Quantitative Economics 1 (2010), 255–277 1759-7331/20100255

Set identification and sensitivity analysis with Tobin regressors

Victor Chernozhukov
Department of Economics, MIT

Roberto Rigobon
Sloan School of Management, MIT

Thomas M. Stoker
Sloan School of Management, MIT

We give semiparametric identification and estimation results for econometric models with a regressor that is endogenous, bound censored, and selected; it is called a Tobin regressor. First, we show that the true parameter value is set-identified and characterize the identification sets. Second, we propose novel estimation and inference methods for this true value. These estimation and inference methods are of independent interest and apply to any problem possessing the sensitivity structure, where the true parameter value is point-identified conditional on some nuisance parameter values that are set-identified. By fixing the nuisance parameter value in some suitable region, we can proceed with regular point and interval estimation. Then we take the union over nuisance parameter values of the point and interval estimates to form the final set estimates and confidence set estimates. The initial point or interval estimates can be frequentist or Bayesian. The final set estimates are set-consistent for the true parameter value, and confidence set estimates have frequentist validity in the sense of covering this value with at least a prespecified probability in large samples. Our procedure may be viewed as a formalization of the sensitivity analysis in the sense of Leamer (1985). We apply our identification, estimation, and inference procedures to study the effects of changes in housing wealth on household consumption. Our set estimates fall in plausible ranges, significantly above low ordinary least squares estimates and below high instrumental variables estimates that do not account for the Tobin regressor structure.

Keywords. Partial identification, endogenous censoring, sample selection.

JEL classification. C14, C24, C26.

Victor Chernozhukov: vchern@mit.edu
Roberto Rigobon: rigobon@mit.edu
Thomas M. Stoker: tstoker@mit.edu

The authors wish to thank the participants of seminars at UCL and Cambridge University, and CEMMAP conference in June 2007 for helpful comments. We particularly thank Richard Blundell, Andrew Chesher, Whitney Newey, Richard Smith, and Elie Tamer.

DOI: 10.3982/QE1

1. Introduction

In economic surveys, financial variables are often mismeasured in nonrandom ways. The largest values of household income and wealth are often eliminated by top-coding above prespecified threshold values. Income and wealth are also typically reported as nonnegative, which may neglect large transitory income losses, large debts (negative components of wealth), or other aspects that could be modeled as bottom-coding below a prespecified threshold value. In addition to mismeasurement problems related to upper and lower bounds, income and wealth are often missing due to nonresponse.[1]

These measurement problems are particularly onerous when they obscure key features of the economic process under study. For instance, suppose one is studying the impact of liquidity constraints on consumption spending using data from individual households. It is a widespread practice to drop all household observations when there are top-coded income values. However, that practice seemingly eliminates households that are the least affected by liquidity constraints, which would provide the most informative depiction of baseline consumption behavior. Likewise, if one is studying the household demand for a luxury good, the most informative data are from rich households, who, for confidentiality reasons, often will not answer detailed questions about their income and wealth situations.

These problems can be compounded when the observed financial variable is itself an imperfect proxy of the economic concept of interest. For instance, suppose one is studying the impact of the availability of cash on a firm's investment decisions. Only imperfect proxies of "cash availability" are observed in balance sheet data, such as whether the firm has recently issued dividends. The mismeasurement of those proxies is not random; positive dividends indicate positive cash availability, but zero dividends can indicate either mild cash availability or severe cash constraints. Thus, observed dividends represent a censored (bottom-coded at zero) version of the cash availability status of a firm.

The study of mismeasurement due to censoring and selection was initiated by the landmark work of Tobin (1958). In the context of analyzing expenditures on durable goods, Tobin showed how censoring of a dependent variable induced biases and how such bias could be corrected in a parametric framework. This work has stimulated an enormous literature on parametric and semiparametric estimation with censored and selected dependent variables. The term "Tobit model" is common parlance for a model with a censored or truncated dependent variable.

We study the situation where an endogenous regressor is censored or selected. This also causes bias to arise in estimation; bias whose sign and magnitude varies with the mismeasurement process as well as the estimation method used (Rigobon and Stoker

[1]For many surveys, extensive imputations are performed to attempt to "fill in" mismeasured or unrecorded data often in ways that are difficult to understand. For instance, in the U.S. Consumer Expenditure survey, every component of income is top-coded; namely wages, interest, gifts, stock dividends and gains, retirement income, transfers, and bequests, and there is no obvious relation between the top-coding on each component and the top-coding on total income. The CEX makes extensive use of ad hoc multiple imputation methods to fill in unrecorded income values.

(2009)). With reference to the title, we use the term "Tobin regressor" to refer to a regressor that is bound censored, selected, and endogenous.

When the mismeasurement of the regressor is exogenous to the response under study,[2] consistent estimation is possible by using only "complete cases" or estimating with a data sample that drops any observations with a mismeasured regressor.[3] When endogenous regressors are censored or selected, the situation is considerably more complicated. Dropping observations with a mismeasured regressor creates a selected sample for the response under study. Standard instrumental variables methods are not consistent when computed from the full data sample and are also not consistent when computed using the complete cases only.

In this paper, we provide a full identification analysis and estimation solution for situations with Tobin regressors. We show how the true parameter value is set-identified and characterize the identification sets. We propose novel estimation and inference methods for this true value. These estimation and inference methods are of independent interest and apply to any problem with the sensitivity structure, where the true parameter value is point-identified conditional on some nuisance parameter values that are set-identified. Indeed, fixing the nuisance parameter value in some suitable region, we can proceed with regular point and interval estimation. Then we take the union over nuisance parameter values of the point and interval estimates to form the final set estimates and confidence set estimates. The initial point or interval estimates can be frequentist or Bayesian. The final set estimates are set-consistent for the true parameter value, and confidence set estimates have frequentist validity in the sense of covering this value with at least a prespecified probability in large samples. Our procedure may be viewed as a formalization of the sensitivity analysis in the sense of Leamer (1985).

Our approach is related to several contributions in the literature. Without censoring or selection, our framework is in line with work on nonparametric estimation of an endogenous model with nonadditivity, as developed by Altonji and Matzkin (2005), Chesher (2003), Imbens and Newey (2005), and Chernozhukov and Hansen (2005), among others. Our accommodation of endogeneity uses the control function approach, as laid out by Blundell and Powell (2003) and Newey, Powell, and Vella (1999). In terms of dealing with censoring, we follow Powell's (1984) lead in using monotonicity assumptions together with quantile regression methods (see Koenker's (2005) excellent review of quantile regression). Our inference results complement the inferential procedures proposed in Chernozhukov, Hong, and Tamer (2007) and other literature. For examples of models of interval data and related problems, where our inference approach could be valuable, see Ponomareva and Tamer (2009).

There is a great deal of literature on mismeasured data, some focused on regressors. Foremost is Manski and Tamer (2002), who used monotonicity restrictions to propose consistent estimation with interval data. For other contributions in econometrics, see Ai (1997), Chen, Hong, and Tamer (2005), Chen, Hong, and Tarozzi (2008), Liang, Wang, Robins, and Carroll (2004), and Tripathi (2004), among many others, which are primar-

[2]That is, both the correctly measured regressor and the censoring/selection process are exogenous.

[3]The existence of consistent estimates allows for tests of whether bias is evident in estimates computed from the full data sample. See Rigobon and Stoker (2006) for regression tests and Nicoletti and Peracchi (2005) for tests in a generalized method of moments framework.

ily concerned with estimation when data are missing at random. The large literature in statistics on missing data is well surveyed by Little and Rubin (2002), and work focused on mismeasured regressors is surveyed by Little (1992).

The exposition proceeds by introducing our approach within a simple framework, in Section 2. Section 3 gives our general framework and a series of generic results on identification and estimation. Section 4 contains an empirical application, where we show how accommodating censoring and selection gives rise to a much larger estimate of the impact of housing wealth on consumption.

2. A simplified framework and the basic identification approach

2.1 *A linear model with a Tobin regressor*

We introduce the ideas in a greatly simplified setting, where a linear model is the object of estimation. We do this to highlight the main concepts of our approach. Our main results do not rely on a linear model, but are based on a very general (parametric or nonparametric) framework. We spell this out in Section 3.

Consider the estimation of a linear model with a (potentially) endogenous regressor:

$$Y = X^*\alpha + U^*, \tag{2.1}$$

$$X^* = Z'\gamma + V^*, \tag{2.2}$$

$$U^* = \beta V^* + \varepsilon, \tag{2.3}$$

where

$$\varepsilon \text{ is mean (or median or quantile) independent of } (V^*, X^*), \tag{2.4}$$

$$V^* \text{ is median independent of } Z. \tag{2.5}$$

Here Y is the dependent variable, X^* is the uncensored regressor, which is endogenous when $\beta \neq 0$, and Z represents valid instruments (without censoring or selection). We make no further assumption on the distribution of ε or V^*.

The regressor X^* is not observed. Rather, we observe a censored version of X^*:

$$X = I\{R = 1\}I\{X^* > 0\}X^*, \tag{2.6}$$

$$R = \begin{cases} 1 & \text{with probability } 1-\pi, \\ 0 & \text{with probability } \pi, \end{cases} \quad \text{independent of } Z. \tag{2.7}$$

The observed X matches X^* unless one of two sources of censoring arises, in which case $X = 0$. The first source is bound censoring, which occurs when $X^* \leq 0$ or $V^* \leq -Z'\gamma$. The second source is an independent censoring method, which selects $X = 0$ when $R = 0$ (or, equivalently, selects to observe positive X^* when $R = 1$). The variable R is not observed. We sometimes refer to the probability π as the selection probability in the following discourse; it would be more complete to refer to π as "the probability of independent selection for censoring."

In this sense, the observed regressor X is *censored*, *selected*, and *endogenous*, which we refer to as a *Tobin regressor*. There exists an instrument Z for the uncensored regressor X^*, but that instrument will typically be correlated with $X - X^*$. Therefore, Z will not be a valid instrument if X is used in place of X^* in the response equation (2.1).

In our simplified setup, censoring is modeled with the lower bound (bottom-coding) of 0, but top-coding or different bound values are straightforward to incorporate. The selection probability π is taken as constant here, but is allowed to vary with covariates in our general framework. We assume that $P[Z'\gamma > 0] > 0$, which is both convenient and empirically testable.

It is also straightforward to include additional controls in the response equation (2.1). With that in mind, we develop some examples for concreteness.

EXAMPLE 1 (Income and Consumption). Suppose X is income and Y is household consumption expenditure. X is typically endogenous, top-coded, and missing for various households. Bound censoring arises for large income values, and selection refers to missing values, possibly due to households who decline to report their income. For instance, if one is estimating a permanent income model of consumption, then X would be observed permanent income (or wealth). If one is investigating excess sensitivity (or liquidity constraints), then X would be observed current income (and the equation would include lagged consumption). Here, the instruments Z could include unanticipated income shocks, lagged income values, and demographic variables that are not included in the consumption equation. Finally, the same censoring issues can arise in an Engel curve analysis, where Y is the expenditure on some commodity and X is total expenditures on all commodities.

EXAMPLE 2 (Dividends and Firm Investment). Suppose X is declared dividends and Y is investment for individual firms. Here X^* is the level of cash availability (or opposite of cash constraints). Positive dividends X indicate positive cash availability, but zero dividends arises with either mild or severe cash constraints (small or large negative X^*). Here bound censoring would be the primary reason for dividends to mismeasure cash availability. The instruments Z could include exogenous variables that affect the cost of debt, such as foreign exchange fluctuations.

EXAMPLE 3 (Day Care Expenditures and Female Wages). Suppose you are studying the economic situation faced by single mothers, where Y is expenditure on day care and X is the observed wage rate. X is potentially endogenous (work more to pay for higher quality day care), and selection can be due to missing values or directly to the labor participation choice. Here, the instruments Z could include job skills and other factors that affect the labor productivity of single mothers, as well as exogenous household income shocks.

2.2 *Basic identification and estimation ideas*

The strategy for identification is to set the amount of selection first, which allows the rest of the model to be identified. That is, suppose we set a value π^* for $\pi = \Pr[R = 0]$. The following steps give identification:

Step 1. Note that the conditional median curve $Q_{X^*}(\frac{1}{2}|Z) = Z'\gamma$ is partially identified from the estimable curve

$$Q_X\left(\frac{1}{2}(1-\pi^*)+\pi^*\Big|Z\right) = \max[Z'\gamma, 0], \tag{2.8}$$

since $Z'\gamma > 0$ with positive probability.

Step 2. Given γ, we can identify the control function

$$V^* = X^* - Z'\gamma = X - Z'\gamma \tag{2.9}$$

whenever $X > 0$.

Step 3. Given the control function V^*, we can recover the regression function of interest (mean, median, or quantile) for the subpopulation where $X > 0$ and $Z'\gamma > 0$. For instance, if ε is mean independent of (V^*, X^*), we can recover the mean regression

$$E[Y|X, V^*] = X'\alpha + \beta V^*. \tag{2.10}$$

If ε is quantile independent of (V^*, X^*), we can recover the quantile regression

$$Q_Y(\tau|X, V^*) = X'\alpha + \beta V^*. \tag{2.11}$$

Step 4. All of the above parameters depend on the value π^*. We recognize this functional dependence by writing $\alpha(\pi^*)$, $\beta(\pi^*)$, and $\gamma(\pi^*)$ for solutions of Steps 1, 2, and 3. For concreteness, suppose the particular value $\alpha_0 = \alpha(\pi_0)$ is of interest, where π_0 is the true value of $\pi = \Pr[R = 0]$. If we can determine the set $\mathcal{P}$ of all feasible values of π^*, the set

$$\mathcal{A}_0 = \{\alpha(\pi), \pi \in \mathcal{P}\}$$

clearly contains α_0. Likewise, if we denote $\theta(\pi) = \{\alpha(\pi), \beta(\pi), \gamma(\pi)\}$, then $\theta_0 = \theta(\pi_0)$ is contained in the set

$$\Theta_0 = \{\theta(\pi), \pi \in \mathcal{P}\}.$$

Step 5. It remains to characterize the set $\mathcal{P}_0$. In the absence of further information, this set is given by

$$\mathcal{P} = \Big[0, \inf_{z \in \text{support}(Z)} \Pr[D = 1|Z = z]\Big], \tag{2.12}$$

where

$$D \equiv 1 - I\{R = 1\}I\{X^* > 0\}$$

is the index of an observation that is censored. In words, $\mathcal{P}$ is the interval that contains all values between 0 and the smallest probability of censoring in the population.

This outlines the basic identification strategy. It is clear that point identification is achieved if π_0 is a known value.[4] For instance, if there is only bound censoring (no R

[4]If R is observed, the true value π_0 would be identified and we would have point identification of $\alpha(\pi_0)$, $\beta(\pi_0)$, and $\gamma(\pi_0)$.

term in (2.6)), then $\pi = 0$. Then estimation (Step 1) uses median regression to identify the (single) control function needed.

Estimation proceeds by the analogy principle: empirical curves are used in place of the population curves above to form estimators. In Section 3, we justify this for our general framework. We also show how inference is possible by simply constructed confidence sets. That is, suppose α_0 is of interest and the set $\mathcal{P}$ is known. Given π, a standard confidence region for $\alpha(\pi)$ is

$$\mathrm{CR}_{1-a}(\alpha(\pi)) = \big[\widehat{\alpha}(\pi) \pm c_{1-a}\mathrm{s.e.}(\widehat{\alpha}(\pi))\big].$$

We note that a $1-a$ confidence region for $\alpha_0 = \alpha(\pi_0)$ is simply the union of these confidence intervals

$$\bigcup_{\pi\in\mathcal{P}} \mathrm{CR}_{1-a}(\alpha(\pi)). \tag{2.13}$$

Reporting such a confidence region is easy by reporting its largest and smallest elements.

We also discuss adjustments that arise because of the estimation of $\mathcal{P}$, the range of selection probabilities. In particular, a confidence region for $\mathcal{P}$ can be developed, as well as adjustments for the level of significance of the parameter confidence regions such as (2.13).

We now turn to our general framework and main results.

3. Generic set identification and inference

3.1 *The general framework*

The stochastic model we consider is given by the system of quantile equations

$$Y = Q_Y(U|X^*, W, V), \tag{3.1}$$

$$X^* = Q_{X^*}(V|W, Z), \tag{3.2}$$

where Q_Y is the conditional quantile function of Y given X^*, W, and V, and Q_{X^*} is the conditional quantile function of X^* given Z. Here U is Skorohod disturbance such that $U \sim U(0,1)|X^*, W, V$ and V is Skorohod disturbance such that $V \sim U(0,1)|W, Z$. The latent true regressor is X^*, which is endogenous when V enters the first equation nontrivially. We assume that X^* is continuously distributed. Z represents "instruments" for X^* and W represents covariates.

The observed regressor X, the Tobin regressor, is given by the equation

$$X = I\{R = 1\}I\{X^* > 0\}X^*, \tag{3.3}$$

where

$$R = \begin{cases} 1 & \text{with probability } 1-\pi(W), \\ 0 & \text{with probability } \pi(W), \end{cases} \quad \text{conditional on } W, Z, V. \tag{3.4}$$

There are two sources of censoring of X^* to 0: First there is bound censoring, which occurs when $X^* \le 0$; second is independent selection censoring, which occurs when $R = 0$ (and, as before, R is unobserved). As such, X is endogenous, censored, and selected.

The model (3.1)–(3.2) is quite general, encompassing a wide range of nonlinear models with an endogenous regressor. The primary structural restriction is that the system be triangular; that is, V can enter both (3.1) and (3.2), but U does not enter (3.2). The Skorohod disturbances U and V are linked to the response variables Y and X^* via the corresponding conditional quantile functions. We have by definition that

$$U = F_Y(Y|X^*, W, V),$$
$$V = F_{X^*}(X^*|W, Z),$$

where F_Y is the conditional distribution function of Y given X^*, W, and V, and F_{X^*} is the conditional distribution function of X^* given W and Z. The random variables U and V provide an equivalent parameterization to the stochastic model as would additive disturbances or other (more familiar) ways to capture randomness. For example, the linear model (2.1)–(2.2) is written in the form of (3.1)–(3.2) as

$$Y = X^*\alpha + Q_{U^*}(U|V),$$
$$X^* = Z'\gamma + Q_{V^*}(V|Z),$$

where the additive disturbances U^* and V^* have been replaced by U and V through the equivalent quantile representations $U^* = Q_{U^*}(U|V)$ and $V^* = Q_{V^*}(V|Z)$.

The primary restriction of the Tobin regressor structure (3.3)–(3.4) is that selection censoring is independent of bound censoring (conditional on W, Z, and V). We have left the selection probability in the general form $\pi(W)$, which captures many explicit selection models. For instance, we could have selection based on threshold crossing. In that case, the selection mechanism is $R \equiv 1[W'\delta + \eta \geq 0]$, with η an independent disturbance, which implies the selection probability is $\pi(W) = \Pr\{\eta < -W'\delta\}$.

3.2 *Set identification without functional form assumptions*

We now state and prove our first main result. We require the following assumption.

ASSUMPTION 1. *We assume that the systems of equations* (3.1)–(3.4) *and independence assumptions hold as specified above, and that* $v \mapsto Q_{X^*}(v|W, Z)$ *is strictly increasing in* $v \in (0, 1)$ *almost surely.*

Our main identification result follows:

PROPOSITION 1. *The identification regions for* $Q_Y(\cdot|X^*, V, W)$ *and* $F_Y(\cdot|X^*, V, W)$ *on the subregion of the support of* (X^*, V, W) *implied by* $X > 0$ *are given by*

$$\mathcal{Q} = \{Q_Y(\cdot|X, V_\pi, W), \pi \in \mathcal{P}\}$$

and

$$\mathcal{F} = \{F_Y(\cdot|X, V_\pi, W), \pi \in \mathcal{P}\},$$

where when $X > 0$,

$$V_\pi = \frac{F_X(X|Z,W) - \pi(W)}{1 - \pi(W)}, \tag{3.5}$$

or, equivalently, when $X > 0$,

$$V_\pi = \int_0^1 1\big\{Q_X\big(\pi(W) + (1 - \pi(W))v|W, Z\big) \le X\big\}\, dv. \tag{3.6}$$

Finally,

$$\mathcal{P} = \Big\{\pi(\cdot)\ \textit{measurable}: 0 \le \pi(W) \le \inf_{z \in \text{support}\,(Z|W)} F_X(0|W,z)\ \textit{a.s.}\Big\}. \tag{3.7}$$

Proposition 1 says that given the level of the selection probability $\pi(W)$, we can identify the quantile function of Y with respect to X^* by using the (identified) quantile function of Y with respect to the observed Tobin regressor X, where we shift the argument V to V_π of (3.5)–(3.6). The identification region comprises the quantile functions for all possible values of $\pi(\cdot) \in \mathcal{P}$. The proof is constructive, including indicating how the quantiles with respect to X^* and to X are connected. It also makes clear that point identification of the functions is possible where $X > 0$ and $\pi(W)$ is known (or point-identified), including the no selection case with $\pi(W) = 0$.

In empirical applications, one is typically not interested in (nonparametrically) estimating the full conditional distribution of Y given X, V, and W, but rather in more interpretable or parsimonious features. That is, one wants to estimate

$$\theta(\pi) = \theta(Q(\cdot;\pi)), \tag{3.8}$$

a functional of π taking values in the parameter space Θ, where the quantile Q can be either the conditional quantile Q_Y or Q_{X^*}, or, equivalently,

$$\theta(\pi) = \theta^*(F(\cdot;\pi)), \tag{3.9}$$

where the conditional distribution F is either F_Y or F_{X^*}. The functional $\theta(\pi)$ can represent parameters of a model of Q or F, average policy effects, average derivatives, local average responses, and other features (including representing the full original functions Q or F). The following corollary establishes identification of such aspects of interest.

COROLLARY 1. *The identification region for the functional* $\theta(\pi_0)$ *is*

$$\{\theta(\pi), \pi \in \mathcal{P}\}.$$

Given Proposition 1, the proof of this corollary is immediate. We now give the proof of our first main result.

PROOF OF PROPOSITION 1 (Identification). We follow the logic of the identification steps outlined in the previous section. Suppose we first set a value $\pi(W)$ for $\Pr[R = 0|W]$. For

$x > 0$, we have that

$$\begin{aligned}\Pr[X \le x|W, Z] &= \Pr[R = 0|W, Z] + \Pr[R = 1 \text{ and } X^* \le x|W, Z] \\ &= \Pr[R = 0|W, Z] + \Pr[R = 1|W, Z] \cdot \Pr[X^* \le x|W, Z] \\ &= \pi(W) + (1 - \pi(W)) \cdot \Pr[X^* \le x|W, Z].\end{aligned}$$

That is,

$$F_X[x|W, Z] = \pi(W) + (1 - \pi(W))F_{X^*}[x|W, Z].$$

In terms of distributions, whenever $X > 0$,

$$V_\pi = F_{X^*}[X|W, Z] = \frac{F_X(X|Z, W) - \pi(W)}{1 - \pi(W)}.$$

Thus V_π is identified from the knowledge of $F_X(X|Z, W)$ and $\pi(W)$ whenever $X > 0$. In addition

$$X^* = X$$

when $X > 0$.

In terms of quantiles, when $X > 0$,

$$\begin{aligned}Q_{X^*}(V_\pi|W, Z) &= Q_{X^*}\left(\frac{F_X(X|Z, W) - \pi(W)}{1 - \pi(W)}\Big|W, Z\right) \\ &= Q_X\big(\pi(W) + (1 - \pi(W))V_\pi|W, Z\big).\end{aligned} \tag{3.10}$$

This implies that for any $X > 0$,

$$\begin{aligned}V_\pi &= \int_0^1 1\{Q_{X^*}(v|W, Z) \le X^*\}\, dv \\ &= \int_0^1 1\big\{Q_X\big(\pi(W) + (1 - \pi(W))v|W, Z\big) \le X\big\}\, dv.\end{aligned}$$

Thus V_π is identified from the knowledge of $Q_X(\cdot|Z, W)$ whenever $X > 0$.

Inserting

$$X, V_\pi \quad \text{for cases } X > 0$$

into the outcome equation, we have a point identification of the quantile functional

$$Q_Y(\cdot|X, V_\pi, W)$$

over the region implied by the condition $X > 0$. This functional is identifiable from the quantile regression of Y on X, V_π, and W.

Likewise, we have the point identification of the distributional functional

$$F_Y(\cdot|X, V_\pi, W)$$

over the region implied by the condition $X > 0$. This functional is identified either by inverting the quantile functional or by the distributional regression of Y on X, V_π, and W.

Now, since the (point) identification of the functions depends on the value $\pi(W)$, by taking the union over all $\pi(\cdot)$ in the class $\mathcal{P}$ of admissible conditional probability functions $w \mapsto \Pr[R = 0|W = w]$, we have the identified sets for both quantities:

$$\{Q_Y(\cdot|X, V_\pi, W), \pi(\cdot) \in \mathcal{P}\}$$

and

$$\{F_Y(\cdot|X, V_\pi, W), \pi(\cdot) \in \mathcal{P}\}.$$

The quantities above are sets of functions.

It remains to characterize the admissible set $\mathcal{P}$. From the relationship

$$F_X[0|W, Z] = \pi(W) + (1 - \pi(W)) \cdot F_{X^*}[0|W, Z],$$

we have

$$0 \le \pi(W) = \frac{F_X[0|W, Z] - F_{X^*}[0|W, Z]}{1 - F_{X^*}[0|W, Z]} \le F_X(0|W, Z),$$

where the last observation is by the equalities

$$0 = \min_{0 \le x \le F} \left(\frac{F - x}{1 - x} \right) \le \max_{0 \le x \le F} \left(\frac{F - x}{1 - x} \right) = F.$$

Taking the best bound over values of z, we have

$$0 \le \pi(W) \le \inf_{z \in \text{support}\{Z|W\}} F_X(0|W, z).$$

Hence

$$\mathcal{P} = \left\{ \pi(\cdot) \text{ measurable} : 0 \le \pi(W) \le \inf_{z \in \text{support}\{Z|W\}} F_X(0|W, z) \text{ a.s.} \right\},$$

which demonstrates Proposition 1. □

3.3 *Estimation and inference*

Our constructive derivation of identification facilitates a general treatment of estimation. Here we present the general results. In the following section, we discuss some particulars of estimation as well as related results in the literature.

The approach described below applies to any problem with the sensitivity structure, where the parameter of interest $\theta(\pi)$ is indexed by some partially identified nuisance parameter $\pi \in \mathcal{P} \subseteq \Pi$ and a consistent estimator $\widehat{\theta}(\pi)$ is available for each $\pi \in \Pi$. In our case, consider a plug-in estimator

$$\widehat{\theta}(\pi) = \theta(\widehat{Q}(\cdot; \pi)) \quad \text{or} \quad \theta^*(\widehat{F}(\cdot; \pi)),$$

where the true quantile or distribution function is replaced by an estimator. We assume that the model structure is sufficiently regular to support a central limit theorem for $\widehat{\theta}(\pi)$

for each value of π, and that estimates of confidence intervals are available for for each π. This is summarized in the following two assumptions.

ASSUMPTION 2.1. *Suppose*

$$\mathcal{Z}_n(\pi) := A_n(\pi)(\widehat{\theta}(\pi) - \theta(\pi)) \quad \Rightarrow \quad \mathcal{Z}_\infty(\pi) \quad \textit{for each } \pi \in \Pi,$$

where convergence occurs in some metric space $(B, \|\cdot\|_B)$*, where* $A_n(\pi)$ *is a sequence of scalers, possibly data dependent, diverging to infinity.*

ASSUMPTION 2.2. *Let*

$$c(1-a, \pi) := (1-a)\textit{-quantile of } \|\mathcal{Z}_\infty(\pi)\|_B$$

and suppose that the distribution function of $\|\mathcal{Z}_\infty(\pi)\|_B$ *is continuous at* $c(1-a, \pi)$ *for each* $\pi \in \Pi$*. Estimates are available such that* $\widehat{c}(1-a, \pi) \to_p c(1-a, \pi)$ *for each* $\pi \in \Pi$*.*

Note that the assumptions above impose convergence requirements only pointwise in π, which is considerably weaker than imposing convergence requirements uniformly in π.

With these assumptions, we can construct the confidence intervals when the set $\mathcal{P}$ is known; we simply construct the confidence interval given each value of $\pi \in \mathcal{P}$ and take the union.

PROPOSITION 2. *Let*

$$C_{1-a}(\pi) := \left\{\theta \in \Theta : \left\| A_n(\pi)(\widehat{\theta}(\pi) - \theta)\right\|_B \leq \widehat{c}(1-a, \pi)\right\}.$$

Let

$$\mathrm{CR}_{1-a} := \bigcup_{\pi \in \mathcal{P}} C_{1-a}(\pi).$$

Then for any $\pi_0 \in \mathcal{P}$*,*

$$\liminf_{n\to\infty} P\{\theta(\pi_0) \in \mathrm{CR}_{1-a}\} \geq 1 - a.$$

PROOF. We have that

$$\begin{aligned} P\{\theta(\pi_0) \in \mathrm{CR}_{1-a}\} &\geq P\{\theta(\pi_0) \in \mathrm{CR}_{1-a}(\pi_0)\} \\ &= P\left\{\|\mathcal{Z}_n(\pi_0)\|_B \leq \widehat{c}(1-a, \pi_0)\right\} \\ &= P\left\{\|\mathcal{Z}_\infty(\pi_0)\|_B \leq c(1-a, \pi_0)\right\} + o(1) \\ &= 1 - a + o(1), \end{aligned}$$

where $P\{\|\mathcal{Z}_n(\pi_0)\|_B \leq \widehat{c}(1-a, \pi_0)\} = P\{\|\mathcal{Z}_\infty(\pi_0)\|_B \leq c(1-a, \pi_0)\} + o(1)$ follows by a standard argument, using Assumption 2.1 to impose convergence in distribution of random variable $\|\mathcal{Z}_n(\pi_0)\|_B$ and Assumption 2.2 to impose continuity of the map $c \longmapsto P\{\|\mathcal{Z}_\infty(\pi_0)\|_B \leq c\}$ at $c = c(1-a, \pi_0)$, as well as the consistency property $\widehat{c}(1-a, \pi_0) = c(1-a, \pi_0) + o_p(1)$. □

In applications, the set $\mathcal{P}$ is a nuisance parameter that needs to be estimated, and the above confidence intervals need to be adjusted for that estimation. In our case, estimation of $\mathcal{P}$ poses some new challenges. From (3.7), estimation of $\mathcal{P}$ is equivalent to estimation of the minimum of a function

$$\ell(W) = \inf_{z \in \text{support}\{Z|W\}} F_X[0|W, z].$$

Let $\widehat{\ell}(W)$ be a suitable estimate of this function. One example is an analog estimator

$$\widehat{\ell}(W) = \inf_{z \in \text{support}\{Z|W\}} \widehat{F}_X[0|W, z], \tag{3.11}$$

where $(w, z) \mapsto \widehat{F}_X[0|w, z]$ is a suitable estimator of the conditional distribution function of X at 0. We make the following assumption on the estimator.

ASSUMPTION 2.3. *Let $\widehat{\kappa}_n(1-b)$ and let the known scaler $B_n(W)$ be such that*

$$\liminf_{n\to\infty} P\{\ell(W) - \widehat{\ell}(W) \le B_n(W)\widehat{\kappa}_n(1-b)\} \ge 1-b.$$

Conservative forms of confidence regions of this type are available from the literature on simultaneous confidence bands. For instance, technical conditions can be stated under which Assumption 2.3 holds for the analog estimator stated above, for parametric, series, and kernel estimators of $\widehat{F}_X[0|\cdot, \cdot]$, and[5]

$$B_n(W) := \left[\text{s.e.}(\widehat{F}(W, z))\right]_{z=\widehat{z}_0(W)}, \qquad \widehat{\kappa}_n(1-b) = 2\log n,$$

where $\widehat{z}_0(W) = \arg\inf_{z \in \text{support}\{Z|W\}} \widehat{F}_X[0|W, z]$. For more refined constructions of confidence intervals for minimized functions, see, for example, Chernozhukov, Lee, and Rosen (2009).

Let $\pi(\cdot)$ belong to the parameter set Π. From Assumption 2.3, the confidence region for $\pi(\cdot)$ is given by

$$\text{CR}'_{1-b} = \{\pi \in \Pi : \pi(W) - \widehat{l}(W) \le B_n(W)\widehat{\kappa}_n(1-b)\}. \tag{3.12}$$

Under Assumption 2.3 we have that for each $\pi_0 \in \mathcal{P}$

$$\liminf_{n\to\infty} P\{\pi_0 \in \text{CR}'_{1-b}\} \ge 1-b. \tag{3.13}$$

The following proposition covers our case as well as more general cases having the sensitivity structure.

PROPOSITION 3. *Let CR'_{1-b} be a region possessing the property (3.13) and let*

$$\text{CR}_{1-a} := \bigcup_{\pi \in \text{CR}'_{1-b}} C_{1-a}(\pi).$$

[5]In practice, it is advisable to choose the constant $\kappa_n(1-b)$ via bootstrap; what we state here is just a simple example of a construction that can work well in many, but not all, cases.

Then for each $\pi_0 \in \mathcal{P}$,

$$\liminf_{n\to\infty} P\{\theta(\pi_0) \in \mathrm{CR}_{1-a}\} \geq 1 - a - b.$$

PROOF. We have that

$$\begin{aligned} P\{\theta(\pi_0) \in \mathrm{CR}_{1-a}\} &\geq P\{\theta(\pi_0) \in \mathrm{CR}_{1-a}(\pi_0) \cap \pi_0 \in \mathrm{CR}'_{1-b}\} \\ &\geq P\{\theta(\pi_0) \in \mathrm{CR}_{1-a}\} - P\{\pi_0 \notin \mathrm{CR}'_{1-b}\}. \end{aligned}$$

By the proof of Proposition 2, the lower limit of the first term is bounded below by $1 - a$, and by construction, the lower limit of the second term is bounded below by $-b$. □

Thus, we construct parameter confidence intervals by taking the union of confidence intervals for all $\pi(W)$ in the confidence interval CR'_{1-b}, which is an expanded version of the estimate of the parameter set $\mathcal{P}$. More conservative parameter intervals are obtained by choosing a larger confidence set CR'_{1-b}.

This completes our general estimation results.

3.4 *Some estimation specifics*

At this point, it is useful to summarize the steps in estimation and relate our general results to the literature.

The first step is to estimate the allowable values of selection probabilities $\mathcal{P}$ using a boundary estimator such as (3.11). Then we widen the set to accommodate estimation, obtaining the confidence set CR'_{1-b} of (3.12). This set gives the values of selection probabilities $\pi(W)$ to be used in the subsequent estimation steps. That is, if π is a scalar parameter, then we choose values in a grid $\{\pi_k^*, k = 1, \ldots, K\}$ that represents CR'_{1-b}. If $\pi(W)$ is modeled to depend nontrivially on covariates W, then the range of values of $\pi(W)$ is represented; for instance, if $\pi(W)$ depends on a vector of parameters, then a grid over the possible parameter values could be used. We summarize the grid as $\{\pi_k^*(W), k = 1, \ldots, K\}$ in the following discussion.

The second step is to estimate the control function for the Tobin regressor for each value $\pi_k^*(W)$. With an estimator $\widehat{F}_X(\cdot|Z, W)$ of the distribution $F_X(\cdot|Z, W)$, we compute (3.5) for $X > 0$ as

$$\widehat{V}_{\pi,k} = \frac{\widehat{F}_X(X|Z, W) - \pi_k(W)}{1 - \pi_k(W)}. \tag{3.14}$$

Alternatively, with an estimator $\widehat{Q}_X(\cdot|W, Z)$ of the quantile function $Q_X(\cdot|W, Z)$, we compute (3.6) for $X > 0$ as

$$\widehat{V}_{\pi,k} = \int_0^1 1\{\widehat{Q}_X(\pi_k(W) + (1 - \pi_k(W))v|W, Z) \leq X\}\, dv. \tag{3.15}$$

Either approach to estimating the control function can be used. If the model is restricted, then simpler methods of estimating the control function may be applicable. For instance, under the linear model discussed in Section 2, the general formulae (3.14) or

(3.15) can be replaced by the simpler linear version (2.9),

$$\widehat{V}_{\pi,k} = X - Z'\widehat{\gamma}_k,$$

where $\widehat{\gamma}_k$ is from the estimation of (2.8) with $\pi = \pi_k$.

The third step estimates the response model for each control function estimate $\widehat{V}_{\pi,k}$. This may involve estimating the conditional distribution $F_Y[\cdot|X, W, \widehat{V}_{\pi,k}]$ or the conditional quantile function $Q_Y[\cdot|X, W, \widehat{V}_{\pi,k}]$, either under a structural parameterization or using a nonparametric procedure. Alternatively, this could involve estimating the mean regression $E[Y|X, W, \widehat{V}_{\pi,k}]$ or some other interpretable function, such as local policy effects, average derivatives, or local average responses, again with either a parametric model or nonparametric procedure. Using our notation for the functional of interest, this step results in the estimate $\widehat{\theta}_k = \widehat{\theta}(\pi_k(W))$ of the parameter of interest. This step also yields an estimate of the confidence interval $C_k = C_{1-a}(\pi_k(W))$ for each component of θ. For expositional ease, now we suppose in the following analysis that θ is a scalar parameter, so that $\widehat{\theta}_k$ is a scalar and C_k is its estimated confidence interval.

The final step is to assemble the results for all the grid values $\{\pi_k(W), k = 1, \ldots, K\}$ into the final estimates. That is, the set estimate for θ is formed as the interval

$$\Big[\min_k \theta_k, \max_k \theta_k\Big].$$

The confidence interval for θ is given as the union of the confidence intervals over all $\pi_k(W)$ values, namely

$$\mathrm{CR} := \bigcup_{k=1}^{K} C_k.$$

For a vector-valued θ, we would compute set estimates and confidence intervals for each component; for a function-valued θ, we could do the same for functional aspects of interest. This completes the procedure we have justified by our general derivations and results.

It is important to stress that our constructive identification approach relates set identification and inference to results for point identification and inference. We were brief in describing details for our third step—estimation of the response model given $\pi_k(W)$ and estimated control—because most of the necessary properties for estimation and inference are established in the existing literature. The foremost reading here is Imbens and Newey (2005), who discussed nonparametric series estimation in triangular equation systems with estimated regressors, and gave detailed coverage to the properties needed for estimating many common functions such as policy effects and average derivatives. For parametric response models with an estimated regressor, much of the theory is available in Newey, Powell, and Vella (2004), as well as in the classic Newey and McFadden (1994). Turning to censored quantile regression in a parametric framework (such as (2.8) here), see Powell (1984) and Chernozhukov and Hong (2002), and for quantile regression with an estimated regressor, see Koenker and Ma (2006) and Lee (2007). Nonparametric quantile regression with estimated regressors is covered in Chaudhuri

(1991), Chaudhuri, Doksum, and Samarov (1997), Belloni and Chernozhukov (2007), and Lee (2007). Finally, for estimation of the conditional distribution function of the response, see Hall, Wolff, and Yao (1999), He and Shao (2000), Chernozhukov, Fernadez-Val, and Melly (2007) among others.

We now turn to a substantive empirical application to illustrate our method including inference.

4. The marginal propensity to consume out of housing wealth

Recent experience in housing markets has changed the composition of household wealth. In many countries such as the United States, housing prices have increased over a long period, followed by substantial softening. The market for housing debt, especially the risky subprime mortgage market, has experienced liquidity shortages that first resulted in increased volatility in many financial markets and later led to the collapse of credit markets.

In terms of economic growth, much interest centers on the impact of changes in housing wealth on household consumption. For instance, several market observers have argued that housing prices were too high because of a speculative bubble. If so, the recent market correction should have a permanent component, and it is natural to ask what the impact will be on household consumption and aggregate demand. For this, one requires an assessment of the marginal propensity to consume out of housing wealth.

Surprisingly, the literature does not agree on the "right" measure of the marginal propensity of consumption out of housing wealth. Some papers find marginal propensities of 15–20% (e.g., Benjamin, Chinloy, and Donald (2004)), while others report relatively low estimates of 2% in the short run and 9% in the long run (e.g., Carroll, Otsuka, and Slacalek (2006)). Research in this area is very active, but no concensus has arisen about the impacts.[6]

One of the problems of estimation is the fact that variables such as income and housing wealth are endogenous and, in most surveys, also censored. The literature typically drops the censored observations and tries to estimate the relationship by incorporating some nonlinearities. As we have discussed, this is likely to bias the results and, therefore, could play a role in why there is no agreement on a standard set of estimates. We feel that the estimation of the marginal propensity to consume out of housing wealth is a good situation for using the methodologies developed here to shed light on a range of parameter values applicable to the design of policy.

We have data on U.S. household consumption and wealth from Parker (1999). These data are constructed by imputing consumption spending for observed households in the Panel Survey of Income Dynamics (PSID), using the Consumer Expenditure Survey (CEX). Income data are preprocessed—original observations on income are top-coded, but all households with a top-coded income value have been dropped in the construction of our data.

[6]This impact of housing wealth is of primary interest for the world economy, not just the United States. See, for instance, Catte, Girouard, Price, and Andre (2004) and Guiso, Paiella, and Visco (2005) for European estimates in the range of 3.5%. Asian estimates are in a similar range; see Cutler (2004) for estimates of 3.5% for Hong Kong.

We estimate a "permanent income" style of consumption model:

$$\begin{aligned}\ln C_{it} &= \alpha + \beta_{PY} \ln PY_{it} + \beta_H \ln H_{it}/WE_{it} + \beta_W \ln WE_{it} + \beta_Y \ln Y_{it} + U_{it}^* \\ &= \alpha + \beta_{PY} \ln PY_{it} + \beta_H \ln H_{it} + (\beta_W - \beta_H) \ln WE_{it} + \beta_Y \ln Y_{it} + U_{it}^*.\end{aligned} \tag{4.1}$$

Here C_{it} is consumption spending, PY_{it} is a constructed permanent component of income (human capital), H_{it} is housing wealth,[7] WE_{it}, is total wealth, and Y_{it} is current income. Our focus is the elasticity β_H, the propensity to consume out of housing wealth.

We view the composition of wealth between housing and other financial assets as endogenous, being chosen as a function of household circumstances and likely jointly with consumption decisions. Observed log housing wealth takes on many lower bound values for several reasons.[8] There is the endogenous choice of renting versus owning a household's residence. Small house values might be disregarded by the interviewer, in part because of problems in the assessment of property at the moment that the survey is performed. We also could have measurement error in the recording process that produces some implausibly low housing values. One approach to modeling these features would be to specify a fully structural model of the censoring process, which, for instance, would specify the relative benefits of renting versus owning a home, and fully specify the processes of censoring by interviewers. Instead, we follow the literature in estimating the reduced form model (4.1) and we model the endogenous, bounded variable $\ln H_{it}$ as a Tobin regressor. We assume that current income, permanent income, and total wealth are exogenous variables. For instruments, we use lagged values of current income, permanent income, and total wealth.

One implication of the Tobin regressor structure is that all standard ordinary least squares (OLS) and instrumental variables (IV) estimates are not consistent; they include estimates that take into account either censoring or endogeneity, but not both. In Table 1, we present OLS and IV estimates of β_H for various subsamples of the data. The OLS estimates are all low: 2.7% for all data, 3.3% for households with observed lag values, and 5.3% for the complete cases or households with nonzero housing values. The IV estimate for the complete cases is roughly a fourfold increase, namely 21.3%.

Estimation begins with establishing a range for the selection probability by studying the probability of censoring. Once the range is set, the estimates are computed in two steps. First, we compute quantile regressions of the Tobin regressor, using censored LAD as in (2.8), for different values of the selection probability,[9] and then estimate the control function for each probability value. Second, we estimate the model (4.1), including the estimated control function, as in (2.10) or (2.11). Our set estimates are given by

[7]The Parker data have observations on housing values. Following the literature, we assume that the (log of the) ratio of housing wealth to housing value is uncorrelated with value, so that house value is a good proxy of housing wealth. This includes the case where mortgage debt is assumed proportional to house value.

[8]We set a positive lower bound of \$4000 for house values to facilitate taking logarithms. This involved recoding mostly house price values observed as zero, plus a few positive values that we deemed implausibly small.

[9]The selection probability π is taken as a scalar parameter in this application and is constant over observations.

TABLE 1. Basic estimates of housing effects.

	All Households	Households With Observed IV	Nonzero Housing Wealth (Complete Cases)
Sample size	8735	3771	2961
OLS	0.027	0.033	0.053
	(0.004)	(0.006)	(0.010)
IV (TSLS)			0.213
			(0.030)

the range of coefficients obtained for all the different selection probability values. Their confidence intervals are given by the range of upper and lower confidence limits for coefficient estimates. All estimates were computed using Stata 10.0; the code is available from the authors.

To set the range for the selection probability, we estimated the probability that $\ln H$ attains its lower bound given values of PY, Y, and WE, and we found its minimum over the range of our data, as in (3.11). Specifically, we used a probit model, including linear and quadratic terms in all regressors. The minimum values were small, so as a result, we chose a rather low yet conservative value of $\widehat{\ell} = 0.04$.[10] After adjusting by two times standard error times a log factor, the upper bound estimate became 0.08. Thus, we set the range for the selection probability to $\pi \in [0, 0.08]$. Specifically, each estimation step is done over the grid of values $0, 0.008, 0.016, \ldots, 0.08$

For the first estimation step, we implement the censored LAD estimation algorithm of Chernozhukov and Hong (2002). This requires (again) estimating the probability of censoring and then performing standard quantile regression on samples with low censoring probability. These estimates are used to construct the control function V_π for each grid value. For the second estimation step, we computed mean regression, median regression, and quantile regression for the 10% and 90% quantiles.

We present some representative estimates in Figures 1 and 2. Figure 1 displays the different estimates of the housing effect and Figure 2 gives the estimates of all coefficients for median regression. Each figure plots the estimates for each grid value of π, as well as the associated confidence interval obtained by bootstrapping (denoted bci in the legend). The set estimates are the projections of those curves onto the vertical axis.

Overall, there is very little variation in the estimates with π, the selection probability. The housing effect increases over quantiles and there are other level differences not displayed. The bootstrap confidence interval values are fairly wide, reflecting variation from censored LAD estimation (as well as the selection probability) as well as the second step regressions. For what they are worth, Figure 2 includes the confidence interval

[10] In particular, 44 observations (out of 3771, or 1.16%) have an estimated censoring probability less than 0.04; in terms of pointwise 95% confidence intervals, only 3 (out of 3771) fell strictly below 0.04. As such, we feel that 0.04 is a conservative estimate of the selection probability (i.e., the minimum possible censoring probability). Note also that to stay on the conservative side, we used the Cauchit specification for probabilities (as in Koenker and Yoon (2009)); using probit specification, we were obtaining a less conservative estimate.

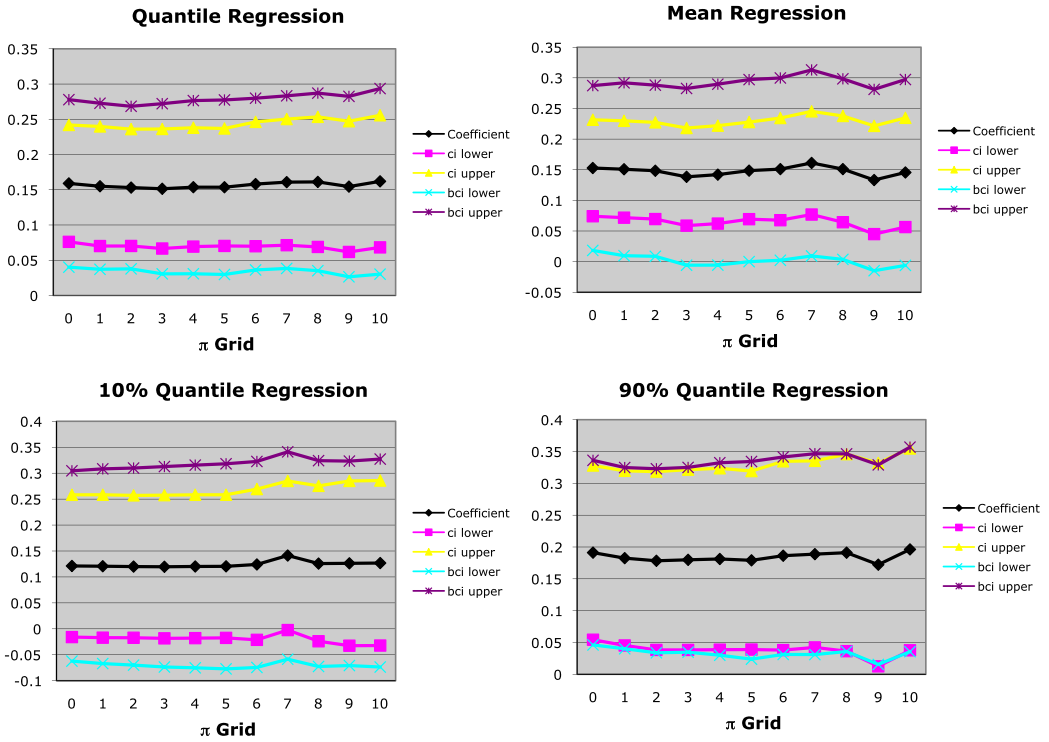

FIGURE 1. Housing coefficient estimates: π grid from 0.0 to 0.08.

estimates from the second step only (denoted ci), so that the difference with the bootstrap intervals gives a sense of the impact of the censored LAD estimation and control function construction.

The results on housing effects are summarized in Table 2. Interval estimates are fairly tight, evidencing the lack of sensitivity with the selection probability. We note that all results are substantially larger than the OLS estimates (2.7–5.3%), which ignore endogeneity. All results are substantially smaller than the IV estimate of 21.3%, which ignores censoring. Relative to the policy debate on the impact of housing wealth, our interval estimates fall in a very plausible range. However, the bootstrap confidence intervals are too wide to discriminate well among these ranges of values. We do see that bootstrap confidence intervals are smaller for median regression than mean regression, and much smaller than for the 10% and 90% quantiles, as expected.

5. SUMMARY AND CONCLUSION

We have presented a general set of identification and estimation results for models with a Tobin regressor—a regressor that is endogenous and mismeasured by bound censoring and (independent) selection. Tobin regressor structure arises very commonly with observations on financial variables, and our results are the first to deal with endogeneity and censoring together. As such, we hope our methods provide a good foundation

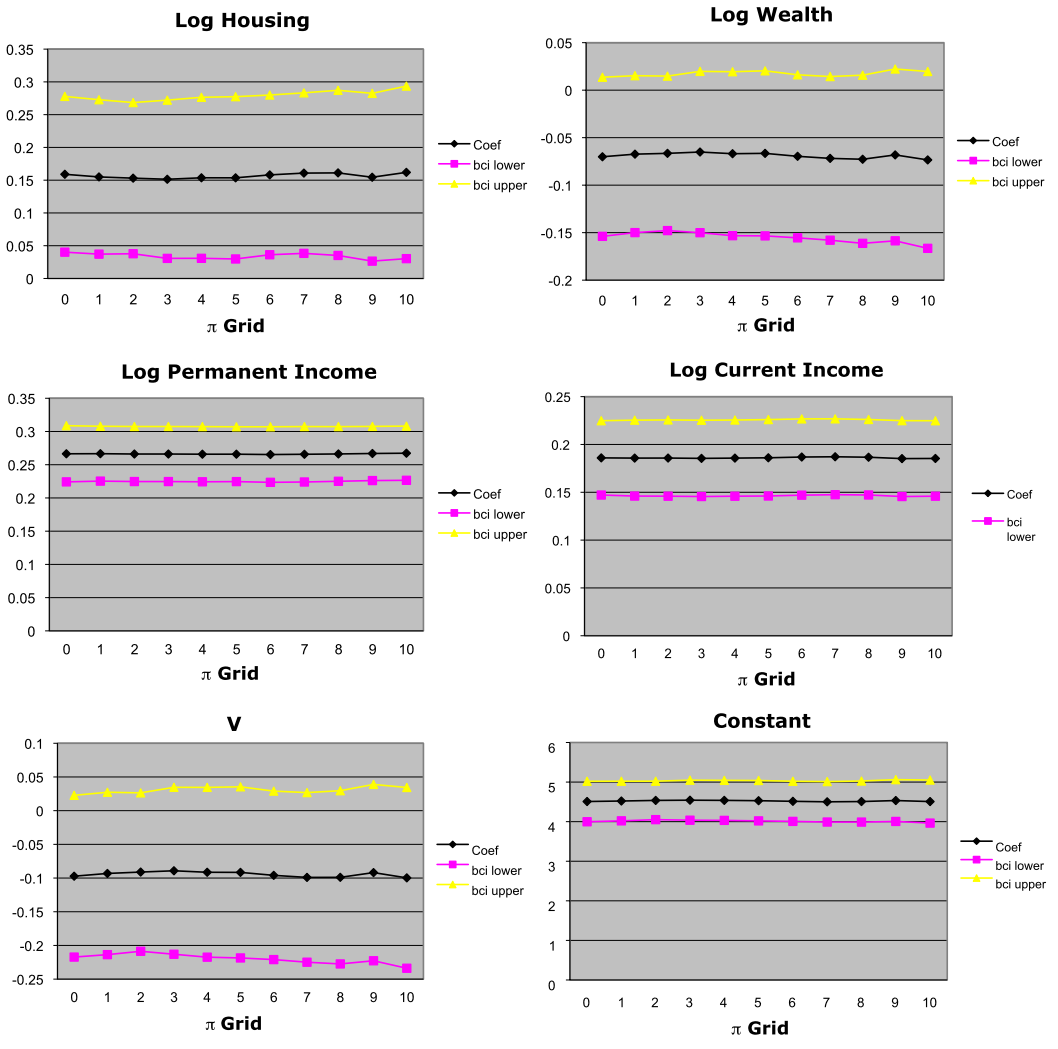

FIGURE 2. Quantile regression estimates: π grid from 0.0 to 0.08.

for understanding how top-coding, bottom-coding, and selection distort the estimated impacts of changes in income, wealth, dividends, and other financial variables.

Our results are restricted to particular forms of censoring. It is not clear how to get around this issue, because endogeneity requires undoing the censoring, and undoing the censoring (seemingly) requires understanding its structure. Here we separate selection and bound censoring with independence, use quantile regression to address bound censoring, and identify parameter sets for the range of possible selection probability values.

We have developed estimation and inference methods for set-identified parameters. In particular, our results apply to any problem with the sensitivity structure, where the parameter value of interest is point-identified conditional on the values of some nui-

TABLE 2. Confidence sets for housing effects.

	Set Estimate	Bootstrap Confidence Region
Housing coefficient β_H		
Mean outcome	[0.133, 0.161]	[−0.015, 0.312]
Median outcome	[0.151, 0.162]	[0.026, 0.293]
90% Quantile	[0.172, 0.196]	[0.015, 0.357]
10% Quantile	[0.120, 0.141]	[−0.077, 0.341]
Selection probability π	[0, 0.04]	[0, 0.08]

sance parameters that are set-identified.[11] The procedure is quite simple: by fixing the nuisance parameter value in some suitable region, one first proceeds with regular point and interval estimation. Then one takes the union over nuisance parameter values of the point and interval estimates to form the final set estimates and confidence set estimates. The final set estimates are set-consistent for the true parameter value, and the confidence set estimates cover this value with at least a prespecified probability in large samples.

REFERENCES

Ai, C. (1997), "An improved estimator for models with randomly missing data." *Nonparametric Statistics*, 7, 331–347. [257]

Altonji, J. and R. Matzkin (2005), "Cross section and panel data estimators for nonseparable models with endogneous regressors." *Econometrica*, 73, 1053–1103. [257]

Belloni, A. and V. Chernozhukov (2007), "Conditional quantile and probability processes under increasing dimension." Preprint, MIT. [270]

Benjamin, J. D., P. Chinloy, and G. D. Donald (2004), "Why do households concentrate their wealth in housing?" *Journal of Real Estate Research*, Oct–Dec, 329–344. [270]

Blundell, R. and J. L. Powell (2003), "Endogeneity in nonparametric and semiparametric regression models." In *Advances in Economics and Econometrics* (M. Dewatripont, L. Hansen, and S. Turnovsky, eds.), Chap. 8, 312–357. Cambridge University Press, Cambridge. [257]

Carroll, C., M. Otsuka, and J. Slacalek (2006), "How large is the housing wealth effect? A new approach." Working paper, NBER. [270]

Catte, P., N. Girouard, R. Price, and C. André (2004), "Housing markets, wealth and the business cycle." Working Paper 394, OECD. [270]

Chaudhuri, P., K. Doksum, and A. Samarov (1997), "On average derivative quantile regression." *The Annals of Statistics*, 25 (2), 715–744. [270]

[11] Several useful examples in Ponomareva and Tamer (2009) are covered by our framework.

Chaudhuri, S. (1991), "Nonparametric estimates of regression quantiles and their local Bahadur representation." *The Annals of Statistics*, 19 (2), 760–777. [270]

Chen, X., H. Hong, and E. Tamer (2005), "Measurement error models with auxiliary data." *Review of Economic Studies*, 72, 343–366. [257]

Chen, X., H. Hong, and A. Tarozzi (2008), "Semiparametric efficiency in GMM models of nonclassical measurement errors, missing data and treatment effects." *The Annals of Statistics*, 36, 808–843. [257]

Chernozhukov, V. and C. Hansen (2005), "An IV model of quantile treatment effects." *Econometrica*, 73, 245–261. [257]

Chernozhukov, V. and H. Hong (2002), "Three-step censored quantile regression and extramarital affairs." *Journal of the American Statistical Association*, 97, 872–882. [269, 272]

Chernozhukov, V., I. Fernadez-Val, and B. Melly (2007), "Inference on Counterfactual Distributions." Preprint, MIT. [270]

Chernozhukov, V., H. Hong, and E. Tamer (2007), "Estimation and confidence regions for parameter sets in econometric models." *Econometrica*, 75, 1243–1284. [257]

Chernozhukov, V., S. Lee, and A. Rosen (2009), "Intersection bounds: Estimation and inference." Working Paper CWP19/09, CEMMAP, Institute for Fiscal Studies, and University College London, Department of Economics. [267]

Chesher, A. (2003), "Identification in nonseparable models." *Econometrica*, 71, 1405–1441. [257]

Cutler, J. (2004), "The relationship between consumption, income and wealth in Hong Kong." Working Paper 1/2004, HKIMR. [270]

Guiso, L., M. Paiella, and I. Visco (2005), "Do capital gains affect consumption? Estimates of wealth effects from Italian households' behavior." Economic Working Paper 555, Bank of Italy. [270]

Hall, P., R. Wolff, and Q. Yao (1999), "Methods for estimating a conditional distribution function." *Journal of the American Statistical Association*, 94, 154–163. [270]

He, X. and Q.-M. Shao (2000), "On parameters of increasing dimensions." *Journal of Multivariate Analysis*, 73 (1), 120–135. [270]

Imbens, G. W. and W. K. Newey (2005), "Identification and estimation of triangular simultaneous equations models without additivity." Working paper, MIT. [257, 269]

Koenker, R. (2005), *Quantile Regression*. Cambridge University Press, Cambridge. [257]

Koenker, R. and L. Ma (2006), "Quantile regression methods for recursive structural equation models." *Journal of Econometrics*, 134, 471–506. [269]

Koenker, R. and J. Yoon (2009), "Parametric links for binary response models." *Journal of Econometrics*, 152, 120–130. [272]

Leamer, E. (1985), "Sensitivity analysis would help." *American Economic Review*, 75 (3), 308–313. [255, 257]

Lee, S. (2007), "Endogeneity in quantile regression models: A control function approach." *Journal of Econometrics*, 141, 1131–1158. [269, 270]

Liang, H., S. Wang, J. M. Robins, and R. J. Carroll (2004), "Estimation in partially linear models with missing covariates." *Journal of the American Statistical Association*, 99, 357–367. [257]

Little, R. J. A. (1992), "Regression with missing X's: A review." *Journal of the American Statistical Association*, 87, 1227–1237. [258]

Little, R. J. A. and D. B. Rubin (2002), *Statistical Analysis With Missing Data*, second edition. John Wiley and Sons, Hoboken, New Jersey. [258]

Manski, C. F. and E. Tamer (2002), "Inference on regressions with interval data on a regressor or outcome." *Econometrica*, 70, 519–546. [257]

Newey, W. and D. McFadden (1994), "Large sample estimation and hypothesis testing." In *Handbook of Econometrics*, Volume IV, 2111–2245, North-Holland, Amsterdam. [269]

Newey, W. K., J. L. Powell, and F. Vella (1999), "Nonparametric estimation of triangular simultaneous equations models." *Econometrica*, 67, 565–603. [257]

Nicoletti, C. and F. Peracchi (2005), "The effects of income imputations on micro analyses: Evidence from the ECHP." Working paper, University of Essex, August. [257]

Parker, J. (1999), "Spendthrift in America? On two decades of decline in the U.S. saving rate?" In *NBER Macroeconomics Annual 1999* (B. Bernanke and J. Rotemberg, eds.), MIT Press, Cambridge. [270]

Ponomareva, M. and E. Tamer (2009), "Misspecification in moment inequality models: Back to moment equalities?" Working paper, Northwestern University. [257, 275]

Powell, J. L. (1984), "Least absolute deviations estimation for the censored regression model." *Journal of Econometrics*, 25, 303–325. [257, 269]

Rigobon, R. and T. M. Stoker (2006), "Testing for bias from censored regressors." Working paper, MIT. [257]

Rigobon, R. and T. M. Stoker (2009), "Bias from censored regressors." *Journal of Business and Economic Statistics*, 27, 340–353. [257]

Tobin, J. (1958), "Estimation of relationships for limited dependent variables." *Econometrica*, 26, 24–36. [256]

Tripathi, G. (2004), "GMM and empirical likelihood with incomplete data." Working paper, University of Wisconsin. [257]

Submitted July, 2009. Final version accepted October, 2010.