QUANTILE REGRESSION
AN INTRODUCTION

ROGER KOENKER AND KEVIN F. HALLOCK
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

ABSTRACT. Quantile regression as introduced in Koenker and Bassett (1978) may be viewed as a natural extension of classical least squares estimation of conditional mean models to the estimation of an ensemble of models for conditional quantile functions. The central special case is the median regression estimator that minimizes a sum of absolute errors. The remaining conditional quantile functions are estimated by minimizing an asymmetrically weighted sum of absolute errors. Taken together the ensemble of estimated conditional quantile functions offers a much more complete view of the effect of covariates on the location, scale and shape of the distribution of the response variable.

This essay provides a brief tutorial introduction to quantile regression methods, illustrating their application in several settings. Practical aspects of computation, inference, and interpretation are discussed and suggestions for further reading are provided.

1. INTRODUCTION

In the classical mythology of least-squares regression the conditional mean function, the function that describes how the mean of y changes with the vector of covariates x, is (almost) all we need to know about the relationship between y and x. In the resilient terminology of Frisch (1934) and Koopmans (1937) it is "the 'systematic component' or 'true value'," around which y fluctuates due to an "erratic component" or "accidental error." The crucial, *and convenient*, thing about this view is that the error is assumed to have precisely the same distribution whatever values may be taken by the components of the vector x. We refer to this as a pure location shift model since it assumes that x affects only the location of the conditional distribution of y, not its scale, or any other aspect of its distributional shape. If this is the case, we can be fully satisfied with an estimated model of the conditional mean function, supplemented perhaps by an estimate of the conditional dispersion of y around its mean.

Version: December 28, 2000. We would like to thank Gib Bassett, John DiNardo, Olga Geling, and Steve Portnoy for helpful comments on earlier drafts. This paper has been prepared for the *Journal of Economic Perspectives* "Symposium on Econometric Tools". The research was partially supported by NSF grant SBR-9911184.

When we add the further requirement that the errors are Gaussian, least-squares methods deliver the maximum likelihood estimates of the conditional mean function and achieve a well-publicized optimality. Indeed, Gauss seems to have "discovered" the Gaussian density as an *ex post* rationalization for the optimality of least-squares methods. But we will argue that there is more to econometric life than is dreamt of in this philosophy of the location shift model. Covariates may influence the conditional distribution of the response in myriad other ways: expanding its dispersion as in traditional models of heteroscedasticity, stretching one tail of the distribution, compressing the other tail, and even inducing multimodality. Explicit investigation of these effects via quantile regression can provide a more nuanced view of the stochastic relationship between variables, and therefore a more informative empirical analysis.

The remainder of the paper is organized as follows. Section 2 briefly explains how the ordinary quantiles, and consequently the *regression* quantiles, may be defined as the solution to a simple minimization of a weighted sum of absolute residuals. We then illustrate the technique in the classical bivariate setting of Ernst Engel's (1857) original food expenditure study. In Section 3 we sketch the outline of a more ambitious application to the analysis of infant birthweights in the U.S. In Section 4 we offer a brief review of recent empirical applications of quantile regression in economics. Section 5 contains some practical guidance on matters of computation, inference, and software. Section 6 offers several thumbnail sketches of "what can go wrong" and some possible remedies, and Section 7 concludes.

2. What is it?

We say that a student scores at the τth quantile of a standardized exam if he performs better than the proportion τ, of the reference group of students, and worse than the proportion $(1-\tau)$. Thus, half of students perform better than the median student, and half perform worse. Similarly, the quartiles divide the population into four segments with equal proportions of the reference population in each segment. The quintiles divide the population into 5 parts; the deciles into 10 parts. The quantiles, or percentiles, or occasionally fractiles, refer to the general case. Quantile *regression* seeks to extend these ideas to the estimation of conditional quantile *functions* – models in which quantiles of the conditional distribution of the response variable are expressed as functions of observed covariates. To accomplish this task we need a new way to define the quantiles.

Quantiles, and their dual identities, the ranks, seem inseparably linked to the operations of ordering and sorting that are generally used to define them. So it may come as a mild surprise to observe that we can define the quantiles, and the ranks, through a simple alternative expedient – as an optimization problem. Just as we can define the sample mean as the solution to the problem of minimizing a sum of squared residuals, we can define the median as the solution to the problem of minimizing a sum of absolute residuals. What about the other quantiles? If the symmetric absolute value function yields the median, maybe we can simply tilt the absolute value to produce the other quantiles. This "pinball logic" suggests solving

$$\min_{\xi \in \Re} \sum \rho_\tau(y_i - \xi) \tag{2.1}$$

where the function $\rho_\tau(\cdot)$ is illustrated in Figure 2.1. To see that this problem yields the sample quantiles as its solutions, it is only necessary to compute the directional derivative of the objective function with respect to ξ, taken from the left and from the right.[1]

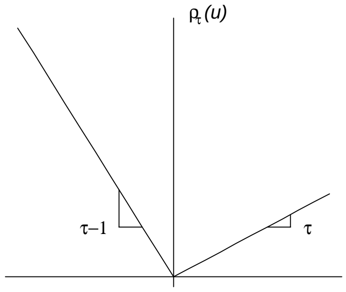

Figure 2.1. Quantile Regression ρ Function

Having succeeded in defining the unconditional quantiles as an optimization problem, it is easy to define conditional quantiles in an analogous fashion. Least squares regression offers a model for how to proceed. If, presented with a random sample $\{y_1, y_2, \dots, y_n\}$, we solve

$$\min_{\mu \in \Re} \sum_{i=1}^{n} (y_i - \mu)^2, \tag{2.2}$$

[1]Consider the median case with τ equal to $\frac{1}{2}$: the directional derivatives are simply one-half the sum of the signs of the residuals, $y_i - \xi$, with zero residuals counted as $+1$ from the right and as -1 from the left. If $\hat{\xi}$ is taken to be a value such that half the observations lie above $\hat{\xi}$ and half lie below, then both directional derivatives will be positive, and since the objective function is increasing in either direction $\hat{\xi}$ must be a local minimum. But the objective function is a sum of convex functions and hence convex, so the local minimum is also a global one. Such a solution is a median by our original definition. The same argument works for the other quantiles, but now we have asymmetric weighting of the number of observations with positive and negative residuals and this leads to solutions $\hat{\xi}(\tau)$ corresponding to the τth quantiles.

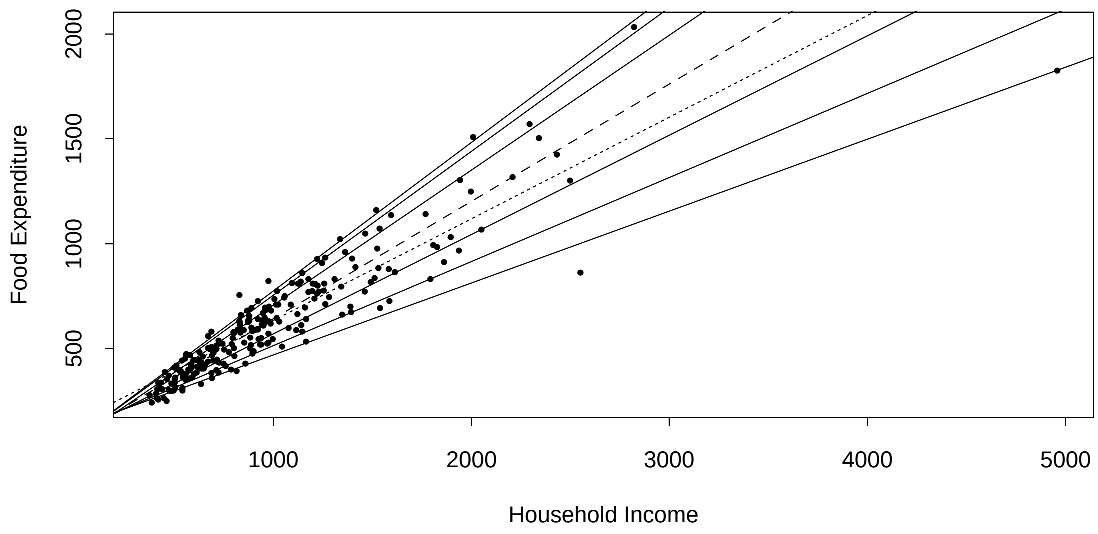

FIGURE 2.2. Engel Curves for Food: This figure plots data taken from Ernst Engel's (1857) study of the dependence of households' food expenditure on household income.

we obtain the sample mean, an estimate of the unconditional population mean, EY. If we now replace the scalar μ by a parametric function $\mu(x, \beta)$ and solve

$$\min_{\beta \in \Re^p} \sum_{i=1}^{n} (y_i - \mu(x_i, \beta))^2 \tag{2.3}$$

we obtain an estimate of the conditional expectation *function* $E(Y|x)$.

In quantile regression we proceed in exactly the same way. To obtain an estimate of the conditional median function, we simply replace the scalar ξ in (2.1) by the parametric function $\xi(x_i, \beta)$ and set τ to $\frac{1}{2}$. Variants of this idea were proposed in the mid 18th century by Boscovich, and subsequently investigated by Laplace and Edgeworth, among others. To obtain estimates of the other conditional quantile functions we simply replace absolute values by $\rho_\tau(\cdot)$, and solve

$$\min_{\beta \in \Re^p} \sum \rho_\tau(y_i - \xi(x_i, \beta)) \tag{2.4}$$

The resulting minimization problem, when $\xi(x, \hat{\beta}(\tau))$ is formulated as a linear function of parameters, can be solved very efficiently by linear programming methods. We will defer further discussion of computational aspects of quantile regression to Section 5.

To illustrate the basic ideas we briefly reconsider a classical empirical application, Ernst Engel's (1857) analysis of the relationship between household food expenditure and household income. In Figure 2.2 we plot the data taken from 235 European working class households. Superimposed on the plot are seven estimated quantile regression lines corresponding to the quantiles $\tau \in \{.05, .1, .25, .5, .75, .9, .95\}$. The median $\tau = .5$ fit is indicated by the dashed line; the least squares fit is plotted as the dotted line.

The plot clearly reveals the tendency of the *dispersion* of food expenditure to increase along with its level as household income increases. The spacing of the quantile regression lines also reveals that the conditional distribution of food expenditure is skewed to the left: the narrower spacing of the upper quantiles indicating high density and a short upper tail and the wider spacing of the lower quantiles indicating a lower density and longer lower tail.

The conditional median and mean fits are quite different in this example, a fact that is partially explained by the asymmetry of the conditional density and partially by the strong effect exerted on the least squares fit by the two unusual points with high income and low food expenditure. Note that one consequence of this nonrobustness is that the least squares fit provides a rather poor estimate of the conditional mean for the poorest households in the sample.

We have occasionally encountered the faulty notion that something like quantile regression could be achieved by segmenting the response variable into subsets according to its unconditional distribution and then doing least squares fitting on these subsets. Clearly, this form of "truncation on the dependent variable" would yield disastrous results in the present example. In general, such strategies are doomed to failure for all the reasons so carefully laid out in Heckman (1979). It is thus worth emphasizing that even for the extreme quantiles *all* the sample observations are actively in play in the process of quantile regression fitting. Each fit depicted in Figure 2.2 is ultimately determined by only a pair of sample points, but *all* n points are needed to determine which pair of points are selected. With p parameters to be estimated, p points determine the fit, but *which* p points depend on the entire sample.

In contrast, segmenting the sample into subsets defined according to the conditioning covariates is always a valid option. Indeed such local fitting underlies all non-parametric quantile regression approaches. In the most extreme cases we have p distinct cells corresponding to different settings of the covariate vector, x, and quantile regression reduces to simply computing ordinary univariate quantiles for each of these cells. In intermediate cases we may wish to project these cell estimates onto a more parsimonious (linear) model. See *e.g.* Chamberlain (1994) and Knight, Bassett, and Tam (2000).

Another variant, one that actually has some merit, is the suggestion that instead of estimating linear conditional quantile models, we could instead estimate a family of binary response models for the probability that the response variable exceeded some prespecified cutoff values. This approach replaces the hypothesis of conditional quantile functions that are linear in parameters with the hypothesis that some transformation of the various probabilities of exceeding the chosen cutoffs, say the logistic, could instead be expressed as linear functions in the observed covariates. In our view the conditional quantile assumption is more natural, if only because it nests within it the iid error location shift model of classical linear regression.

3. "Let's Do It"

In this section we reconsider a recent investigation by Abreveya (2001) of the impact of various demographic characteristics and maternal behavior on the birthweight of infants born in the U.S. Low birthweight is known to be associated with a wide range of subsequent health problems, and has even been linked to educational attainment and labor market outcomes. Consequently, there has been considerable interest in factors influencing birthweights, and public policy initiatives that might prove effective in reducing the incidence of low birthweight infants.

Although most of the analysis of birthweights has employed conventional least squares regression methods it has been recognized that the resulting estimates of various effects on the conditional mean of birthweights were not necessarily indicative of the size and nature of these effects on the lower tail of the birthweight distribution. In an effort to focus attention more directly on the lower tail, several studies have recently explored binary response, *e.g.* probit, models for the occurrence of low birthweights – conventionally defined to be infants weighing less than 2500 grams, or about 5 pounds 9 ounces. Quantile regression offers a natural complement to these prior modes of analysis. A more complete picture of covariate effects can be provided by estimating a family of conditional quantile functions, as we will now illustrate.

Our analysis is based on the June, 1997, Detailed Natality Data published by the National Center for Health Statistics. Like Abreveya (2001), we limit the sample to live, singleton births, with mothers recorded as either black or white, between the ages of 18 and 45, residing in the U.S. Observations with missing data for any of the variables described below were dropped from the analysis. This process yielded a sample of 198,377 babies. Birthweight, the response variable, is recorded in grams. Education of the mother is divided into four categories: less than high school, high school, some college, and college graduate. The omitted category is less than high school so coefficients may be interpreted relative to this category. The prenatal medical care of the mother is also divided into 4 categories: those with no prenatal visit, those whose first prenatal visit was in the first trimester of the pregnancy, those with first visit in the second trimester, and those with first visit in the last trimester. The omitted category is the group with a first visit in the first trimester; they constitute almost 85 percent of the sample. An indicator of whether the mother smoked during pregnancy is included in the model, as well as mother's reported average number of cigarettes smoked per day. The mother's reported weight gain during pregnancy (in pounds) is included as a quadratic effect.

In Figure 3.1 we present a concise visual summary of the quantile regression results for this example. Each plot depicts one of the 16 coefficient in the quantile regression model. The solid line with filled dots represents the 19 point estimates of the coefficient for τ's ranging from 0.05 to 0.95. The shaded grey area depicts a 90 percent pointwise confidence band. Superimposed on the plot is a dashed line representing the ordinary least squares estimate of the mean effect, with two dotted lines representing again a 90 percent confidence interval for this coefficient. Note that with the exception of the "high-school" and "some college" coefficients, the quantile regression estimates lie outside mean regresion confidence interval indicating that the location shift interpretation of the covariate "effect" is implausible.

In the first panel of the figure the intercept of the model may be interpreted as the estimated conditional quantile function of the birthweight distribution of a girl born to an unmarried, white mother with less than a high school education, who is 27 years old and had a weight gain of 30 pounds, didn't smoke, and had her first prenatal visit in the first trimester of the pregnancy. The mother's age and weight gain are chosen to reflect the means of these variables in the sample.[2] Note that the $\tau = .05$ quantile of the distribution for this group is just at the margin of the conventional definition of a low birthweight baby.

At any chosen quantile we can ask how different are the corresponding weights of boys and girls, given a specification of the other conditioning variables. The second panel answers this question. Boys are obviously larger than girls, by about 100 grams according to the OLS estimates of the mean effect, but as is clear from the quantile regression results the disparity is much smaller in the lower quantiles of the distribution and somewhat larger than 100 grams in the upper tail of the distribution.

Perhaps surprisingly, the marital status of the mother seems to be associated with a rather large positive effect on birthweight especially in the lower tail of the distribution. The public health implications of this finding should, of course, be viewed with caution.

The disparity between birthweights of infants born to black and white mothers is disturbing particularly at the left tail of the distribution. At the 5th percentile of the conditional distribution the difference is roughly one third of a kilogram.

Mother's age enters the model as a quadratic. At the lower quantiles the mother's age tends to be more concave, increasing birthweight from age 18 to about age 30, but tending to decrease birthweight when the mother's age is beyond 30. At higher quantiles this optimal age becomes gradually older. At the third quantile it is about 36, and at $\tau = .9$ it is almost 40.

[2]It is conducive for interpretation to center covariates so that the intercept can be interpreted as the conditional quantile function for some representative case – rather than as an extrapolation of the model beyond the convex hull of the data. This may be viewed as adhering to John Tukey's admonition: "Never estimate intercepts, always estimate *centercepts*!"

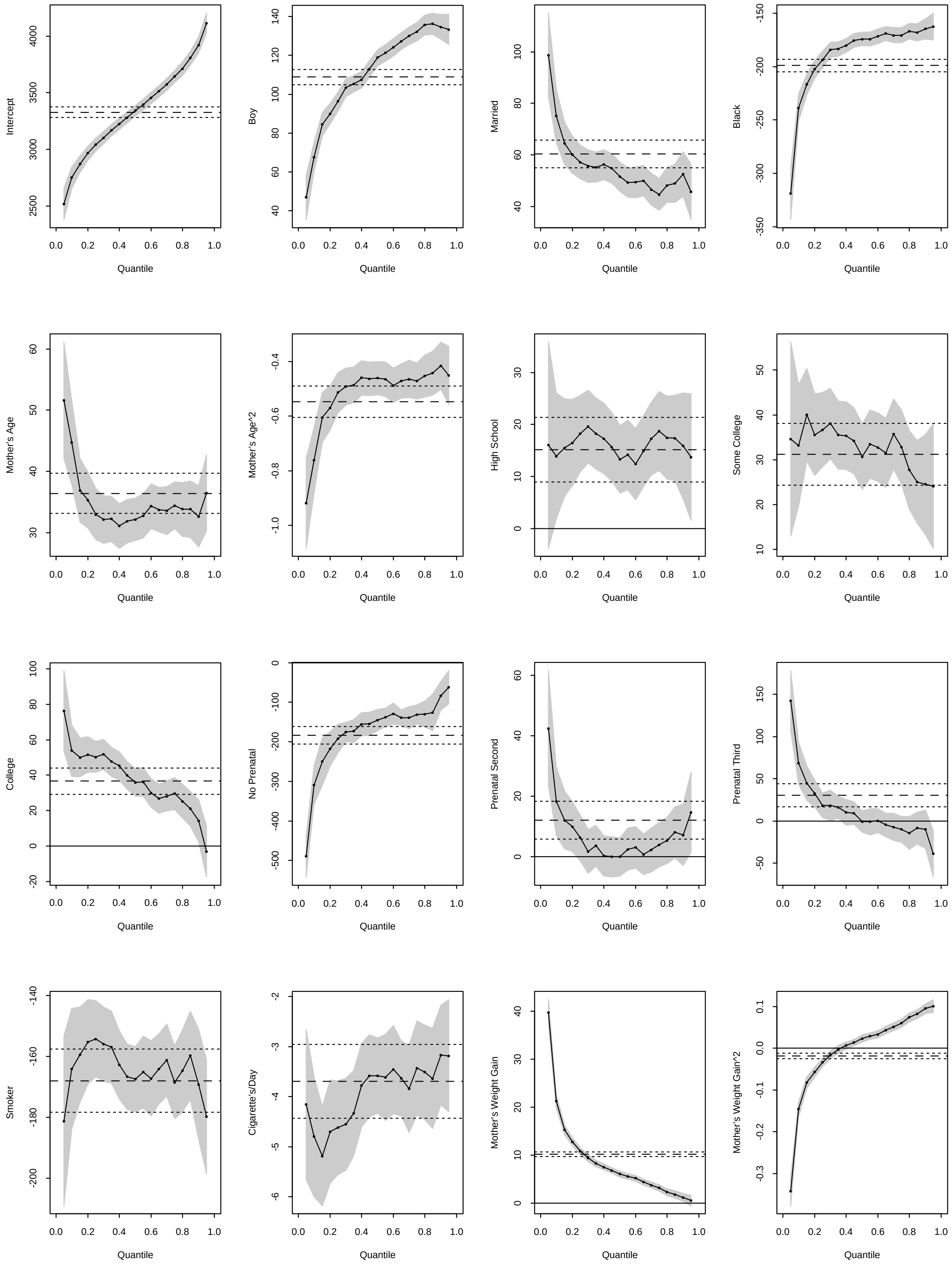

FIGURE 3.1. Quantile Regression for Birthweights

Education beyond high school is associated with a modest increase in birthweights. High school graduation has a quite uniform effect over the whole range of the distribution of about 15 grams. This is a rare example of an effect that really does appear to exert a pure location shift effect on the conditional distribution. Some college education has a somewhat more positive effect in the lower tail than in the upper tail, varying from about 35 grams in the lower tail to 25 grams in the upper tail. A college degree has an even more substantial positive effect, but again much larger in the lower tail and declining to a negligible effect in the upper tail.

The effect of prenatal care is of obvious policy interest. Since individuals self-select into prenatal care results must be interpreted with considerable caution. Those receiving no prenatal care are likely to be at risk in other dimensions as well. Nevertheless, the effects are sufficiently large to warrant considerable further investigation. Babies born to mothers who received no prenatal care were on average about 150 grams lighter than those who had a prenatal visit in the first trimester. In the lower tail of the distribution this effect is considerably larger – at the 5th percentile it is nearly half a kilogram! In contrast, mothers who delayed prenatal visits until the second or third trimester have substantially *higher* birthweights in the lower tail than mothers who had a visit in the first trimester. This might be interpreted as the self-selection effect of mothers confident about favorable outcomes. In the upper 3/4 of the distribution there seems to be no significant effect.

Smoking has a clearly deleterious effect. The indicator of whether the mother smoked during the pregnancy is associated with a decrease of about 175 grams in birthweight. In addition, there is an effect of about 4 to 5 grams per cigarette per day. Thus a mother smoking a pack per day appears to induce a birthweight reduction of about 250 to 300 grams, or from about half to two-thirds of a pound.

Lest this smoking effect be thought to be attributable to some associated reduction in the mothers weight gain, we should hasten to point out that the weight gain effect is explicitly accounted for with a quadratic specification. Not suprisingly, the mother's weight gain has a very strong influence on birthweight, and this is reflected in the very narrow confidence band for both linear and quadratic coefficients. At low weight gains by the mother the marginal effect of another pound gained is about 30 grams at the lowest quantiles and declines to only about 5 grams at the upper quantiles. This pattern of declining marginal effects is maintained for large weight gains until we begin to consider extremely large weight gains at which point the effect is reversed. The quadratic specification of the effect of mother's weight gain offers a striking example of how misleading the OLS estimates can be. Note that the OLS estimates strongly suggest that the effect is linear with an essentially negligible quadratic effect. The quantile regression estimates give a very different picture, one in which the quadratic effect of the weight gain is very significant except where it crosses the zero axis at about $\tau = .33$.

4. Who's Doing It?

In Spain, the best upper sets do it...
Some Argentines, without means, do it
People say in Boston even beans do it
Cole Porter (1928)

There is a rapidly expanding *empirical* quantile regression literature in economics which, taken as a whole, makes a persuasive case for the value of "going beyond models for the conditional mean" in empirical economics. Catalysed by Gary Chamberlain's invited address to the 1990 World Congress of the Econometric Society, Chamberlain (1994), there has been considerable work in labor economics: on union wage effects, returns to education, and labor market discrimination. Chamberlain finds, for example, that for manufacturing workers, the union wage premium at the first decile is 28 percent and declines monotonically to a negligible 0.3 percent at the upper decile. The least squares estimate of the mean union premium of 15.8 percent is thus captured mainly by the lower tail of the conditional distribution. The conventional location shift model thus delivers a rather misleading impression of the union effect. Other contributions exploring these issues in the U.S. labor market include the influential work of Buchinsky (1994, 1995, 1997, 1998, 2001) . Arias, Hallock, and Sosa (2001) using data on identical twins interpret observed heterogeneity in the estimated return to education over quantiles as indicative of an interaction between observed educational attainment and unobserved ability.

There is also a large literature dealing with related issues in labor markets outside the U.S. including Fitzenberger and Kurz (1997), Büttner and Fitzenberger (1998), Fitzenberger (1999), and Fitzenberger, Hujer, MaCurdy, and Schnabel (2001), on Germany; Machado and Mata (2001) on Portugal; Abadie (1997) and Lopez, Hernandez, and Garcia (2001) on Spain; Schultz and Mwabu (1998) on South Africa; Montenegro (1998) on Chile; and Kahn (1998) on international comparisons. The work of Machado and Mata (2001) is particularly notable since it introduces a useful way to extend the counterfactual decomposition approach of Oaxaca to quantile regression and suggests a general strategy for simulating marginal distributions from the quantile regression process. Tannuri (2000) has employed this approach in a study of assimilation of U.S. immigrants.

In other applied micro areas Eide and Showalter (1998), Knight, Bassett, and Tam (2000), and Levin (2001) have addressed school quality issues. Levin's study of a panel survey of the performance of Dutch school children finds little support for the claim that reducing class size improves student outcomes, but he does find some evidence of positive peer effects particularly in the lower tail of the achievement distribution. Poterba and Rueben (1995) and Mueller (2000) study public-private wage differentials in the U.S. and Canada. Work by Viscusi and Born (1995) considers liability reform effects on medical malpractice. Viscusi and Hamilton (1999) consider public decision making on hazardous waste cleanup.

Deaton (1997) offers a nice introduction to quantile regression for demand analysis. In a study of Engel curves for food expenditure in Pakistan he finds that although the median engel elasticity of 0.906 is similar to the OLS estimate of 0.909, the coefficient at the 10th quantile is 0.879 and the estimate at the 90th percentile is 0.946. In another demand application, Manning, Blumberg, and Moulton (1995) study demand for alcohol using survey data from the National Health Interview Study and find considerable heterogeneity in the price and income elasticities over full range of the conditional distribution. Hendricks and Koenker (1991) investigate demand for electricity by time of day.

Earnings inequality and mobility is a natural arena of applications for quantile regression. Conley and Galenson (1998) explore wealth accumulation in several U.S. cities in the mid-19th century. Gosling, Machin, and Meghir (1996) study the income and wealth distribution in the UK using 27 years of the UK Family Expenditure Survey. Trede (1998) and Morillo (2000) compare earnings mobility in the U.S. and Germany.

There is also a growing literature in empirical finance employing quantile regression methods. One strand of this literature is the rapidly mushrooming literature on value at risk: this connection is developed in Taylor (1999), Chernozhukov and Umantsev (2001), and Engle and Manganelli (1999). Bassett and Chen (2001) consider quantile regression index models to characterize mutual fund investment styles.

5. How to do it

The diffusion of technological change throughout statistics is closely tied to its embodiment in statistical software. This is particularly true of quantile regression methods since the linear programming algorithms that underlie reliable implementations of the methods appear somewhat esoteric to some users. This section offers a critical review of existing algorithms and inference strategies for quantile regression. Readers interested in the practicalities of current software may prefer to skip immediately to the final subsection where this is discussed.

5.1. **Algorithms.** Since the early 1950's it has been recognized that median regression methods based on minimizing sums of absolute residuals can be formulated as linear programming problems and efficiently solved with some form of the simplex algorithm. For this purpose, the median regression algorithm of Barrodale and Roberts (1974) has proven particularly influential. Median regression algorithms can be easily adapted for general quantile regression problems. Koenker and d'Orey (1987, 1993) describe one implementation.

The Barrodale and Roberts approach typifies the class of exterior point algorithms for solving linear programming problems: we travel from vertex to vertex along the edges of the polyhedral constraint set, chosing at each vertex the path of steepest descent, until we arrive at the optimum. The work of Karmarker (1984) initiated a dramatic reappraisal of computational methods for linear programming. Instead, of traversing the outer surface, we take Newton steps from the interior of a deformed version of the constraint set toward the boundary. This approach has produced extremely effective interior point algorithms that are closely related to the log barrier methods pioneered by Frisch in the 1950's for solving constrained optimization problems. These methods are particularly effective for large scale quantile regression problems. For such problems, Portnoy and Koenker (1997) have shown that a combination of interior point methods and effective problem preprocessing render large scale quantile regression computation competitive with least squares computations for problems of comparable size.

Sample Size	Barrodale-Roberts	Frisch-Newton	Preprocessing
100	0.03	0.04	0.05
1000	0.57	0.14	0.47
10000	17.96	1.49	1.61
100000	1317.24	24.59	11.69

TABLE 5.1. Timings for three alternative quantile regression algorithms. Timings are reported in cpu seconds using Splus's `unix.time` command on a Sun Sparc Ultra II. The data for each row of the table consists of the first n rows of the natality data described above. The model has been simplified to include only 9 parameters rather than the original 16 to avoid some obvious singularity problems that arise at small sample sizes.

To illustrate the performance of the three algorithms we provide some timings based on a restricted version of our natality example. From the results reported in Table 5.1, it is clear that the exterior point algorithm of BR is competitive only for relatively small samples. The pure interior point algorithm that we characterize as the Frisch- Newton method is considerably faster for large problems, nearly 60 times faster when n is 100,000. Preprocessing can yield a further significant improvement, but only for quite large problems. The lesson we would draw from this experience is that the traditional reliance on simplex methods in most implementations of quantile regression in commercial software is well founded only for rather small problems. We hope that reemphasizing this point, which is developed in considerably more detail in Portnoy and Koenker (1997), will encourage software developers to consider implementing more efficient interior point methods as an option for larger problems.

It may also be worth emphasizing that the proposed preprocessing strategy also provides, in principle, a way to design algorithms for problems so large that data cannot be accommodated in machine memory. In effect, preprocessing substitutes solution of several smaller $\mathcal{O}(n^{2/3})$ problems for solving one large $\mathcal{O}(n)$ problem.[3]

5.2. **What should we put in parentheses?** It is a basic principle of sound econometrics that *every serious estimate deserves a standard error*. There are two general approaches to filling parentheses. Either we offer a procedure to estimate the asymptotic standard error of the estimator, or we suggest some form of the bootstrap. We will briefly consider both approaches.

5.3. **Asymptotics.** In its most elementary form the asymptotic theory of quantile regression is provided in Koenker and Bassett (1978). In fact, the finite sample distribution is also explicitly given there, although the combinatorial form of the exact density is unlikely to prove practical for several more cycles of Moore's law. The elementary asymptotics of quantile regression are based on the assumption of an iid error, pure location-shift version of the model. In this case the limiting behavior of $\hat{\beta}_n(\tau)$ is normal with covariance matrix $\omega^2(\tau)(X'X)^{-1}$, where $\omega^2(\tau)$ denotes the quantity $\tau(1-\tau)/f^2(F^{-1}(\tau))$ and $f(F^{-1}(\tau))$ denotes the density of the error distribution evaluated at the τ^{th} quantile. This corresponds closely to the behavior of the ordinary least squares estimator under the same conditions except that ω^2 is replaced by σ^2, the variance of the underlying error distribution.

Why does $f^2(F^{-1}(\tau))$ appear in the asymptotic covariance matrix of $\hat{\beta}(\tau)$? A heuristic explanation may be useful. Estimation of the τth conditional quantile function relies, at least asymptotically, only on the observations near the τth quantile, but the number of such observations is proportional to $f(F^{-1}(\tau))$ in the iid case, and the variability of the estimate, by δ method considerations, decreases like the squared reciprocal. It is easy to estimate the nuisance parameter $\omega^2(\tau)$ based on the quantile regression residuals, but as we have already seen in the examples the iid error assumptions seems highly questionable in many application settings.

When the observations are independent but not identically distributed, as we would expect in most microeconometric applications it is quite straightforward to extend the iid theory to produce a version of the Huber-Eicker-White sandwich formula for the limiting covariance matrix of $\hat{\beta}(\tau)$. Estimation is somewhat more complicated, but there are several reasonable candidate estimators that have been proposed in the literature. In addition inversion of a rank test as described in Koenker (1994,1996) offers a reliable method of constructing confidence intervals in the non-iid error context. These intervals are constructed by inverting a univariate rank test introduced in Gutenbrunner, Jurečková , Koenker, and Portnoy (1993) to find a set of hypothetical values of the parameter that would not lead to rejection at the proscribed level. The test, in turn, is based on a fundamental link between the formal linear programming dual of the quantile regression optimization problem and the theory of rank statistics,

[3]Thus, for example for a problem with a million observations we need only solve problems with n roughly 10,000 that as Table 5.1 indicates are very efficiently handled by the Frisch-Newton method.

introduced ie Gutenbrunner and Jurečková (1992). We compare several methods of confidence interval construction based on asymptotic theory below.

5.4. **Bootstrap.** Efron's ubiquitous bootstrap offers an expanded range of options for computing confidence intervals and standard errors. Indeed, Efron himself proposed using the bootstrap to estimate confidence intervals for median regression as early as his 1987 Ferber lecture at the University of Illinois. Since that time, there has been considerable further consideration of the bootstrap for quantile regression from both theoretical and applied perspectives.

In the simplest independent, but not identically distributed settings the standard approach is the so-called (x, y)-pair bootstrap. Pairs (x_i, y_i) $i = 1, \ldots, n$ are drawn at random from the original observations *with replacement.* For each resampling the estimator $\hat{\beta}_n^*(\tau)$ is recomputed. Repeating this procedure B times yields a sample of B p-vectors whose sample covariance matrix constitutes a valid estimator of the covariance matrix of the original estimator. This procedure is automated in Stata's `bsqreg` command, for example. Alternatively, and in our view preferably, confidence regions for the quantile regression parameters can be computed from the empirical distribution of the sample of bootstrapped $\hat{\beta}_n^*(\tau)$'s, the so-called percentile method. These procedures are easily extended to deal with the joint distribution of several distinct quantile regression estimators $\{\hat{\beta}_n(\tau_k)\ \ k = 1, \ldots, K\}$, as would be needed to test equality of slope parameters across quantiles, for example. In practice it is often desirable, particularly when the initial sample size is large to consider subsampling, or what is sometimes called the m out of n bootstrap. In this case the sample size of each bootstrap sample is $m \ll n$, and the resulting bootstrapped covariance matrix is rescaled by $\sqrt{m/n}$. This approach has been used extensively by Buchinsky (1994) and also by Abreveya (2001) in large scale applications.[4]

Parzen, Wei, and Ying (1994) have suggested that rather than bootstrapping (x_i, y_i) pairs, one can instead bootstrap the quantile regression gradient condition. This involves sampling Bernoulli random variables, leads to an asymptotically pivotal approach, and can be implemented very efficiently. More recently, Bilias, Chen, and Ying (2000) have suggested a formulation of this approach as a resampling strategy for censored quantile regression.

5.5. **Comparison of Methods.** We have occasionally encountered the view that because there are several available approaches to inference for quantile regression, or because it involves a form of density estimation, it is a subject too arcane to offer

[4]Recently, Bickel and Sakov (2000) have shown that choice of m such that m/n tends to zero for bootstrapping the sample median achieves a higher-order asymptotic refinement. Extension of this finding to quantile regression is an intriguing open problem. Choosing $m/n \to 0$ may be viewed as a smoothing device. Alternative smoothing schemes, e.g. Horowitz (1998), have been proposed, but they have the disadvantages of being computationally more demanding and introducing additional tuning constants.

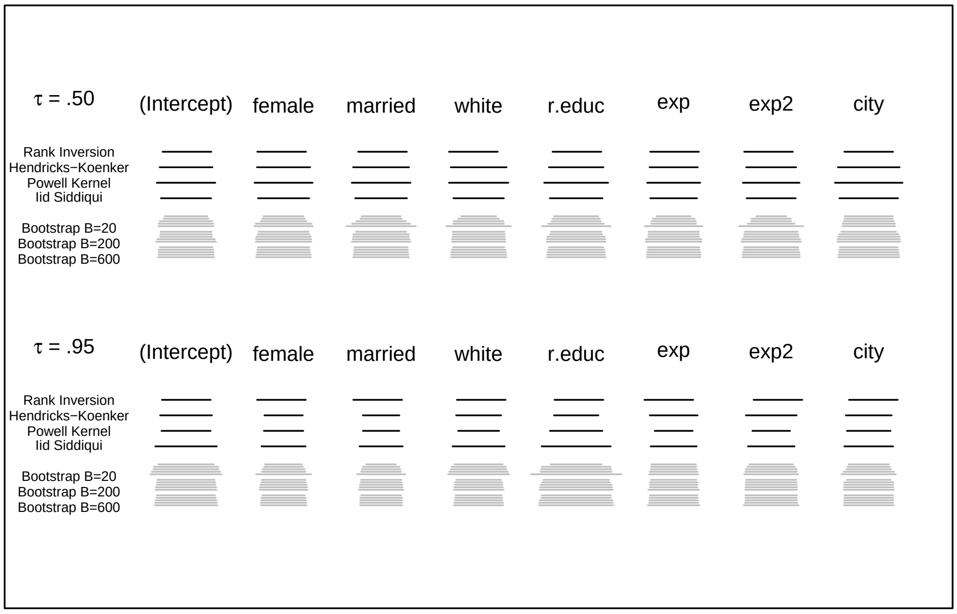

Figure 5.1. Comparison of Confidence Interval Lengths: The figure plots relative confidence interval lengths for several different proposed methods. The rank-inversion method is normalized to have unit length, so all other intervals are relative to the rank inversion interval. The bootstrap intervals are represented by 5 independent realizations that are plotted separately.

reliable conclusions.[5] In our experience the discrepancies between reasonable competing methods are slight, and inference for quantile regression is, if anything, more robust than many other forms of inference commonly encountered in econometrics.

To illustrate this point, we compare in Figure 5.1 several competing methods of estimating confidence intervals for quantile regression parameters in a standard log wage

[5]Of course, with the advent of the Eicker-White-Newey-West proliferation of standard error proposals for the ordinary least squares estimator one might be led to similar conclusions for mean regression.

equation model based on 39,466 observations from the 1999 March Current Population Survey. The specification includes a linear schooling term, potential experience, and its square, and indicators for gender, marital status, race, and residence in an urban area. We compare 5 methods: the rank inversion intervals Koenker (1994), the sandwich formula estimate suggested in Hasan and Koenker (1997), the sandwich formula proposal of Powell (1989), the naive Siddiqui estimator of the asymptotic covariance matrix under the assumption of iid errors, and three versions of the bootstrap using B=20, 200, and 600. For each of the bootstrap methods we plot five realizations of each bootstrapped confidence interval to provide some indication of the variability of the method. In each case we have normalized the length of the rank inversion interval to be one. Note that the rank inversion interval is allowed to asymmetric, while all the other intervals are symmetric by construction. Of course, in such a large samples, we would expect that the estimator would be nearly normal and hence the intervals would be approximately symmetric.

Subject to the obvious proviso that one should be highly suspicious of any *one* exercise like this, we boldly draw the following conclusions. None of the methods display embarrassing performance. Even the iid Siddiqui approach that we expected to yield overly optimistic, short intervals isn't *that* bad. Powell's approach looks somewhat more pessimistic at the median, but seems somewhat more optimistic at .95. The bootstrap intervals are obviously too variable with only 20 replications, and perhaps also at 200. The more stable behavior of the bootstrap at 600 replications is quite consistent with the guidelines provided by Andrews and Buchinsky (2000).

Stata's command `qreg` also produces estimates of asymptotic standard errors based on iid error assumptions. Although they are designated as "Koenker-Bassett standard errors" the method bears little resemblance to the histospline approach of the cited reference. As described by Rogers (1993) the `qreg`'s standard errors appear to be a variant of the iid Siddiqui method with a rather unfortunate choice of bandwidth.[6] A consequence of the undersmoothing implied by the Stata rule is that the resulting standard errors are frequently considerably smaller than would be obtained with a more conventional bandwidth selection rule. This conclusion is supported by the Monte Carlo comparison reported in Rogers (1992).

5.6. **Software Packages.** Among commercial programs in common use in econometrics only Stata offers some basic functionality for quantile regression within the central core of the package distributed by the vendor. Since the mid-1980's one of us has maintained a public domain package of quantile regression software designed for the S language of Becker, Chambers, and Wilks (1988), and the related commercial package Splus. Recently, this package has been extended to include a version for

[6]Stata's bandwidth rule appears to be $h = [n^{-1/2}]$. In contrast the default Hall and Sheather (1988) rule employed in our S implementation is $h = n^{-1/3}\Phi^{-1}(\tau)^{1/3}[3/2\phi^2(\xi)/(2\xi_1^2)]^{1/3}$ where $\xi = \Phi^{-1}(\tau)$. The Stata rule produces bandwidths that are generally considerably smaller than the S rule.

R, the very impressive GNU dialect of S. The current repository for this software is `www.econ.uiuc.edu/ roger/research/rq/rq.html`. The R version is also available at `lib.stat.cmu.edu/R/CRAN/`. The former website also provides software for the Ox language of Doornik (1996), and for Matlab. In addition to including code for the algorithms described above in fortran, this website offers a considerable variety of related code designed to facilitate inference and presentation of quantile regression results.

Although there is quite a large theoretical literature dealing with quantile regression inference, these developments have been slow to percolate into standard econometric software. With the notable exceptions of Stata and Xplore, see Cizek (2000), and the packages available for `Splus` and `R`, none of the implementations of quantile regression include any functionality for inference. We would welcome a concerted effort by readers to encourage software developers to rectify this situation.

6. Perils of Doing It

In mean regression there is a familiar litany of ailments that plague the econometric mind with doubt and dread. In a few happy cases the practitioner of quantile regression can rest a little easier than his mean-spirited colleague. In other cases, doubts are magnified. We will very briefly survey this territory and try to offer some guidance to available remedies and related literature. For severe or persistent symptoms consult your local econometrician.

6.1. **Robustness.**

From the outset an important motivation for quantile regression has been its inherent robustness to outlying observations in the response variable. While mean regression tends to follow a single outlier like a rat behind the Pied Piper, the influence of an outlying observation on $\hat{\beta}(\tau)$ is bounded. Indeed, moving observations away from the quantile regression fit in the y-direction has no effect whatsoever on the fit. This insensitivity has been occasionally misinterpreted, but it is fundamental to the nature of the quantiles. As in mean regression, outlying x observations can be highly influential in quantile regression, but the problem is not quite so severe. Several proposals have been made to "robustify" quantile regression against such points. See *e.g.* Rousseeuw and Hubert (1999) and the associated discussion.

6.2. **Heteroscedasticity.**

The recognition that covariates can exert a significant effect on the dispersion of the response variable as well as its location is the first step toward a general acceptance of the expanded flexibility of covariate effects in quantile regression. Some simple tests for heteroscedasticity were suggested in Koenker and Bassett (1982), and there has been considerable subsequent work on models for linear location and scale shifts. As in mean regression, variation in conditional dispersion creates opportunities for improved efficiency of estimation through weighted quantile regression. See Koenker and Zhao (1994) and Zhao (2000) for further details.

6.3. **Censoring.** In cases of fixed censoring like the well-known tobit model parametric maximum likelihood methods have attracted considerable criticism due to their sensitivity to both distributional assumptions and potential heteroscedasticity. See *e.g.* Goldberger (1983). The crucial observation of Powell (1986) that linear conditional quantile models could easily accommodate fixed censoring by a simple nonlinear modification of the response function has been enormously influential. More generally, as elaborated by Powell (1989) the quantiles are equivariant with respect to any monotone increasing transformation, so the transformed random variable $h(Y)$ has conditional quantile functions $Q_{h(Y)}(\tau) = h(Q_Y(\tau))$, a fact that considerably simplifies the interpretation of a wide variety of transformation models. See *e.g.* Buchinsky (1995) and Machado and Mata (2000). Random censoring constitutes a somewhat more challenging context for quantile regression, but there have been several important recent developments including: Honoré, Powell, and Khan (2000), Lipsitz, Fitzmaurice, Molenberghs, and Zhao (1997), Yang (1999), and Portnoy (2001).

6.4. **Sample Selection.** Semiparametric models of sample selection have received considerable attention in recent years as researchers explored various schemes to relax the parametric specifications employed in the seminal work of Heckman (1979). A valuable survey is provided by Manski (1993). A unified approach to sample selection for quantile regression remains a challenging open problem. See Buchinsky (1997, 2001) for some recent developments in this direction.

6.5. **Binary Response Models.** The maximum score estimator proposed by Manski (1975,1985) for the binary response model chooses $\hat{\beta}$ to maximize the agreement between the signs of the observed response variable, coded ± 1, and the signs of the linear predictor $x_i\beta$. Manski (1985) observed that this problem could be reformulated as a general quantile regression problem, but for reasons that are not altogether clear subsequent attention seems to have been focused almost entirely on the median case. Kordas (2000) has recently explored the consequences of estimating an ensemble of these binary quantile regression models, an approach that leads to a much more flexible view of how covariates influence the response. The asymptotic behavior of the joint distribution of these binary regression quantiles is studied following the smoothing approach introduced by Horowitz (1992). This approach seems to offer a very appealing empirical strategy for many discrete choice applications where heterogeneity of covariate effects is an important consideration.

6.6. **Duration Models.** Chaudhuri, Doksum, and Samarov (1997) have argued that quantile regression serves as a unifying concept for a variety of duration models: proportional hazards, proportional odds, accelerated failure time, etc. There are many potential econometric applications. Koenker and Geling (2001) describe a large scale application in experimental demography, and discuss an important link to the general notion of treatment effects introduced by Lehmann (1974). Fitzenberger (1996) and Koenker and Bilias (2001) discuss applications to unemployment duration models.

6.7. **Panel Data.** Additive random effects models for panel data would seem to offer a fertile field for the growth of quantile regression methods. Caution is called for, however. Quantiles of convolutions of random variables are rather intractable objects, and preliminary differencing strategies familiar from Gaussian models have sometimes unanticipated effects. Some progress has been made with minimum distance methods based on unrestricted quantile regression fitting to several cross sections, as in Chay (1995).

6.8. **Endogeneity.** Amemiya (1982) and Powell (1983) consider analogues of the two stage least squares estimator for median regression. And as noted in the previous section, there have been several recent applications of these methods. Recently, Abadie, Angrist, and Imbens (2000) have proposed a weighted quantile regression approach to estimating endogenous treatment effects in observational studies. That causal inference will continue to generate important (and controversial) work seems safe to predict.

6.9. **Time-Series.** There is a growing literature devoted to applications of quantile regression to time-series: Weiss (1991), Koul and Saleh (1995), Koul and Mukherjee (1994), Hallin and Jurečková (1999), Davis and Dunsmuir (1997), Hasan and Koenker (1997), Koenker and Zhao (1994). Interest has focused almost entirely, however, on the classical iid innovation model. The toy example of Koenker (1999) of a first order quantile autoregression model for daily temperature suggests that there may be interesting scope for non-iid innovation models as well.

6.10. **Extremes.** There is often a compelling substantive case for focusing attention on the behavior of conditional extreme values. This is certainly the case, for example, in auction applications in economics where such observations represent the winning bids. Or, in production-cost models where they represent firms near the technological frontier. Until recently the literature on this subject, like the data, has been rather sparse. Chernozhukov (2000) has dramatically altered this state of affairs by providing a very complete theory of the large sample behavior of extreme quantile regression, incorporating cases in which $n\tau_n$ converges to a constant as well as intermediate cases with $n\tau_n$ diverging, but $\tau_n \to 0$. This new theory significantly expands the domain of applicability of these methods, and introduces some important new ideas about inference.

6.11. **Nonlinear Models.** As for models of conditional mean functions, it is frequently useful to consider models for conditional quantile functions that are nonlinear in parameters. Powell's censored regression estimator is a prominent example. Nonlinear models pose some new computational problems since the strict linear programming formulation, and its underlying convexity is no longer available. Nevertheless, effective algorithms can be designed and the large sample theory of linear quantile regression is quite easily adapted to these cases. Fitzenberger (1997) considers the

Powell example in considerable detail. Koenker and Park (1996) describe a general approach to computation for nonlinear quantile regression problems based on interior point ideas.

6.12. **Nonparametric Models.** There is also a rapidly growing literature on nonparametric quantile regression. Locally polynomial methods have been extensively explored by Chaudhuri (1991), Welsh (1996) and others. Regression splines have been considered by Hendricks and Koenker (1991) and several variants of quantile regression smoothing splines have also been suggested. Koenker, Ng, and Portnoy (1994) suggest that total variation of the first derivative of the fitted function provides a natural alternative to the classical $\mathcal{L}_2$ smoothing penalty for univariate nonparametric quantile regression applications. Tibshirani (1996) has explored related penalties for other model selection problems.

6.13. **Multivariate Quantile Regression.** Quantile regression offers a coherent strategy for exploring the conditional distribution of the response variable in a wide variety of univariate regression-type settings. It is thus natural to ask whether there are associated methods for multivariate response models such as seemingly unrelated regression. A prior question is: how should we define a multivariate quantile? Unfortunately, this question has resisted a satisfactory resolution. Even in the case of the multivariate median, there has been considerable controversy as the excellent survey of Small (1990) portrays. Despite the very interesting recent work of Chaudhuri (1996) and Koltchinskii (1997), it seems fair to say that a fully successful approach remains elusive.

7. CONCLUSION

Much of applied statistics may be viewed as an elaboration of the linear regression model and associated estimation methods of least-squares. In beginning to describe these techniques Mosteller and Tukey (1977) in their influential text remark:

> What the regression curve does is give a grand summary for the averages of the distributions corresponding to the set of x's. We could go further and compute several different regression curves corresponding to the various percentage points of the distributions and thus get a more complete picture of the set. Ordinarily this is not done, and so regression often gives a rather incomplete picture. Just as the mean gives an incomplete picture of a single distribution, so the regression curve gives a correspondingly incomplete picture for a set of distributions.

We would like to think that quantile regression is gradually developing into a comprehensive strategy for completing the regression picture.

References

Abadie, A. (1997): "Changes in Spanish Labor Income Structure during the 1980's: A Quantile Regression Approach," *Investigaciones Economicas*, 21, 253–272.

Abadie, A., J. Angrist, and G. Imbens (2000): "Instrumental Variables Estimation of Quantile Treatment Effects," preprint.

Abreveya, J. (2001): "The effects of demographics and maternal behavior on the distribution of birth outcomes," *Empirical Economics*, forthcoming.

Amemiya, T. (1982): "Two Stage Least Absolute Deviations Estimators," *Econometrica*, 50, 689–712.

Andrews, D. W. K., and M. Buchinsky (2000): "A three-step method for choosing the number of bootstrap repetitions," *Econometrica*, 68, 23–52.

Arias, O., K. Hallock, and W. Sosa (2001): "Indidividual heterogeneity in the returns to schooling: instrumental variables quantile regression using Twins data," *Empirical Economics*.

Barrodale, I., and F. Roberts (1974): "Solution of an overdetermined system of equations in the ℓ_1 norm," *Communications of the ACM*, 17, 319–320.

Bassett, G., and H. Chen (2001): "Quantile Style: Return-Based Attribution Using Regression Quantiles," forthcoming.

Becker, R., J. Chambers, and A. Wilks (1988): *The New S Language: A Programming Environment for Data Analysis and Graphics*. Wadsworth: Pacific Grove.

Bickel, P., and A. Sakov (2000): "An Edgeworth expansion for the m out of n bootstrapped median," *Stat. and Prob. Letters*.

Bilias, Y., S. Chen, and Z. Ying (2000): "Simple resampling methods for censored regression quantiles," forthcoming.

Buchinsky, M. (1994): "Changes in US Wage Structure 1963-87: An Application of Quantile Regression," *Econometrica*, 62, 405–458.

——— (1995): "Quantile Regression, the Box-Cox Transformation Model and U.S. Wage Structure 1963-1987," *J. of Econometrics*, 65, 109–154.

——— (1997): "Women's return to education in the U.S.: Exploration by Quantile Regression," *J. Applied Econometrics*, 13, 1–30.

——— (1998): "Recent Advances in Quantile Regression Models: A practical guide for empirical research," *J. of Human Resources*, 33, 88–126.

——— (2000): "Quantile regression with sample selection: Estimating women's return to education in the U.S.," *Empirical Economics*, forthcoming.

Büttner, T., and B. Fitzenberger (1998): "Central Wage Bargaining and Local Wage Flexibility: Evidence from the Entire Wage Distribution," preprint.

Chamberlain, G. (1994): "Quantile regression censoring and the structure of wages," in *Advances in Econometrics*, ed. by C. Sims. Elsevier: New York.

Chaudhuri, P. (1991): "Global Nonparametric Estimation of Conditional Quantile Functions and Their Derivatives," *J. Mult. Analysis.*, 39, 246–269.

Chaudhuri, P. (1996): "On a geometric notion of quantiles for multivariate data," *J. of the American Statistical Association*, 91, 862–872.

Chaudhuri, P., K. Doksum, and A. Samarov (1997): "On average derivitive quantile regression," *Annals of Stat.*, 25, 715–744.

Chay, K. (1995): "Evaluating the impact of the Civil Rights Act of 1964 of 1964 on the Economic Status of Black Men Using Censored Longitudinal, Earnings Data," preprint.

Chernozhukov, V. (2000): "Conditional Extremes and Near-Extremes," preprint.

Chernozhukov, V., and L. Umantsev (2001): "Conditional Value at Risk: Aspects of Modeling and Estimation," forthcoming.

CIZEK, P. (2000): "Quantile regression," in *XploRe Application Guide*, ed. by W. Härdle , Z. Hlavka, and S. Klinke. Springer-Verlag: Heidelberg.

CONLEY, T., AND D. GALENSON (1998): "Nativity and Wealth in Mid-Nineteenth-Century Cities," *J. of Economic History*, 58, 468–93.

DAVIS, R. A., AND W. T. M. DUNSMUIR (1997): "Least Absolute Deviation Estimation for Regression With ARMA Errors," *Journal of Theoretical Probability*, 10, 481–497.

DEATON, A. (1997): *The Analysis of Household Surveys*. Johns Hopkins: Baltimore.

DOORNIK, J. (1996): *Ox: An object oriented matrix programming language*. Thompson: London.

EIDE, E., AND M. SHOWALTER (1998): "The Effect of School Quality on Student Performance: A Quantile Regression Approach," *Economics Letters*, 58, 345–50.

ENGEL, E. (1857): "Die Productions-und Consumtionsver-haltnisse des Konigreichs Sachsen," *Zeitschrift des Statistischen Bureaus des Koniglich Sachsischen Misisteriums des Innern*, 8, 1–54.

ENGLE, R., AND S. MANGANELLI (1999): "CaViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," preprint.

FITZENBERGER, B. (1996): "A Guide to Censored Quantile Regressions," in *Handbook of Statistics*, ed. by C. Rao, and G. Maddala. North-Holland: New York.

——— (1997): "Computational aspects of censorted quantile regression," in L_1-*Statistical Procedures and Related Topics*, ed. by Y. Dodge, IMS Lecture Notes-Monograph Series.

——— (1999): *Wages and employment across skill groups*. Physica-Verlag: Heidleberg.

FITZENBERGER, B., R. HUJER, T. MACURDY, AND R. SCHNABEL (2001): "Testing for uniform wage trends in West-Germany: A cohort analysis using quantile regressions for censored data," *Empirical Economics*, forthcoming.

FITZENBERGER, B., AND C. KURZ (1997): "New insights on earnings trends across skill groups and industries in West Germany," preprint.

FRISCH, R. (1934): *Statistical Confluence Analysis by Means of Complete Regression Systems*. Universitetets Økonomiske Institutt: Oslo.

GOLDBERGER, A. S. (1983): "Abnormal Selection Bias," in *Studies in Econometrics, Time Series, and Multivariate Statistics*, pp. 67–84. Academic (New York; London).

GOSLING, A., S. MACHIN, AND C. MEGHIR (1996): "What Has Happened to the Wages of Men since 1966?," in *New inequalities: The changing distribution of income and wealth in the United Kingdom*, ed. by J. Hills.

GUTENBRUNNER, C., AND J. JUREČKOVÁ (1992): "Regression quantile and regression rank score process in the linear model and derived statistics," *Ann. Statist.*, 20, 305–330.

GUTENBRUNNER, C., J. JUREČKOVÁ , R. KOENKER, AND S. PORTNOY (1993): "Tests of linear hypotheses based on regression rank scores," *J. of Nonparametric Statistics*, 2, 307–33.

HALL, P., AND S. SHEATHER (1988): "On the distribution of a studentized quantile," *J. Royal Stat. Soc. (B)*, 50, 381–391.

HALLIN, M., AND J. JUREČKOVÁ (1999): "Optimal tests for autoregressive models based on autoregression rank scores," *The Annals of Statistics*, 27, 1385–1414.

HASAN, M., AND R. KOENKER (1997): "Robust rank tests of the unit root hypothesis," *Eonometrica*, 65, 133–161.

HECKMAN, J. J. (1979): "Sample Selection Bias As a Specification Error," *Econometrica*, 47, 153–161.

HENDRICKS, W., AND R. KOENKER (1991): "Hierarchical spline models for conditional quantiles and the demand for electricity," *J. of Am. Stat. Assoc.*, 87, 58–68.

HONORÉ, B., J. POWELL, AND S. KHAN (2000): "Quantile regression under random censoring," preprint.

Horowitz, J. (1992): "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, 60, 505–531.

——— (1998): "Bootstrap methods for median regression models," *Econometrica*, 66, 1327–1352.

Kahn, L. (1998): "Collective Bargaining and the Interindustry Wage Structure: International Evidence," *Economica*, 65, 507–34.

Karmarker, N. (1984): "A new polynomial time algorithm for linear programming," *Combinatorica*, 4, 373–395.

Knight, K., G. Bassett, and M. Tam (2000): "Comparing Quantile Estimators for the linear model," preprint.

Koenker, R. (1994): "Confidence intervals for regression quantiles," in *Asymptotic Statistics: Proceedings of the 5th Prague Symposium on Asymptotic Statistics*, ed. by M. P., and M. Hušková . Physica-Verlag: Heidleberg.

——— (1996): "Rank Tests for Linear Models," in *Handbook of Statistics*, ed. by C. Rao, and G. Maddala. North-Holland: New York.

——— (1999): "Galton, Edgeworth, Frisch and Prospects for Quantile Regression in Econometrics," *J. Econometrics*, forthcoming.

Koenker, R., and G. Bassett (1978): "Regression Quantiles," *Econometrica*, 46, 33–50.

——— (1982): "Robust tests for heteroscedasticity based on regression quantiles," *Econometrica*, 50, 43–61.

Koenker, R., and Y. Bilias (2001): "Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments," *Empirical Economics*, forthcoming.

Koenker, R., and V. d'Orey (1987): "Computing Regression Quantiles," *Applied Statistics*, 36, 383–393.

——— (1993): "A Remark on Computing Regression Quantiles," *Applied Statistics*, 36, 383–393.

Koenker, R., and O. Geling (2001): "Reappraising Medfly Longevity: A quantile regression survival analysis," *J. of Am. Stat. Assoc.*, forthcoming.

Koenker, R., P. Ng, and S. Portnoy (1994): "Quantile Smoothing Splines," *Biometrika*, 81, 673–80.

Koenker, R., and B. Park (1996): "An interior point algorithm for nonlinear quantile regression," *Journal of Econometrics*, 71, 265–283.

Koenker, R., and Q. Zhao (1994): "L-Estimation for linear heteroscedastic models," *J. of Nonparametric Statistics*, 3, 223–235.

Koltchinskii, V. (1997): "M-estimation, convexity, and quantiles," *Annals of Stat.*, 25, 435–477.

Koopmans, T. (1937): *Linear Regression Analysis of Economic Time Series*. Netherlands Economic Institute: Haarlem.

Kordas, G. (2000): "Binary Regression Quantiles," preprint.

Koul, H., and R. Mukherjee (1994): "Regression quantiles and related processes under long range dependent errors," *J. Mult. Anal*, 51, 318–337.

Koul, H., and A. Saleh (1995): "Autoregression quantiles and related rank score processes," *Annals of Statistics*, 23, 670–689.

Lehmann, E. (1974): *Nonparametrics: Statistical Methods based on Ranks*. Holden-Day.

Levin, J. (2001): "Where the reductions count: A quantile regression analysis of effects of class size on scholastic acheivement," *Empirical Economics*, forthcoming.

Lipsitz, S., G. Fitzmaurice, G. Molenberghs, and L. Zhao (1997): "Quantile regression methods for longitudinal data with Drop-outs: Application CD4 cell counts of patients infected with HIV," *Applied Stat.*, pp. 463–476.

Lopez, A., P. Hernandez, and J. Garcia (2001): "How wide is the gap? An investigation of gender wage differences using quantile regression," *Empirical Economics*, forthcoming.

MACHADO, J., AND J. MATA (2000): "Box-Cox Quantile regression and the distribution of firm sizes," *Journal of Applied Econometrics*, pp. 253–274.
——— (2001): "Counterfactual decomposition of changes in wage distributions using quantile regression," *Empirical Economics*, forthcoming.
MANNING, W., L. BLUMBERG, AND L. MOULTON (1995): "The demand for alcohol: the differential response to price," *Journal of Health Economics*, 14, 123–148.
MANSKI, C. (1975): "Maximum score estimation of the stochastic utility model of choice," *J. of Eonometrics*, 3, 205–228.
——— (1985): "Semiparametric analysis of discrete response asymptotic properties of the maximum score estimator," *J. of Eonometrics*, 27, 205–228.
MANSKI, C. (1993): "The Selection Problem in Eocnometrics and Statistics," in *Handbook of Statistics*, ed. by C. R. G.S. Maddala, and H. Vinod. Elsevier: New York.
MONTENEGRO, C. (1998): "The Structure of Wages in Chile 1960-1996: An Application of Quantile Regression," *Estudios de Economia*, 25, 71–98.
MORILLO, D. (2000): "Income mobility with nonparametric quantiles: A comparison of the U.S. and Germany," preprint.
MOSTELLER, F., AND J. TUKEY (1977): *Data Analysis and Regression: A Second Course in Statistics.* Addison-Wesley, Reading Mass.
MUELLER, R. (2000): "Public and Private Wage Differentials in Canada Revisited," *Industrial Relations*, 39, 375–400.
PARZEN, M. I., L. WEI, AND Z. YING (1994): "A resampling method based on pivotal estimating functions," *Biometrika*, 81, 341–350.
PORTNOY, S. (2001): "Censored Quantile Regression," preprint.
PORTNOY, S., AND R. KOENKER (1997): "The Gaussian Hare and the Laplacian Tortoise: Computability of squared-error versus absolute-error estimators, with discusssion," *Stat. Science*, 12, 279–300.
POTERBA, J., AND K. RUEBEN (1995): "The Distribution of Public Sector Wage Premia: New Evidence Using Quantile Regression Methods," NBER Working paper #4734.
POWELL, J. L. (1983): "The Asymptotic Normality of Two-stage Least Absolute Deviations Estimators," *Econometrica*, 51, 1569–1576.
——— (1986): "Censored Regression Quantiles," *J. Econometrics*, 32, 143–155, Asymptotics for censored quantile regression.
——— (1989): "Estimation of monotonic regression models under quantile restrictions," in *Nonparametric and Semiparametric Methods in Econometrics*, ed. by W. Barnett, J. Powell, and G. Tauchen. Cambridge U. Press: Cambridge.
ROGERS, W. (1992): "Quantile regression standard errors," *Stata Technical Bulletin*, 9, 16–19.
——— (1993): "Calculation of quantile regression standard errors," *Stata Technical Bulletin*, 13, 18–19.
ROUSSEEUW, P., AND M. HUBERT (1999): "Regression Depth," *J. of Am. Stat. Assoc.*, 94, 388–433.
SCHULTZ, T., AND G. MWABU (1998): "Labor unions and the distribution of wages and employment in South Africa," *Ind. and Labor Relations Rev.*, 51, 680–703.
SMALL, C. (1990): "A survey of multidimensional medians," *International Statistical Review*, 58, 263–277.
TANNURI, M. (2000): "Are the assimilation processes of low and high income immigrants distinct? A Quantile Regression Approach," preprint.
TAYLOR, J. (1999): "A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns," *Journal of Derivatives*, pp. 64–78.

TIBSHIRANI, R. (1996): "Regression Shrinkage and Selection Via the Lasso," *J. Royal Stat. Soc. (B)*, 58, 267–288.

TREDE, M. (1998): "Making Mobility Visible: A Graphical Device," *Economics Letters*, 59, 77–82.

VISCUSI, W., AND P. BORN (1995): "Medical Malpractice Insurance in the Wake of Liability Reform," *J. of Legal Studies*, 24, 463–90.

VISCUSI, W., AND J. HAMILTON (1999): "Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions," *American Economic Review*, 89, 1010–27.

WEISS, A. A. (1991): "Estimating Nonlinear Dynamic Models Using Least Absolute Error Estimation," *Econometric Theory*, 7, 46–68.

WELSH, A. H. (1996): "Robust Estimation of Smooth Regression and Spread Functions and Their Derivatives," *StSinica*, 6, 347–366.

YANG, S. (1999): "Censored median regression using weighted empirical survival and hazard functions," *J. of Am. Stat. Assoc.*, 94, 137–145.

ZHAO, Q. (2000): "Asymptotically efficient median regression in the presence of heteroscedasticity of unknown form," *J. Econometrics*, forthcoming.