NBER WORKING PAPER SERIES

TOBIT AT FIFTY: A BRIEF HISTORY OF TOBIN'S REMARKABLE ESTIMATOR, OF RELATED EMPIRICAL METHODS, AND OF LIMITED DEPENDENT VARIABLE ECONOMETRICS IN HEALTH ECONOMICS

Kohei Enami
John Mullahy

Working Paper 14512
http://www.nber.org/papers/w14512

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2008

Thanks are owed to Art Goldberger for sharing his insights about the evolution of Tobit and Orbit, and to Don Hester for sharing his recollections about computational issues in the production of Tobin's paper. The first author's work was supported in part by a Merck Quantitative Sciences Fellowship in Health Economics. Some of the second author's research was conducted when he was a visiting scholar at the Geary Institute at UCD. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2008 by Kohei Enami and John Mullahy. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Tobit at Fifty: A Brief History of Tobin's Remarkable Estimator, of Related Empirical Methods, and of Limited Dependent Variable Econometrics in Health Economics
Kohei Enami and John Mullahy
NBER Working Paper No. 14512
November 2008
JEL No. I1

ABSTRACT

Practitioners of empirical health economics might be forgiven for paying little heed to the recent 50th anniversary of the publication of one of the most important papers in its methodological heritage: James Tobin's widely-cited 1958 Econometrica paper that developed what later became known as the Tobit estimator. This golden anniversary milestone provides a fitting opportunity to reflect on Tobin's contribution and to assess the role that econometric limited dependent variable modeling has played in empirical health economics. Of primary focus here is how Tobin's estimator came to be and came to take root in empirical health economics. The paper provides a brief history of Tobin's estimator and related methods up through about 1971, discusses the early applications of Tobit and related estimators in health economics, i.e. the "technology diffusion" of Tobit in health economics, and offers some concluding remarks.

Kohei Enami
University of Wisconsin-Madison
Dept. of Population Health Sciences
707 WARF, 610 N. Walnut St.
Madison, WI 53726
enami@wisc.edu

John Mullahy
University of Wisconsin-Madison
Dept. of Population Health Sciences
787 WARF, 610 N. Walnut Street
Madison, WI 53726
and NBER
jmullahy@facstaff.wisc.edu

1. Introduction

Because we tend generally to focus our efforts on applications of modern quantitative methods, practitioners of empirical health economics might be forgiven for paying little heed to the recent 50th anniversary of the publication of one of the most important papers in its methodological heritage: James Tobin's widely-cited *Econometrica* paper (Tobin, 1958) that developed what later became known as the Tobit estimator. This golden anniversary milestone provides a fitting opportunity to reflect on Tobin's contribution and to assess the role that econometric limited dependent variable modeling has played in empirical health economics.

Roughly contemporaneous with the publication of Tobin's paper in January 1958, other developments in econometric methodology and inquiry were from today's perspective -- in particular, the perspective of applied health economics -- quite noteworthy. In the October 1957 issue of *Econometrica*[1] appeared the classic "Hybrid Corn" paper on technological change by Zvi Griliches (Griliches, 1957).[2] Later in 1958 *Econometrica* published John Denis Sargan's seminal paper on instrumental variables estimation that derived the asymptotic properties of IV estimators in a modern, moment-based context (Sargan, 1958), and another paper by Irving Hoch that -- while less well-known among today's applied economists -- instructed importantly about biases that are likely to arise when production functions are estimated econometrically in contexts where input choices may depend on unobservable determinants of outputs (Hoch, 1958).[3] Also at this time

---

[1] In 1958, *Econometrica* was published in four issues or numbers compared with today's six, and French-language articles were still relatively common. In January 1958, the journal was edited by managing editor Robert Strotz and co-edited by Edmond Malinvaud. All six members of the January 1958 Editorial Board ultimately became Nobel Laureats (Ragnar Frisch, Milton Friedman, J.R. Hicks, Tjalling Koopmans, Paul Samuelson, and J.R.N. Stone), as did six of the fourteen associate editors (Lawrence Klein, Tjalling Koopmans, Herbert Simon, Robert Solow, J.R.N. Stone, and J. Tinbergen).

[2] See Jonathan Skinner and Douglas Staiger, 2005, for an application of Griliches's idea to technology adoption in a health care setting.

[3] On the silver anniversary of the publication of their own seminal work, it might be noted

(continued)

appeared in *JASA* Joseph Berkson's classic discussion of smoking and lung cancer (Berkson, 1958), as well as Daniel Suits's seminal paper on the coding and interpretation of categorical dummy variables in regression models (Suits, 1957).

In 1958 the impact that Tobin's contribution would ultimately have on applied microeconometric work could hardly have been anticipated. And while the applications of the specific censored-normal regression estimator Tobin devised are many,[4] the ultimate impact of Tobin's paper has been far greater: Tobin's contribution raised awareness of the implications of analyzing any economic outcome having limited range, and spawned an enormous body of methodological research on limited dependent variable estimation strategies.

Of primary focus here is how Tobin's estimator came to be, came to take root in empirical health economics, and came to influence subsequent analytical developments. This paper provides a brief history of Tobin's estimator and related methods up through about 1971 (section 2); discusses the early applications of Tobit and related estimators in health economics, i.e. the "technology diffusion" of Tobit in health economics (section 3); and offers some concluding remarks (section 4). The scope is intentionally modest; readers interested in other perspectives on the history and applications of econometric methods in health economics are referred to Martin Feldstein, 1974, Joseph Newhouse, 1987, and Andrew Jones, 2000.

(cont.)
that Hoch's paper was not cited by Mark Rosenzweig and Paul Schultz, 1983, in their treatment of the estimation of health production functions in settings where input choices may be influenced by factors unobserved by econometricians but possibly known by health producers. Rosenzweig and Schultz do, however, cite the better-known paper by Yair Mundlak and Hoch, 1965, that is an extension of Hoch's 1958 paper.

[4] Tobin's estimation strategy has been influential not only within economics but also beyond. A JSTOR full-text search on "tobit" conducted on Oct. 9, 2008, had 2,074 hits in economics, but also 149 in statistics, 193 in political science, and 402 in sociology. There were also 735 hits in religion, but most of these relate to the Book of Tobit in the Old Testament.

2. Tobin and Tobit

Probit and Beyond

Tobin's censored-regression estimator grew out of earlier unpublished work he did on what would now be considered standard binary probit analysis, i.e.

$$\Pr\left(y = 1 \middle| X_1, \ldots, X_k\right) = \Phi\left(X_1\beta_1 + \ldots + X_k\beta_k\right),$$

with

$$y = 1\left(y^* > 0\right) \text{ and } y^* \sim N\left(X_1\beta_1 + \ldots + X_k\beta_k, 1\right).$$

Binary probit analysis had just come to be recognized in the 1940s in biometric, typically dose-response, contexts.[5] Tobin's extension and application (Tobin, 1955) of binary probit analysis to the multivariate regression context in a model of determinants of households' decisions to engage in any spending on consumer durables was itself a remarkable contribution.[6] Notably, this paper appeared as No. 1 in the Cowles Foundation Discussion

---

[5] See the citations in Tobin, 1955.

[6] The application of binary probit analysis to microdata was itself novel in economics, although the methodology had been exposited much earlier in biometrics by Frank Garwood, 1941. Garwood's paper derives the probit MLE for grouped data, but his likelihood function is also suitable for individual data (zero or n successes in groups of size n, where n could be 1). Garwood's likelihood function was written in the original paper as

$$\prod \frac{n!}{s!(n-s)!} P^{n-s} Q^{s}$$

with corresponding log-likelihood function, which was written as

$$\ln L = \text{const.} + S\left[(n-s)\log P + s\log Q\right].$$

(In modern practice this would typically be written as

$$\ln L = \text{const.} + \sum (n-y)\ln(1-p) + y\ln p,$$

where p would be $p(\mathbf{x};\boldsymbol{\beta})$ in a regression setting, typically $\Phi(\mathbf{x}\boldsymbol{\beta})$ for probit. "S" was the common summation operator in earlier days). With binary microdata, const.=ln(1)=0, n=1 and $y \in \{0,1\}$.

Paper Series (henceforth CFDP-1); the original draft was dated July 14, 1955, with a revision dated December 1, 1955.[7]

The sample Tobin used for his probit estimation comprised 1,036 observations drawn from the 1952 and 1953 Surveys of Consumer Finances; the model included two explanatory variables (X_1=disposable income, X_2=liquid asset holdings) plus a constant term. Not least remarkable in this research was its computation of the maximum likelihood estimates of his probit model by brute-force calculation and table lookup.

Of ultimately greater interest for purposes at hand were the concluding sentences of CFDP-1, which provide some suggestion that Tobin was already imagining the development of "Tobin's probit":

> The example has been presented for illustration of a method rather than for substantive results. Still more variables would be needed for a better explanation of durable goods purchasing behavior. Moreover, it is wasteful of information to disregard amounts spent by those who purchased. *Some combination of probit analysis and regression is indicated, to handle a variable with a large probability of having zero value and the remaining probability spread over a positive interval*. [Italics added]

Tobin's Maximum Likelihood Estimator for Censored Dependent Variables

Tobin wasted little time. Dated September 15, 1955, "Estimation of Relationships for Limited Dependent Variables" was issued as Cowles Foundation Discussion Paper No. 3, with a revision dated July 25, 1956 issued as CFDP No. 3R. A year and a half later a revised version of CFDP-3R was published in the January, 1958 issue of *Econometrica*.[8]

Tobin began his discussion with the following:

---

[7] Tobin became Director of the Cowles Foundation in 1955 at the time the Foundation moved from Chicago to Yale (Robert Shiller and Tobin, 1999).

[8] The published version was nearly identical to CFDP-3R. In the published paper: the reference to CFDP-1 that appeared in CDFP-3R was deleted; a reference to the "method of scoring" with a corresponding citation to C.R. Rao's *Advanced Statistical Methods in Biometric Research* was added; several sentences on the estimation algorithm were added on p. 30; figure 2 and table VI were added; and a variety of minor grammatical and typographical modifications were made.

> In economic surveys of households, many variables have the following characteristics: The variable has a lower, or upper, limit and takes on the limiting value for a substantial number of respondents. For the remaining respondents, the variable takes on a wide range of values above, or below, the limit.
>
> Tobin, 1958, p. 24

Picking up on the theme with which he concluded CFDP-1, Tobin then went on to write:

> If only the probability of limit and non-limit responses, without regard for the value of non-limit responses were to be explained, *probit analysis* provides a suitable statistical model.... But it is inefficient to throw away information on the value of the dependent variable when it is available. If only the value of the variable were to be explained, if there were no concentration of observations at a limit, *multiple regression* would be an appropriate statistical technique. But when there is such concentration, the assumptions of the multiple regression model are not realized....
>
> A hybrid of probit analysis and multiple regression seems to be called for, and it is the purpose of this paper to present such a model.
>
> Tobin, 1958, p. 25

One standard modern description of the limited dependent variable structure thus described would be:

$$y^* = \mathbf{x}\boldsymbol{\beta} + u, \quad u \sim N\left(0, \sigma^2\right)$$

$$E\left[y^* \mid \mathbf{x}\right] = \mathbf{x}\boldsymbol{\beta}$$

$$y = \begin{cases} L, & y^* < L \\ y^*, & y^* \geq L \end{cases}.$$

In Tobin's notation, the role of y was played by "W" and the role of $E[y^*|\mathbf{x}]$ was played by "Y"; there was no specific symbol used to describe the latent random variable y^*. Specifically, Tobin uses

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_m X_m$$

and then bases estimation on the normalized coefficients

$$\left(a_0, a_1, a_2, \ldots, a_m, a\right) = \left(\beta_0 / \sigma, \beta_1 / \sigma, \beta_2 / \sigma, \ldots, \beta_m / \sigma, 1 / \sigma\right).$$

Figure 1 reproduces figure 1 from Tobin's paper and exhibits one of the main foci of Tobin's paper, i.e. the discrepancy between E[W] and Y (using his notation).

Tobin's empirical analysis again drew on the 1952 and 1953 Surveys of Consumer Finances, now using a sample of 735 primary nonfarm spending units. The observed censored dependent variable, W, is defined as "the ratio of 1951-52 total durable goods expenditure to 1951-52 total disposable income." The explanatory variables are $\mathbf{x}=[1,X_1,X_2]$, with X_1 defined as "the age of the head of the spending unit, as reported in 1953,"[9] and X_2 defined as "the ratio of liquid asset holdings at the beginning of 1951 to 1951-52 total disposable income."[10] W is defined as a ratio of expenditure to income, but a ratio that can presumably exceed one in situations where spenders are dissavers.[11] The maximum likelihood estimates of the normalized parameters a_j are presented in Table II in Tobin's paper and reproduced here in table 2. Age and liquid assets contribute negatively and positively, respectively, to the latent spending index.

Many aspects of the paper are noteworthy. Tobit's derivation allowed the lower limit of observations to be an arbitrary real value L -- thus enhancing the applicability of the estimator -- instead of zero, which would have been a natural threshold given the

[9] This variable was coded as "1" for ages 18-24, "2" for ages 25-34, "3" for ages 35-44, "4" for ages 45-54, "5" for ages 55-64, and "6" for ages 65 or more. Tobin makes no mention of why this coding strategy was used instead of, e.g., simply including age recorded in natural units.

[10] Note that this is not the same **x**-vector as used in CFDP-1.

[11] Given the data as reproduced here in Table 1, the unconditional mean and variance of W can be computed as .083 and .011, respectively.

expenditure data he was analyzing. Moreover, while much subsequent discussion in the literature of Tobin's estimator focused on estimation of the parameters of the latent index $E[y^* \mid \mathbf{x}] = \mathbf{x\beta}$,[12] Tobin derived the explicit formula for the conditional mean of W -- which he termed "the expected value of W for given values of Y and L" -- and devoted considerable attention to the properties of this parameter (see esp. pp. 35-36 in Tobin, 1958).

In addition, one and a half printed pages were dedicated to the derivation of the likelihood function and the first and second derivatives of its log, with another one and a half pages used to describe the Newton solution algorithm. The empirical parameter estimate iterative solution sequence, a detail that would surely never appear in a modern publication, was presented in two full tables. Given the significant effort involved in 1955 in undertaking what would by today's standards be considered a routine nonlinear regression estimation exercise, it is hardly surprising that the mechanics of the computations were noteworthy and publishable. Regarding computation, Donald Hester recalled (personal communication, Nov. 17, 2008):

> I was an undergraduate research assistant...for Tobin at Yale in 1954-55 and did the calculations on a Monroe desk calculator. It was a bit of a challenge, because one of the papers Tobin was drawing on (by Cornfield and Mantel if I recall correctly) had a missing negative sign and for about six weeks I was minimizing rather than maximizing a likelihood function. After Tobin derived the function himself, the iterations converged rather nicely.[13]

---

[12] Postulating Tobin's conditional mean structure

$$E[y|\mathbf{x}] = \Phi(\mathbf{x\beta}/\sigma)\mathbf{x\beta} + \sigma\phi(\mathbf{x\beta}/\sigma)$$

as a *prior* conditional mean model for a nonnegative dependent variable seems far-fetched. For instance, juxtapose this with streamlined current-day specifications like an exponential conditional mean $E[y|\mathbf{x}] = \exp(\mathbf{x\beta})$.

[13] See Ray Fair, 1977, and Randall Olsen, 1978, on further aspects of Tobit model estimation.

Early Related Approaches

Randall Olsen, 1978, credits Anders Hald, 1949, with the development of the maximum likelihood estimator for outcomes of what Hald termed "censored" distributions. Olsen wrote: "The basic form of the model was suggested by Hald **[5]** with a later independent paper by Tobin **[7]** developing the model in a regression setting." Hald carefully distinguished the "censored-" and the "truncated-" distribution cases, and provided a unified estimation framework in a non-regression framework to accommodate both situations. Yet Hald, in turn, credits W.L. Stevens (see Bliss, 1937) with the development of the "censored" normal model and its estimation.

Pre-dating the publication of Tobin's paper by three years, John Aitchison, 1955, exposited a precursor of what has subsequently come to be known as the two-part model in econometrics.[14] While Aitchison's exposition was not in a regression setting, it provides nonetheless an analytical foundation to limited dependent variable estimation that clearly complements Tobin's:

> In a study of household expenditures it is often of interest to estimate, from a sample of household budgets, the mean expenditure per household on a certain commodity, say children's clothing. Over the period of investigation it may well happen that a number of households in the sample spend nothing on children's clothing whereas the expenditures by the remainder of the households necessarily arise from the distribution of a positive variable, probably skew and possibly approximated by a lognormal curve. If such is the case, then clearly the correct procedure in any analysis is to recognize explicitly this dichotomy of the population into the categories, spender and non-spender. (Aitchison, 1955)

Aitchison's model for outcome variable X (using his notation) is:

[14] Neither Aitchison's nor Hald's paper was cited by Tobin in his 1958 *Econometrica* paper. Younger readers may be interested to learn, however, that Tobin did not have access to Google searching in the late 1950s.

$$P\{X=0\}=\theta, \quad P\{X>0\}=1-\theta$$

$$P\{X\subset(x,x+dx)|x>0\}=g(x)dx$$

$$P\{X\subset(x,x+dx)\}=(1-\theta)g(x)dx, \quad x>0$$

Aitchison defined further α and β as the mean and variance of g(x) and γ and δ as the mean and variance of the distribution of X. The emphasis in his paper was primarily on efficient estimation of γ and δ.

In a remarkable one-two punch in his classic textbook *Econometric Theory*, Arthur Goldberger, 1964, not only coined the term "Tobit model" in expositing Tobin's 1958 estimator,[15] but in the preceding page provided probably the first systematic exposition of an econometric two-part regression model (2PM) as an estimator of E[y|**x**] (where "y" is the observed nonnegative outcome), which he terms the "twin linear probability function" (Goldberger, 1964, pp. 252-253). Goldberger credits[16] Guy Orcutt with the development of the twin linear probability model, noting further that it was "known locally as the *Orbit* model." In the Orbit model, linear regressions are specified for both the binary {0,1} outcome as well as for the magnitude of the positive outcome with a restriction to the sample of positives, and E[y|**x**] is estimated as the product over the entire sample of the

---

[15] Tobin's response to Shiller's (1999) interview inquiry about "Tobit":

> Tobit -- well that's related to "probit," so that's understandable. But it was also related to a reference to me in a novel by Herman Wouk, a friend of mine in the officers' training school in 1942, called *The Caine Mutiny*, where I appear for one or two sentences in the first chapter, and I'm named in a thinly disguised way as Tobit. I asked Arthur Goldberger why he used this label in his statistics text, whether it was *The Caine Mutiny* or just the elision of Tobin and "probit." He wouldn't say. So I don't know.

Yet in a personal communication (Nov. 6, 2008) Art Goldberger noted that he didn't know of *The Caine Mutiny* reference at the time, and that his "...coinage of Tobit was simply a slurring of Tobin & Probit..."

[16] Personal communication, Nov. 6, 2008.

two component estimates.[17] Goldberger's 1964 exposition of the twin linear probability function was:

$$
\begin{aligned}
Ey &= \sum y \bullet f(y) = \sum_{y=0} 0 \bullet f(0) + \sum_{y>0} y \bullet f(y) \\
&= 0 + \sum_{y>0} \left\{ y \bullet \left[f(y \mid y > 0) \right] \bullet f(y > 0) \right\} \\
&= f(y > 0) \bullet \sum_{y>0} \left\{ y \bullet \left[f(y \mid y > 0) \right] \right\} \\
&= \text{Prob}\left\{ y > 0 \right\} \bullet E(y \mid y > 0)
\end{aligned}
$$

Indeed, in the early 1960s a research program focusing on durable goods and other consumer expenditure patterns at the Social Sciences Research Institute at the University of Wisconsin-Madison[18] deployed regression variants of such two-part models (e.g. Goldberger and Maw Lin Lee, 1962; Janet Fisher, 1962; De-Min Wu, 1965). In this research, two regression equations were estimated for each expenditure category: a linear probability model for the {0,1} indicator of any spending; and a linear regression model for the magnitude of spending estimated only on the observations for which positive expenditure was recorded. Arnold Zellner, 1990, offered the following assessment that, among other things, links this research program to the two-part modeling strategy in health econometrics:

[17] Goldberger (perhaps) anticipates some recent one-part modeling strategies (e.g. John Mullahy, 1998, and Joshua Angrist, 2001) with "...the procedure does not directly attempt to fit Ey to y...."

[18] See Orcutt, 1990, and Harold Watts, 1991, on the development of the SSRI under Orcutt's leadership.

> The two-equation model used by Chau, Goldberger, Orcutt, and others at Wisconsin in the 1960s is still central in analyses of micro-data in many areas and has resulted in reinterpretations of firms' investment and divided **[sic]** behavior -- see Laub (1972) and Peck (1974) and consumers' behavior with respect to medical expenditures. Indeed, it has been the subject of a lively, recent exchange, Duan et al. (1983) on the relative merits of the "sample selection bias" model and the "two-part" and "four-part" models (new names for old Wisconsin models) in Duan et al. (1984).

Several years later, John Cragg, 1971, provided a comprehensive treatment of such 2PM or Orbit models. He offered this rationale for why such structures might describe well certain outcomes[19]:

> In some situations the decision to acquire and the amount of the acquisition may not be so intimately related. In particular, even when the probability of a non-zero value is less than one half, one might not feel that the values close to zero are more probable than the ones near some larger value, given that a positive value will occur.

In Cragg's framework, the second part of the model (for y>0) could be treated in a linear regression framework use a log-transformation of the positive y values (which works whether or not these positive y values are lognormally distributed) *or* could be treated in a maximum likelihood context by assuming the positive y values followed a truncated normal distribution. Cragg's probit/truncated-normal model was exposited (using now-standard notation) as:

$$\text{Prob}(y_t = 0 \mid \mathbf{x}_{1t}, \mathbf{x}_{2t}) = \Phi(-\mathbf{x}_{1t}\boldsymbol{\beta})$$

and

$$f(y_t \mid \mathbf{x}_{1t}, \mathbf{x}_{2t}) = (2\pi)^{-1/2}\ \sigma^{-1} \exp\left\{-(y_t - \mathbf{x}_{2t}\gamma)^2 / 2\sigma^2\right\} \Phi(\mathbf{x}_{1t}\beta) / \Phi(\mathbf{x}_{2t}\gamma / \sigma), \quad \text{for } y_t > 0$$

Cragg noted that the Tobit model arose as a restricted version of his probit/truncated-

[19] Cragg's paper appears to have coined the now-familiar term "hurdle" model.

normal 2PM with $\mathbf{x}_1 = \mathbf{x}_2$ and $\beta = \gamma / \sigma$.[20]

3. Early Limited Dependent Variable Modeling in Health Economics

Given its computational complexity and the relative paucity of microdata at the time, it is not surprising that the Tobit estimation technology had initially a slow diffusion rate. Yet given the LDV measures that described interesting outcomes in health individual- or firm-level data available even then, it is also not surprising that Tobin's estimator and related approaches attracted the attention of empirical economists working in health as early as the mid-1960s.

Among the earliest recognitions of the LDV phenomenon in health economics was that by William Comanor, 1965. In his study of technological change in the pharmaceutical industry, one of the main dependent variables is a firm's total sales over two years of all new chemical entities introduced by the firm of the period 1955-1960. For 17 of the 57 firms in Comanor's sample, this measure was zero. Using least squares regression, Comanor estimated his model of sales determinants both with and without the inclusion of these seventeen observations (in modern terminology, a "one-part model" and "part two of a two-part model," respectively), finding sizable (threefold) differences in the point estimates of the key R&D parameter between the two. Comanor commented in a footnote:

> While a more sophisticated technique exists for dealing with this matter, the computational problems involved were too great to warrant its use. See James Tobin, 'Estimation of....'

As software capable of tackling nonlinear estimation with microdata samples became more widely available in the mid-1970s,[21] the nascent health economics literature witnessed

[20] Tsai-Fen Lin and Peter Schmidt, 1984, devised a Lagrange multiplier strategy for testing this parameter restriction.

[21] For an informative overview of the econometric software evolution, see Charles Renfro, (continued)

a surge of scholarly empirical work in which limited dependent variable model estimators were deployed in applied work.

Richard Rosett and Lien-fu Huang's, 1973, study of health insurance effects on medical care use and spending appears to be the first published application of Tobit analysis of a topic in health economics. Rosett and Huang used data from the 1960 Survey of Consumer Expenditures to estimate price and income demand elasticities. The dependent variable was actual or imputed spending on health services, specifically hospitalization plus physican services. Rosett and Huang used what they called "the probit-regression model of Tobin" to model this censored-at-zero outcome as a function of covariates. In the same year, Charles Phelps, 1973 (his dissertation), analyzed using Tobit and other methods data from a national survey undertaken by the Center for Health Administration Studies at the University of Chicago to study determinants of spending on health services and health insurance premiums.

The LDV-based health economics literature began to expand rapidly soon thereafter. In 1975, three important papers were, at least in part, Tobit-based. Frank Sloan and Somchai Richupan, 1975, soon followed with an analysis of nurse labor supply based on data from the 1960 U.S. Census. They used both Tobit analysis as well as the twin linear probability model (or linear-linear 2PM) sketched by Goldberger.[22] In fact, Sloan and Richupan used an instrumental variable version of the Tobit estimator in which predicted wage rates were used as a RHS covariate.

Jan Acton's classic study (Acton, 1975) of nonmonetary cost impacts on several forms of health care utilization in New York City uses linear simultaneous equations methods to control for endogeneity of price (distance) and utilization measures. Anticipating a

---

(cont.)
2004, esp. pp. 22-33.

[22] It is noteworthy that Sloan and Richupan anticipated some of the later debates about the properties of the 2PM.

literature to emerge years later, Acton explicitly eschews using Tobit methods in his analysis: "Estimating a simultaneous Tobit system with 12 or so endogenous variables is probably unwarranted with this data base." Finally, Karen Davis and Roger Reynolds, 1975, use data from the 1969 National Health Interview Survey to examine determinants of health care utilization for Medicare beneficiaries. Since their outcomes are counts of various forms of healthcare utilization with nontrivial zero frequencies, they utilize Tobit methods to estimate their utilization models.[23]

Soon to follow this pivotal mid-1970s literature were the aforementioned innovations in[24] and controversies surrounding[25] estimation methodologies that arose from the RAND HIE, a surge of development of generally likelihood-based estimation methodologies for LDV models[26], and the literature on selection bias correction stimulated by James Heckman's work[27]. A comprehensive assessment of this subsequent literature is beyond the scope of the present paper.

4. Conclusions

While some limitations of Tobin's original estimator have come to light since 1958,[28] from the perspective of the mid-1950s when Tobin wrote, the development of the Tobit estimator was nothing short of an econometric *tour de force*. The advances in

---

[23] The application of Poisson-type count data estimators in health economics was still several years in the future. See Jerry Hausman et al., 1984, for one of the earliest examples.

[24] See, e.g., Naihua Duan, et al., 1983, 1984, Newhouse, et al., 1980, and Willard Manning, et al., 1987.

[25] See, e.g., Joel Hay and Olsen, 1984.

[26] See G.S. Maddala, 1983.

[27] See Heckman, 1976, 1979, and Robert Haveman, 1997, for an overview of the use of selection models in policy contexts.

[28] For instance, several studies have identified the lack of robustness of the Tobit MLE when the baseline latent $N\left(x\beta, \sigma^2\right)$ assumptions fail to hold. See, e.g., Abbas Arabmazar and Schmidt, 1981, 1982, and James Powell, 1984, 1986.

microeconometric estimation it stimulated have had extraordinary influence on the practice of applied microeconomics in general, and -- given the vast array of LDV structures in health data -- on the evolution of scholarly empirical work in health economics in particular.

Acknowledgments

Thanks are owed to Art Goldberger for sharing his insights about the evolution of Tobit and Orbit, and to Don Hester for sharing his recollections about computational issues in the production of Tobin's paper. The first author's work was supported in part by a Merck Quantitative Sciences Fellowship in Health Economics. Some of the second author's research was conducted when he was a visiting scholar at the Geary Institute at UCD.

References

Acton, J.P. 1975. "Nonmonetary Factors in the Demand for Medical Services: Some Empirical Evidence." *Journal of Political Economy* 83: 595-614.

Aitchison, J. 1955. "On the Distribution of a Positive Random Variable Having a Discrete Probability Mass at the Origin." *Journal of the American Statistical Association* 50: 901-908.

Angrist, J.D. 2001. "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business and Economic Statistics* 19: 2-16.

Arabmazar, A. and P. Schmidt. 1981. "Further Evidence on the Robustness of the Tobit Estimator to Heteroskedasticity." *Journal of Econometrics* 17 253-258.

Arabmazar, A. and P. Schmidt. 1982. "An Investigation of the Robustness of the Tobit Estimator to Non-Normality." *Econometrica* 50: 1055-1063.

Berkson, J. 1958. "Smoking and Lung Cancer: Some Observations on Two Recent Reports." *JASA* 53: 28-38.

Bliss, C.I. 1937. "The Calculation of the Time-Mortality Curve." Appendix by W.L. Stevens: "The Truncated Normal Distribution." *Annals of Applied Biology* 24: 815-852.

Comanor, W.S. 1965. "Research and Technical Change in the Pharmaceutical Industry." *Review of Economics and Statistics* 47: 182-190.

Cragg, J.G. 1971. "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods." *Econometrica* 39: 829-844.

Davis, K. and R. Reynolds. 1975. "Medicare and the Utilization of Health Care Services by the Elderly." *Journal of Human Resources* 10: 361-377.

Duan, N. et al. 1983. "A Comparison of Alternative Models for the Demand for Medical Care." *Journal of Business & Economic Statistics* 1: 115-126.

Duan, N. et al. 1984. "Choosing between the Sample-Selection Model and the Multi-Part Model." *Journal of Business and Economic Statistics* 2: 283-289.

Fair, R.C. 1977. "A Note on the Computation of the Tobit Estimator." *Econometrica* 45: 1723-1727.

Feldstein, M.S. 1974. "Econometric Studies of Health Economics." Chapter 7 in M.D. Intriligator and D.A. Kendrick, Eds. *Frontiers of Quantitative Economics*, Volume II. Amsterdam: North-Holland.

Fisher, J.A. 1962. "An Analysis of Consumer Goods Expenditures in 1957." *Review of Economics and Statistics* 44: 64-71.

Garwood, F. 1941. "The Application of Maximum Likelihood to Dosage-Mortality Curves." *Biometrika* 32: 46-58.

Goldberger, A.S. and M.L. Lee. 1962. "Toward a Microanalytic Model of the Household Sector." *American Economic Review* 52: 241-251 (Papers and Proceedings).

Goldberger, A.S. 1964. *Econometric Theory*. New York: Wiley.

Griliches, Z. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25: 501-522.

Hald, A. 1949. "Maximum Likelihood Estimation of the Parameters of a Normal Distribution Which is Truncated at a Known Point." *Skandinavisk Aktuarietidskrift* 32: 119-134.

Hausman, J.A. et al. 1984. "Air Pollution and Lost Work." NBER WP 1263.

Haveman, R.H. 1997. *Poverty Policy and Poverty Research*. Madison: University of Wisconsin Press.

Hay, J.W. and R.J. Olsen. 1984. "Let Them Eat Cake: A Note on Comparing Alternative Models of the Demand for Medical Care." *Journal of Business and Economic Statistics* 2: 279-282.

Heckman, J.J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5: 475-492.

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47: 153-161.

Hoch, I. 1958. "Simultaneous Equation Bias in the Context of the Cobb-Douglas Production Function." *Econometrica* 26: 566-578.

Jones, A.M. 2000. "Health Econometrics." Chapter 6 in *Handbook of Health Economics*, Vol. 1, A.J. Culyer and J.P. Newhouse, Eds. Amsterdam: Elsevier.

Laub, P.M. 1972. "Some Aspects of the Aggregation Problem in the Dividend-Earnings Relationship." *Journal of the American Statistical Association* 67: 552-559.

Lin, T.-F. and P. Schmidt. 1984. "A Test of the Tobit Specification against an Alternative Suggested by Cragg." *Review of Economics and Statistics* 66: 174-177.

Maddala, G.S., 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Manning, W.G. et al. 1987. "Monte Carlo Evidence on the Choice between Sample Selection and Two-Part Models." *Journal of Econometrics* 35: 59-82.

Mullahy, J. 1998. "Much Ado about Two: Reconsidering Retransformation and the Two-Part Model in Health Econometrics." *Journal of Health Economics* 17: 247-281.

Mundlak, Y. and I. Hoch. 1965. "Consequences of Alternative Specifications in Estimation of Cobb-Douglas Production Functions." *Econometrica* 33: 814-828.

Newhouse, J.P. 1987. "Health Economics and Econometrics." *American Economic Review Papers and Proceedings* 77: 269-274.

Newhouse, J.P. et al. 1980. "On Having Your Cake and Eating It Too." *Journal of Econometrics* 13: 365-390.

Olsen, R. 1978. "Note on the Uniqueness of the Maximum Likelihood Estimator for the Tobit Model." *Econometrica* 46: 1211-1215.

Orcutt, G.H. 1990. "From Engineering to Microsimulation: An Autobiographical Reflection." *Journal of Economic Behavior and Organization* 14: 5-27.

Peck, S.C. 1974. "Alternative Investment Models for Firms in the Electric Utilities Industry." *Bell Journal of Economics and Management Science* 5: 420-457.

Phelps, C.E. 1973. "Demand for Health Insurance: A Theoretical and Empirical

Investigation." Rand Corp. Report R-1054-OEO, Prepared for the Office of Economic Opportunity.

Powell, J.L. 1984. "Least Absolute Deviations Estimation for the Censored Regression Model." *Journal of Econometrics* 25: 303-325.

Powell, J.L. 1986. "Symmetrically Trimmed Least Squares Estimation for Tobit Models." *Econometrica* 54: 1435-1460.

Renfro, C.G. 2004. "Econometric Software: The First Fifty Years in Perspective." *Journal of Economic and Social Measurement* 29: 9-107.

Rosenzweig, M.R. and T.P. Schultz. 1983. "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight." *Journal of Political Economy* 91: 723-746.

Rosett, R.N. and L.-F. Huang. 1973. "The Effect of Health Insurance on the Demand for Medical Care." *Journal of Political Economy* 81: 281-305.

Sargan, J.D. 1958. "The Estimation of Economic Relationships using Instrumental Variables." *Econometrica* 26: 393-415.

Shiller, R.J. and J. Tobin. 1999. "The ET Interview: Professor James Tobin." *Econometric Theory* 15: 867-900.

Skinner, J. and D. Staiger. 2005. "Technology Adoption from Hybrid Corn to Beta Blockers." NBER WP 11251.

Sloan, F.A. and S. Richupan. 1975. "Short-Run Supply Responses of Professional Nurses: A Microanalysis." *Journal of Human Resources* 10: 241-257.

Suits, D.B. 1957. "Use of Dummy Variables in Regression Equations." *JASA* 52: 548-551.

Tobin, J. 1955. "The Application of Multivariate Probit Analysis to Economic Survey Data." Cowles Foundation Discussion Paper No. 1. New Haven: Cowles Foundation.

Tobin, J. 1956. "Estimation of Relationships for Limited Dependent Variables." Cowles Foundation Discussion Paper No. 3R. New Haven: Cowles Foundation.

Tobin, J. 1958. "Estimation of Relationships for Limited Dependent Variables." *Econometrica*

26: 24-36.

Watts, H.W. 1991. "Distinguished Fellow: An Appreciation of Guy Orcutt." *Journal of Economic Perspectives* 5: 171-179.

Wu, D.-M. 1965. "An Empirical Analysis of Household Durable Goods Expenditures." *Econometrica* 33: 761-780.

Zellner, A. 1990. "Guy H. Orcutt: Contributions to Economic Statistics." *Journal of Economic Behavior and Organization* 14: 43-51.

Figure 1
Reproduction of Figure 1 from Tobin, 1958 (Panel A)
With y-axis Extended (Panel B)

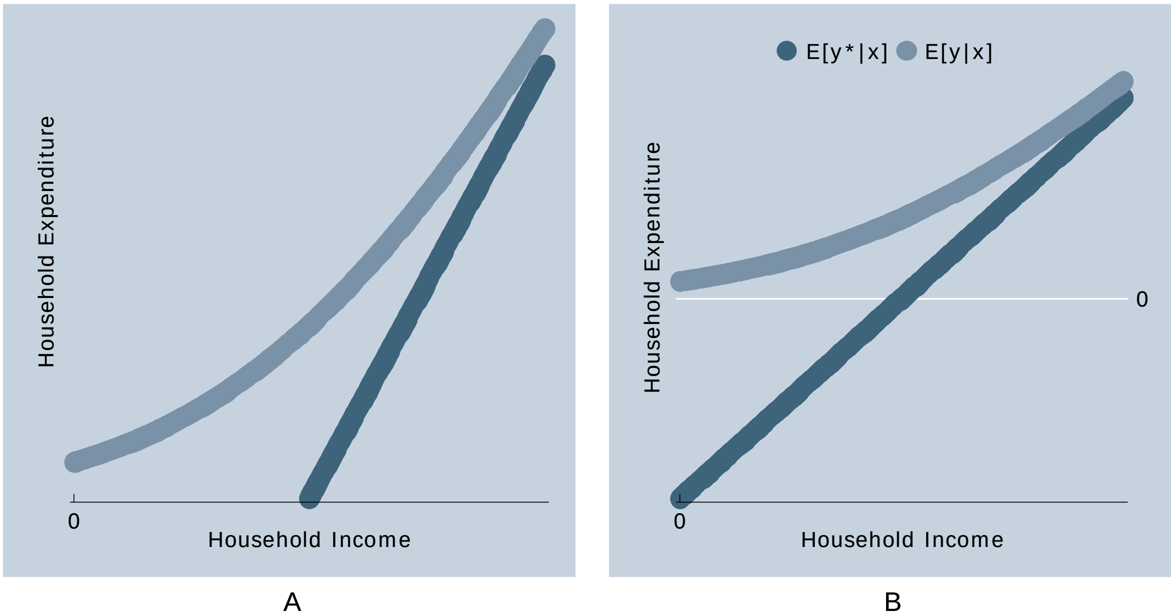

Table 1
Reproduction of TABLE I from Tobin, 1958

TABLE I

SUMS OF SQUARES AND CROSS PRODUCTS

183 limit observations					552 non-limit observations				
	$X_0'=1$	X_1'	X_2'	W'		$X_0=1$	X_1	X_2	W
$X_0'=1$	183				$X_0=1$	552			
X_1'	824	4056			X_1	1976	8060		
X_2'	102.15	552.03	402.3333		X_2	168.06	751.54	255.6740	
W'	0	0	0	0	W	61.449	207.598	20.559	13.113087

Table 2
Reproduction of Excerpt from TABLE II from Tobin, 1957

TABLE II

ITERATIVE ESTIMATION OF PARAMETERS

	a_0	a_1	a_2	a
		.		
		.		
		.		
Final estimates	**1.3392**	**-.2247**	**.0350**	**8.022**
Standard errors	**(.118)**	**(.0295)**	**(.0495)**	**(.252)**