	0.4994	0.01593	-10143
LogS_B1	-2.932	0.3876			-0.00930	0.09360	0.9834	0.04387	-10164.8
Beta1	874190	.			215.7	.	-1158800	.	-10867.6
Logbeta1	2.536	0.4844			2.331	0.1524	3.304	0.6858	-10155.3

Table 9. Zheng test statistics

	Not log	Log
Normal	156.52	**1.39**
Gamma	16.12	34.60
Uniform	382.07	51.95
Triangular	34.19	9.69
S_B	**0.89**	**1.35**
S_B1	**1.89**	5.39
Beta	15.58	**1.40**
Beta1	107.98	3.94

Table 10. Mean VTTS from parametric distributions estimated on Klein-Spady residuals, truncated below at zero

	Not log	Log
Normal	83.3	**183.6**
Gamma	87.2	2.9E+09
Uniform	198.7	122.8
Triangular	200.8	+inf
S_B	**155.4**	**170.4**
S_B1	**105.1**	97.8
Beta	82.9	**181.4**
Beta1	65.7	100.5