Federal Reserve Bank of Chicago

Estimation of Panel Data Regression Models with Two-Sided Censoring or Truncation

Sule Alan, Bo E. Honoré, Luojia Hu, and Søren Leth–Petersen

WP 2011-08

Estimation of Panel Data Regression Models with Two-Sided Censoring or Truncation*

Sule Alan† Bo E. Honoré‡ Luojia Hu § Søren Leth–Petersen ¶

November 14, 2011

Abstract

This paper constructs estimators for panel data regression models with individual specific heterogeneity and two–sided censoring and truncation. Following Powell (1986) the estimation strategy is based on moment conditions constructed from re–censored or re–truncated residuals. While these moment conditions do not identify the parameter of interest, they can be used to motivate objective functions that do. We apply one of the estimators to study the effect of a Danish tax reform on household portfolio choice. The idea behind the estimators can also be used in a cross sectional setting.

Key Words: Panel Data, Censored Regression, Truncated Regression.

JEL Code: C20, C23, C24.

---

*This research was supported by NSF Grant No. SES-0417895 to Princeton University, the Gregory C. Chow Econometric Research Program at Princeton University, and the Danish National Research Foundation, through CAM at the University of Copenhagen (Honoré) and the Danish Social Science Research Council (Leth–Petersen). We thank Christian Scheuer and numerous seminar participants for helpful comments. The opinions expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

†Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge, UK, CB3 9DD. Email: sule.alan@econ.cam.ac.uk.

‡Department of Economics, Princeton University, Princeton, NJ 08544-1021. Email: honore@Princeton.edu.

§Economic Research Department, Federal Reserve Bank of Chicago, 230 S. La Salle Street, Chicago, IL 60604. Email: lhu@frbchi.org.

¶Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K. and SFI, The Danish National Centre for Social Research, Herluf Trolles Gade 11, DK-1052. Email : soren.leth-petersen@econ.ku.dk.

1 Introduction

This paper generalizes a class of estimators for truncated and censored regression models to allow for two–sided truncation or censoring. The class of estimators is based on pairwise comparisons and the proposed generalizations therefore apply to both panel data and cross sectional data.

A leading example of when *two–sided censored regression models* are useful is when the dependent variable is a fraction. For example, Alan and Leth-Petersen (2006) estimate a portfolio share equation where the portfolio shares are between 0 and 1, with a significant number of observations on either of the limits. Other recent applications in economics of regression models with two–sided censoring include Lafontaine (1993), Petersen and Rajan (1994), Petersen and Rajan (1995), Houston and Ryngaert (1997), Fehr, Kirchler, Weichbold, and Gachter (1998), Huang and Hauser (1998), McMillan and Woodruff (1999), de Figueriredo and Tiller (2001), Huang and Hauser (2001), Fenn and Liang (2001), Poterba and Samwick (2002), Nickerson and Silverman (2003), Officer (2004), Charness, Frechette, and Kagel (2004), Andrews, Schank, and Simmons (2005) and Gifford and Bernard (2005). We formulate the two–sided censored regression model as observations of (y, x, L, U) from the model

$$y^* = x'\beta + \varepsilon \tag{1}$$

where y^* is unobserved, but we observe

$$y = \begin{cases} L & \text{if} \quad y^* < L \\ y^* & \text{if} \quad L \leq y^* \leq U \\ U & \text{if} \quad y^* > U \end{cases} \tag{2}$$

and β is the parameter of interest. When y is a share, L and U will typically be 0 and 1, respectively.

We will focus on panel data settings so the observations are indexed by i and t where $i = 1, \ldots, n$ and $t = 1, \ldots T_i$. This allows for unbalanced panels, but we will maintain the restrictive assumption that T_i is exogenous in the sense that it satisfies all the assumptions made on the explanatory variables. In a panel data setting, it is also important to allow for individual specific effects in the errors ε_{it}. We will do this implicitly by making assumptions of the type that ε_{it} is stationary conditional on $(x_{i1} \ldots x_{iT_i})$ or that ε_{it} and ε_{is} are independent and identically distributed conditional on some unobserved component ν_{ist}. These will have the textbook specification $\varepsilon_{it} = v_i + \epsilon_{it}$ ($\{\epsilon_{it}\}$ *i.i.d.*) as a special case. In section 4.2, we discuss how to apply the same ideas to construct estimators of the cross sectional version of the model.

For the *two–sided truncated regression model*, we assume observations of (y, x, L, U) from the

distribution of (y^*, x, L, U) conditional on $L \leq y^* \leq U$. Two–sided truncated regression models are less common than two–sided censored regression models, but they play a role in duration models. Suppose, for example, that one wants to study the effect of early life circumstances on longevity by linking the Social Security Administration's Death Master File to the 1900–1930 U.S. censuses[1]. This will miss a substantial number of deaths: (1) individuals who died under age 65 (since they were less likely to be collecting social security benefits and thus their deaths were less likely to be captured); (2) individuals who died before the year 1965 (the beginning of the computerized Social Security files); and (3) individuals who died after the last year the data is available. This is a case of two-sided truncation because we observe an individual only if the dependent variable, age at death, is greater than 65 years and if the death occurs between 1965 and the last year the data is available. This is essentially the empirical setting in Ferrie and Rolf (2011), although they only consider data from the 1900 census, so right–truncation is unlikely to be an issue in their application.

Honoré (1992) constructed moment conditions for similar panel data models with one–sided truncation or censoring and showed how they can be interpreted as the first–order conditions for a population minimization problem that uniquely identifies the parameter vector, β. This paper generalizes that approach to the case when the truncation or censoring is two–sided. The main contribution of the paper is to show that some of those moment conditions can be turned into a minimization problem that actually uniquely identifies β when there is two–sided censoring or truncation. This is an important step because the moment conditions that we derive do not identify the parameters of the model. This is a generic problem with constructing estimators based on moment conditions. For example, Powell (1986) constructed moment conditions for a related cross sectional truncated and censored regression models based on symmetry of the error distribution. He also pointed out that while these moment conditions did not identify the parameter of interest, minimization of an objective function based on them did lead to identification.

The rest of the paper is organized as follows. Section 2 derives the moment conditions and the associated objective function for models with two–sided censoring and two–sided truncation. Section 3 then discusses how these can be used to estimate the parameters of interest. Generalizations of the two models considered in Section 2 are discussed in Section 4. We present the empirical application in Section 5 and Section 6 concludes.

[1] 1930 is the latest year for which this linkage is feasible.

2 Identification: Moment Conditions and Objective Functions

The challenge in constructing moment conditions in models with censoring and truncation is that one typically starts with assumptions on ε conditional on x in (1). If one had a random sample of (y^*, x) then these assumptions could be used immediately to construct moment conditions. For example if $E[\varepsilon x] = 0$, then one has the moment conditions $E[(y^* - x'\beta)\, x] = 0$. However with truncation or censoring, $y - x'\beta$ will not have the same properties as ε. The idea employed in Powell (1986), Honoré (1992) and Honoré and Powell (1994) is to apply additional censoring and truncation to $y - x'\beta$ in such a way that the the resulting re–censored or re–truncated residual satisfies the conditions assumed on ε. For example, Powell (1986) assumed that ε is symmetric conditional on x in a censored regression model with censoring from below at 0. If that is the case, then $y - x'\beta = \max\{\varepsilon, -x'\beta\}$ will clearly not be symmetric conditional on x, but the re–censored residuals, $\min\{y - x'\beta, x'\beta\}$ will be. This implies moment conditions of the type $E[\min\{y - x'\beta, x'\beta\}\, x] = 0$. Unfortunately, this moment condition will not in general identify β, but Powell (1986) was able to prove that the integral (as a function of b) of $E[\min\{y - x'b, x'b\}\, x]$ is uniquely minimized at $b = \beta$ under appropriate regularity assumptions. Honoré (1992) applied the same insight to a panel setting where the assumption was that ε is stationary conditional on the entire sequence of explanatory variables. Again, censoring or truncation destroys this stationarity, but it can be restored for a pair of residuals by additional censoring. Honoré and Powell (1994) then applied the same idea to any pair of observations in a cross section, and Hu (2002) generalized it to allow for lagged latent dependent variables as covariates.

The contribution of this paper is to generalize the approach in Honoré (1992) to the case with two–sided censoring or truncation. As in Powell (1986), Honoré (1992) and Honoré and Powell (1994), it is straightforward to construct moment conditions based on some re–censored or re–truncated residuals. However, it is not clear that these moment conditions will identify the parameters of interest, and we therefore construct (population) objective functions from these moment conditions, and then explicitly verify that these objective functions are uniquely minimized at the parameter. It is the construction of the objective functions and verifying that they are uniquely minimized at the true parameter value that constitute the methodological contribution of the paper.

The general approach is to start with a comparison of two observations for a given individual in a panel. Based on these observations we will construct re–censored or re–truncated residuals

$e_{its}(y_{it}, x_{it}, x_{is}, L_{it}, L_{is}, U_{it}, U_{is}, b)$ and $e_{ist}(y_{is}, x_{is}, x_{it}, L_{it}, L_{is}, U_{it}, U_{is}, b)$ that have the same properties as ε_{it} and ε_{is} when $b = \beta$. This will then imply that if ε_{it} and ε_{is} are identically distributed conditional on (x_{it}, x_{is}), then

$$E\left[\left(e_{its}(y_{it}, x_{it}, x_{is}, L_{it}, L_{is}, U_{it}, U_{is}, \beta) - e_{ist}(y_{is}, x_{is}, x_{it}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)\right)\middle| x_{it}, x_{is}\right] = 0. \quad (3)$$

This will form the basis for construction of our estimators. Of course, once it has been established that $e_{its}(y_{it}, x_{it}, x_{is}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)$ and $e_{ist}(y_{is}, x_{is}, x_{it}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)$ are identically distributed then for any function, $\xi(\cdot)$, we also have the moment condition

$$E\left[\left(\xi\left(e_{its}(y_{it}, x_{it}, x_{is}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)\right) - \xi\left(e_{ist}(y_{is}, x_{is}, x_{it}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)\right)\right)\middle| x_{it}, x_{is}\right] = 0. \quad (4)$$

provided that the moment exists.

Moreover, if the errors ε_{it} and ε_{is} are also independent conditional on (x_{it}, x_{is}), then so are $e_{its}(y_{it}, x_{it}, x_{is}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)$ and $e_{ist}(y_{is}, x_{is}, x_{it}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)$. This implies that their difference is symmetrically distributed around 0, so for any odd function $\xi(\cdot)$,

$$E\left[\left(\xi\left(e_{its}(y_{it}, x_{it}, x_{is}, L_{it}, L_{is}, U_{it}, U_{is}, \beta) - e_{ist}(y_{is}, x_{is}, x_{it}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)\right)\right)\middle| x_{it}, x_{is}\right] = 0. \quad (5)$$

provided that the moment exists.[2]

In this paper we will focus on (3) and the generalization (5). The reason for this is that in a linear model without censoring or truncation, (3) will correspond to OLS on the differenced data, whereas (5) will also accommodate least absolute deviation estimation on the differenced data as a special case. As already mentioned, the construction of the residuals $e_{its}(y_{it}, x_{it}, x_{is}, L_{it}, L_{is}, U_{it}, U_{is}, \beta)$ is fairly straightforward, and the challenge is to show that although (3) and (5) may not identify β, and unconditional version of them can be integrated to yield a population objective function that is uniquely minimized at $b = \beta$. This will involve the integral of $\xi(\cdot)$, which we will denote by $\Xi(\cdot)$, and in addition to being odd, we will assume that $\xi(\cdot)$ is also increasing, so $\Xi(\cdot)$ is a convex symmetric function. The leading cases are $\Xi(d) = |d|$ and $\Xi(d) = d^2$.

In order to simplify the exposition, we will first develop the case when $L_{it} = 0$ and $U_{it} = 1$. We will then demonstrate that the result can be adapted to the general case.

[2] Of course one could combine the insight in (4) and (5) to get even more general moment conditions. See also the discussion in Arellano and Honoré (2001) and Honoré and Hu (2004).

2.1 Two–Sided Censoring

Consider first the situation with two–sided censoring. Consider an individual, i, in two time periods, t and s, and assume that ε_{it} and ε_{is} are identically distributed. The distribution of $y_{it} - x_{it}'\beta$ will be the same as that of ε_{it} except that the former is censored from below at $-x_{it}'\beta$ and from above at $1 - x_{it}'\beta$. Figure 1 illustrates this. The dotted line depicts the distribution of ε_{it}, while the solid line gives the distribution of $y_{it} - x_{it}'\beta$, which typically has point mass at $-x_{it}'\beta$ and $1 - x_{it}'\beta$ (illustrated by the fatter vertical lines). Since $x_{it}'\beta$ will typically differ from $x_{is}'\beta$, the distributions of $y_{it} - x_{it}'\beta$ and $y_{is} - x_{is}'\beta$ (given (x_{it}, x_{is})) will differ even if $\{\varepsilon_{it}\}$ is stationary (given (x_{it}, x_{is})). However, it is clear that one could obtain identically distributed "residuals" by artificially censoring $y_{it} - x_{it}'\beta$ and $y_{is} - x_{is}'\beta$ from below at $\max\{-x_{it}'\beta, -x_{is}'\beta\}$ and from above at $\min\{1 - x_{it}'\beta, 1 - x_{is}'\beta\}$. See the dashed lines in Figure 1. One can then form moment conditions from the fact that the difference in these "re–censored" residuals will be orthogonal to functions of (x_{it}, x_{is}).[3] Of course, this construction is only useful if $-1 < x_{it}'\beta - x_{is}'\beta < 1$, because otherwise, the supports of $y_{it} - x_{it}'\beta$ and $y_{is} - x_{is}'\beta$ will not overlap.

In order to proceed, we need explicit expressions for the difference in these "re–censored" residuals. Consider first the case when $x_{it}'\beta \geq x_{is}'\beta$. Then the difference in the artificially censored residuals for individual i in periods t and s is

$$\begin{aligned} & \max\left\{y_{it} - x_{it}'\beta, -x_{is}'\beta\right\} - \min\left\{y_{is} - x_{is}'\beta, 1 - x_{it}'\beta\right\} \qquad (6) \\ = \; & \left(\max\left\{y_{it} - \left(x_{it}'\beta - x_{is}'\beta\right), 0\right\} - x_{is}'\beta\right) - \left(\min\left\{y_{is}, 1 - \left(x_{it}'\beta - x_{is}'\beta\right)\right\} - x_{is}'\beta\right) \\ = \; & \max\left\{y_{it} - \left(x_{it}'\beta - x_{is}'\beta\right), 0\right\} - \min\left\{y_{is}, 1 - \left(x_{it}'\beta - x_{is}'\beta\right)\right\}, \end{aligned}$$

and when $x_{it}'\beta \leq x_{is}'\beta$

$$\begin{aligned} & \min\left\{y_{it} - x_{it}'\beta, 1 - x_{is}'\beta\right\} - \max\left\{y_{is} - x_{is}'\beta, -x_{it}'\beta\right\} \qquad (7) \\ = \; & \left(\min\left\{y_{it}, 1 + \left(x_{it}'\beta - x_{is}'\beta\right)\right\} - x_{it}'\beta\right) - \left(\max\left\{y_{is} + \left(x_{it}'\beta - x_{is}'\beta\right), 0\right\} - x_{it}'\beta\right) \\ = \; & \min\left\{y_{it}, 1 + \left(x_{it}'\beta - x_{is}'\beta\right)\right\} - \max\left\{y_{is} + \left(x_{it}'\beta - x_{is}'\beta\right), 0\right\}. \end{aligned}$$

[3] Clearly, one can also use the fact that differences in functions of the re–censored residuals will be orthogonal to functions for the explanatory variables. As discussed in Arellano and Honoré (2001), one can also construct moment conditions based on symmetry under the additional assumption that $(\varepsilon_{i1}, ..., \varepsilon_{iT})$ is exchangeable conditional on $(x_{i1}, ..., x_{iT})$. This is the motivation for the approach in Honoré (1992).

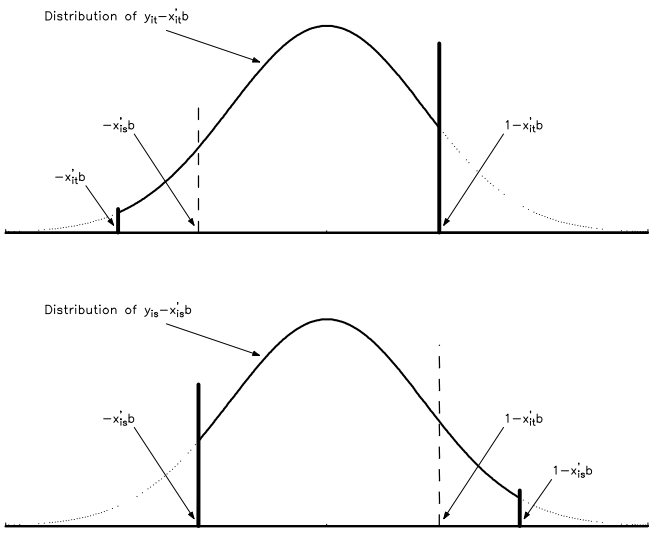

Figure 1: Illustration of Re–Censored Residuals when $x_{it}'\beta > x_{is}'\beta$.

If we define

$$u(y_1, y_2, d) = \begin{cases} 0 & \text{for} \quad d < -1 \\ 1 + d & \text{for} \quad -1 < d < c_1 \\ \min\{1 - y_2, y_1\} & \text{for} \quad c_1 < d < c_2 \\ y_1 - y_2 - d & \text{for} \quad c_2 < d < c_3 \\ \max\{y_1 - 1, -y_2\} & \text{for} \quad c_3 < d < c_4 \\ d - 1 & \text{for} \quad c_4 < d < 1 \\ 0 & \text{for} \quad d > 1 \end{cases} \tag{8}$$

where $c_1 = \min\{-y_2, y_1 - 1\}$, $c_2 = \max\{-y_2, y_1 - 1\}$, $c_3 = \min\{1 - y_2, y_1\}$ and $c_4 = \max\{1 - y_2, y_1\}$ then $u(y_{it}, y_{is}, x_{it}'\beta - x_{is}'\beta)$ will give the difference in the re–censored residuals discussed above (see Appendix 1). Hence the moment conditions are

$$E\left[u\left(y_{it}, y_{is}, x_{it}'\beta - x_{is}'\beta\right) \middle| x_{it}, x_{is}\right] = 0,$$

which implies the unconditional moments

$$E\left[u\left(y_{it}, y_{is}, (x_{it} - x_{is})'\beta\right)(x_{it} - x_{is})\right] = 0. \tag{9}$$

Panel A of Figures 2–4 depict the contribution to the moment condition function $u(y_1, y_2, d)$ for pairs of uncensored observations (Figure 2), pairs with one censored and one uncensored

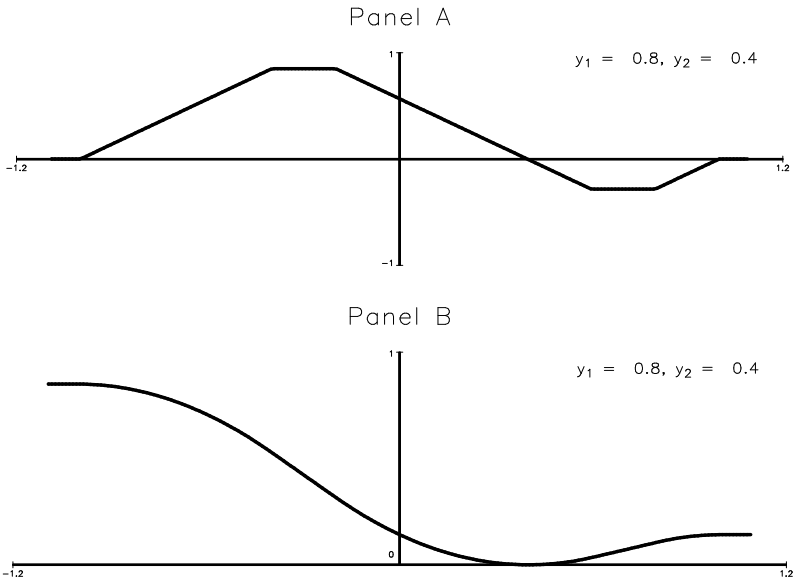

Figure 2: The Functions $u(y_1, y_2, \cdot)$ and $U(y_1, y_2, \cdot)$. Neither Observation Censored.

observation (Figure 3) and pairs with one observation censored from above and one from below (Figure 4).

Although the true parameter value will satisfy (9), it is not in general the unique solution to the moment condition. This is illustrated in Panel A of Figure 5. It considers the case when $y_{i1}^* \sim N(0.5, 1)$ and $y_{i2}^* \sim N(0.4, 1)$ and both are censored from below at 0 and from above at 1. Note that this is the data generation process that one would get with $\beta = 0.1$, $x_{i1} = 1$ and $x_{i2} = 0$ for all i.

It is clear from Figure 5 that the moment condition $E[u(y_{it}, y_{is}, x_{it}'b - x_{is}'b) | x_{it}, x_{is}] = 0$ does not identify the parameter, β, of the model. The most obvious reason is that, as mentioned, only observations for which $-1 < (x_{i1} - x_{i2})'b < 1$ will contribute. In this case $x_{i1} - x_{i2} = 1$ for all observations, so the moment condition is automatically satisfied when $|b| > 1$.

Following Powell (1986), we attempt to overcome the non–identification based on the moment condition by turning it into the first order condition for a minimization problem. It is easy to see that (9) is (half of minus) the first order condition for minimizing

$$E\left[U\left(y_{it}, y_{is}, (x_{it} - x_{is})'b\right)\right], \tag{10}$$

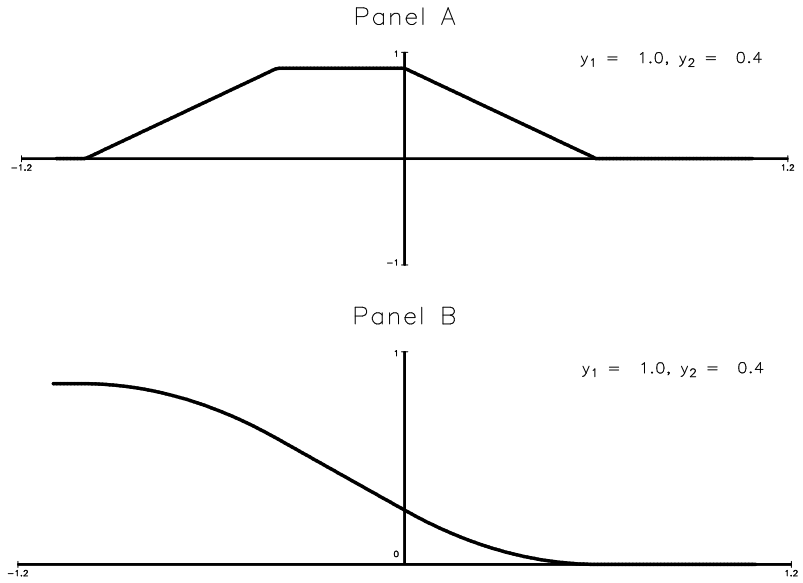

Figure 3: The Functions $u(y_1, y_2, \cdot)$ and $U(y_1, y_2, \cdot)$. One Observation Censored.

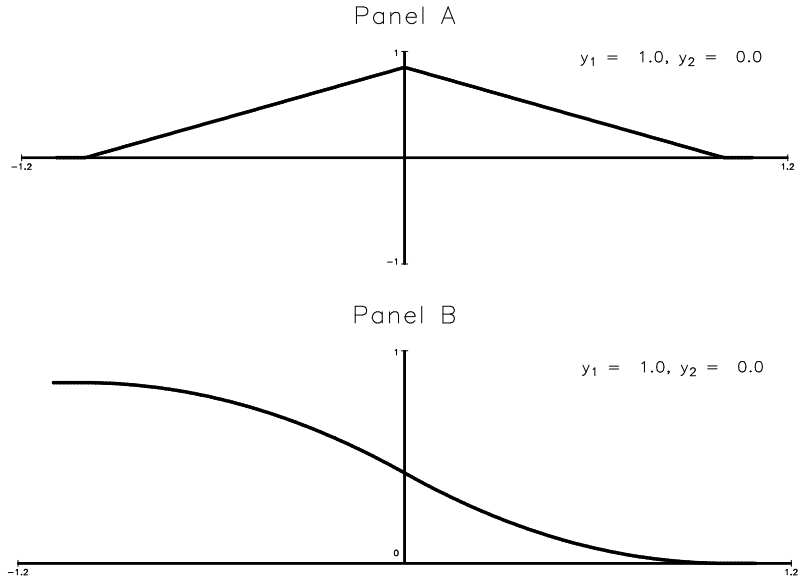

Figure 4: The Functions $u(y_1, y_2, \cdot)$ and $U(y_1, y_2, \cdot)$. Both Observations Censored.

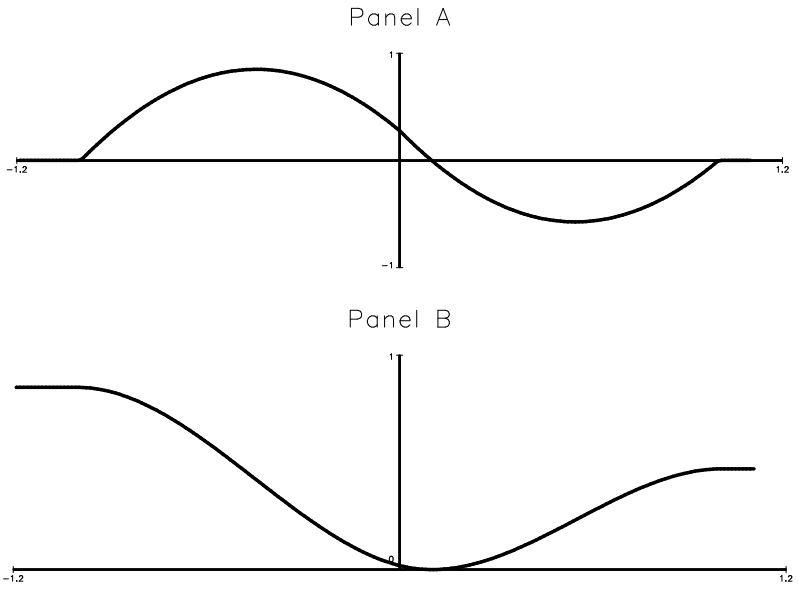

Figure 5: The Functions $E\left[u\left(y_1, y_2, \cdot\right)\right]$ and $E\left[U\left(y_1, y_2, \cdot\right)\right]$.

where

$$U\left(y_1, y_2, d\right) = \begin{cases} 1 + 2c_1 + c_1^2 - 2c_3c_1 + 2c_3c_2 + \left(y_1 - y_2 - c_2\right)^2 & \text{for} \quad d < -1 \\ -2d - d^2 + 2c_1 + c_1^2 - 2c_3c_1 + 2c_3c_2 + \left(y_1 - y_2 - c_2\right)^2 & \text{for} \quad -1 < d < c_1 \\ -2c_3d + 2c_3c_2 + \left(y_1 - y_2 - c_2\right)^2 & \text{for} \quad c_1 < d < c_2 \\ \left(y_1 - y_2 - d\right)^2 & \text{for} \quad c_2 < d < c_3 \\ -2c_2d + 2c_2c_3 + \left(y_1 - y_2 - c_3\right)^2 & \text{for} \quad c_3 < d < c_4 \\ -d^2 + 2d + c_4^2 - 2c_4 - 2c_2c_4 + 2c_2c_3 + \left(y_1 - y_2 - c_3\right)^2 & \text{for} \quad c_4 < d < 1 \\ 1 + c_4^2 - 2c_4 - 2c_2c_4 + 2c_2c_3 + \left(y_1 - y_2 - c_3\right)^2 & \text{for} \quad d > 1 \end{cases}.$$

Panel B of Figures 2–4 depict the contribution to the objective function $U\left(y_1, y_2, d\right)$ for pairs of uncensored observations (Figure 2), pairs with one censored and one uncensored observation (Figure 3) and pairs with one observation censored from above and one from below (Figure 4). Like the estimator for the panel data one–sided censored regression model developed in Honoré (1992), the objective function is piecewise quadratic or linear. However, surprisingly, going from one–sided to two–sided censoring ruins the convexity of the objective function, and the shape of the function is more similar to the objective function in Powell (1986) although that estimator was developed for a cross section model with symmetrically distributed errors.

It is clear from the discussion above that the true β will solve the first order condition for

minimizing the population objective function in (10). However, since U is constant, linear, quadratic and convex, and quadratic and concave over different regions, it is not at all obvious that β will be the unique solution to these first order conditions. The key step for establishing identification of β is therefore to establish that the function in (10) is minimized at β. We establish this in Appendix 1 (Section 7.1), and the result is illustrated in Panel B of Figure 5.

As mentioned, it is also possible to construct moment conditions based on (5). Let Ξ be convex and symmetric, and let $\xi(\cdot) = \Xi'(\cdot)$ (when it exists). When ε_{it} and ε_{is} are independent and identically distributed, we also have the moment conditions

$$E\left[\xi\left(u\left(y_{it}, y_{is}, x_{it}'\beta - x_{is}'\beta\right)\right) \middle| x_{it}, x_{is}\right] = 0$$

which imply the unconditional moments

$$E\left[\xi\left(u\left(y_{it}, y_{is}, (x_{it} - x_{is})'\beta\right)\right)(x_{it} - x_{is})\right] = 0 \tag{11}$$

where

$$\xi(u(y_1, y_2, d)) = \begin{cases} 0 & \text{for} \quad d < -1 \\ \xi(1+d) & \text{for} \quad -1 < d < \min\{-y_2, y_1 - 1\} = c_1 \\ \xi(c_3) & \text{for} \quad c_1 < d < \max\{-y_2, y_1 - 1\} = c_2 \\ \xi(y_1 - y_2 - d) & \text{for} \quad c_2 < d < \min\{1 - y_2, y_1\} = c_3 \\ \xi(c_2) & \text{for} \quad c_3 < d < \max\{1 - y_2, y_1\} = c_4 \\ \xi(d-1) & \text{for} \quad c_4 < d < 1 \\ 0 & \text{for} \quad d > 1 \end{cases}$$

Except for a multiplicative constant, (11) is the first order condition for minimizing

$$E\left[U_\Xi\left(y_{it}, y_{is}, (x_{it} - x_{is})'b\right)\right] \tag{12}$$

where U_Ξ is found by integrating $\xi(u(y_1, y_2, d))$ over each of the regions and insisting on continuity at the boundaries between the regions:

$$U_\Xi(y_1, y_2, d) = \begin{cases} \Xi(0) - \Xi(1+c_1) - \xi(c_3)c_1 + \xi(c_3)c_2 + \Xi(y_1 - y_2 - c_2) & \text{for} \quad d < -1 \\ \Xi(1+d) - \Xi(1+c_1) - \xi(c_3)c_1 + \xi(c_3)c_2 + \Xi(y_1 - y_2 - c_2) & \text{for} \quad -1 < d < c_1 \\ -\xi(c_3)d + \xi(c_3)c_2 + \Xi(y_1 - y_2 - c_2) & \text{for} \quad c_1 < d < c_2 \\ \Xi(y_1 - y_2 - d) & \text{for} \quad c_2 < d < c_3 \\ -\xi(c_2)d + \xi(c_2)c_3 + \Xi(y_1 - y_2 - c_3) & \text{for} \quad c_3 < d < c_4 \\ \Xi(d-1) - \Xi(c_4 - 1) - \xi(c_2)c_4 + \xi(c_2)c_3 + \Xi(y_1 - y_2 - c_3) & \text{for} \quad c_4 < d < 1 \\ \Xi(0) - \Xi(c_4 - 1) - \xi(c_2)c_4 + \xi(c_2)c_3 + \Xi(y_1 - y_2 - c_3) & \text{for} \quad d > 1 \end{cases}.$$

The appendix establishes that (12) is uniquely minimized at $b = \beta$.

2.2 Two–sided Truncation

Mimicing the argument for the model with two–sided censoring, it is clear that if ε_{it} and ε_{is} are independent and identically distributed conditional on (x_{it}, x_{is}), then the observed errors, $y_{it} - x'_{it}\beta$, will be i.i.d. except that the sampling scheme will have truncated them at different points, $-x'_{it}\beta$ and $-x'_{is}\beta$ from below and $1 - x'_{it}\beta$ and $1 - x'_{is}\beta$ from above. We can then construct identically distributed residuals by artificially truncating $y_{it} - x'_{it}\beta$ at $-x'_{is}\beta$ from below and at $1 - x'_{is}\beta$ from above (and similarly for $y_{is} - x'_{is}\beta$). This yields many moment conditions, including

$$\begin{aligned} E\left[r\left(y_{it}, y_{is}, (x_{it} - x_{is})'\beta\right)\right] &\equiv E\left[\left(y_{it} - y_{is} - (x_{it} - x_{is})'\beta\right) \cdot 1\left\{-x'_{is}\beta \leq y_{it} - x'_{it}\beta \leq 1 - x'_{is}\beta\right\}\right. \\ &\quad \left. \cdot\, 1\left\{-x'_{it}\beta \leq y_{is} - x'_{is}\beta \leq 1 - x'_{it}\beta\right\}(x_{it} - x_{is})\right] \qquad (13) \\ &= 0 \end{aligned}$$

It is an easy exercise to see that except for a multiplicative constant (13) is the first order condition for minimizing

$$E\left[R\left(y_{it}, y_{is}, (x_{it} - x_{is})'b\right)\right] \qquad (14)$$

where $R(y_1, y_2, d)$ is defined by

$$\begin{cases} \frac{1}{2}\left(y_1 - y_2 - \max\left\{y_1 - 1, -y_2\right\}\right)^2 & \text{if} \quad \max\left\{y_1 - 1, -y_2\right\} > d \\ \frac{1}{2}\left(y_1 - y_2 - d\right)^2 & \text{if} \quad \max\left\{y_1 - 1, -y_2\right\} \leq d \leq \min\left\{y_1, 1 - y_2\right\} \\ \frac{1}{2}\left(y_1 - y_2 - \min\left\{y_1, 1 - y_2\right\}\right)^2 & \text{if} \quad d > \min\left\{y_1, 1 - y_2\right\} \end{cases}$$

Figure 6 depicts the function R and its derivative, whereas Figure 7 shows their expectation when $y_1^* \sim N(0.5, 1)$ and $y_2^* \sim N(0.4, 1)$ and both are truncated from below at 0 and from above at 1.

More generally, again let Ξ be convex and symmetric, and let $\xi(\cdot) = \Xi'(\cdot)$ (when it exists). Then

$$E\left[\xi\left(y_{it} - y_{is} - (x_{it} - x_{is})'\beta\right) \cdot 1\left\{-x'_{is}\beta \leq y_{it} - x'_{it}\beta \leq 1 - x'_{is}\beta\right\}\right.$$
$$\left. \cdot\, 1\left\{-x'_{it}\beta \leq y_{is} - x'_{is}\beta \leq 1 - x'_{it}\beta\right\}(x_{it} - x_{is})\right] = 0$$

or

$$E\left[\xi\left(y_{it} - y_{is} - (x_{it} - x_{is})'\beta\right) \cdot 1\left\{0 \leq y_{it} - (x_{it} - x_{is})'\beta \leq 1\right\}\right.$$
$$\left. \cdot\, 1\left\{0 \leq y_{is} + (x_{it} - x_{is})'\beta \leq 1\right\}(x_{it} - x_{is})\right] = 0$$

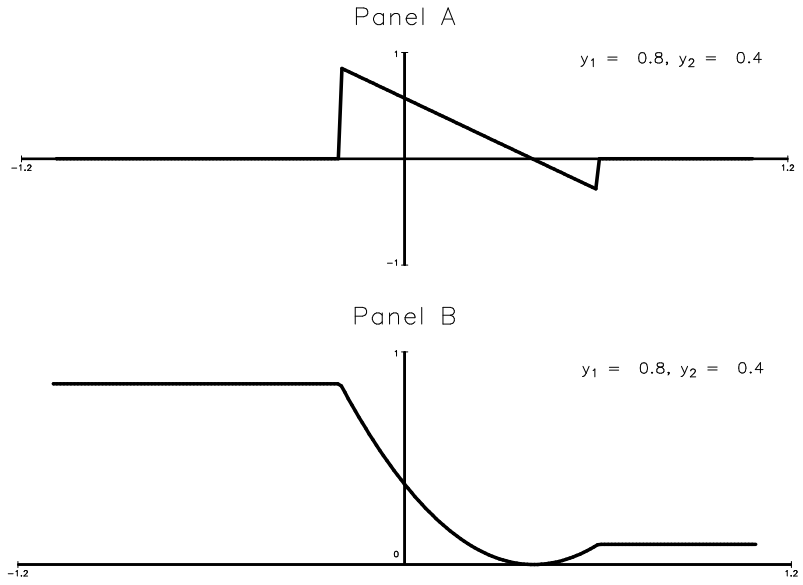

Figure 6: The Functions $r(y_1, y_2, \cdot)$ and $R(y_1, y_2, \cdot)$.

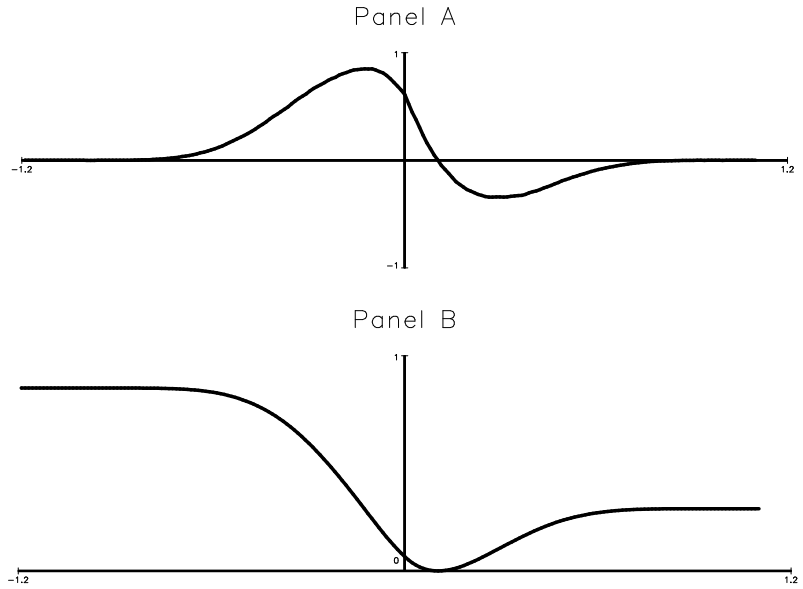

Figure 7: The Functions $E[r(y_1, y_2, \cdot)]$ and $E[R(y_1, y_2, \cdot)]$.

or with $\delta = (x_{it} - x_{is})' \beta$,

$$\begin{aligned}
0 &= E\left[\xi\left(y_{it} - y_{is} - \delta\right) \cdot 1\left\{0 \leq y_{it} - \delta \leq 1\right\} \cdot 1\left\{0 \leq y_{is} + \delta \leq 1\right\}\left(x_{it} - x_{is}\right)\right] \\
&= E\left[\xi\left(y_{it} - y_{is} - \delta\right) \cdot 1\left\{0 \leq y_{it} - \delta\right\} \cdot 1\left\{y_{it} - \delta \leq 1\right\} \cdot 1\left\{0 \leq y_{is} + \delta\right\} \cdot 1\left\{y_{is} + \delta \leq 1\right\}\left(x_{it} - x_{is}\right)\right] \\
&= E\left[\xi\left(y_{it} - y_{is} - \delta\right) \cdot 1\left\{\delta \leq y_{it}\right\} \cdot 1\left\{y_{it} - 1 \leq \delta\right\} \cdot 1\left\{-y_{is} \leq \delta\right\} \cdot 1\left\{\delta \leq 1 - y_{is}\right\}\left(x_{it} - x_{is}\right)\right] \\
&= E\left[\xi\left(y_{it} - y_{is} - \delta\right) \cdot 1\left\{\max\left\{y_{it} - 1, -y_{is}\right\} \leq \delta \leq \min\left\{y_{it}, 1 - y_{is}\right\}\right\}\left(x_{it} - x_{is}\right)\right]
\end{aligned}$$

This is minus the derivative of $E\left[R_{\Xi}\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)\right]$ evaluated at $b = \beta$, where

$$R_{\Xi}\left(y_1, y_2, d\right) = \begin{cases} \Xi\left(\max\left\{y_1 - 1, -y_2\right\}\right) & \text{for} \quad d < \max\left\{y_1 - 1, -y_2\right\} \\ \Xi\left(y_1 - y_2 - d\right) & \text{for} \quad \max\left\{y_1 - 1, -y_2\right\} \leq d \leq \min\left\{y_1, 1 - y_2\right\} \\ \Xi\left(\min\left\{y_1, 1 - y_2\right\}\right) & \text{for} \quad d > \min\left\{y_1, 1 - y_2\right\} \end{cases}$$

As was the case for the censored model, the argument above only establishes that the true β will solve the first order condition for minimizing $E\left[R_{\Xi}\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)\right]$. On the other hand, it is clear that without additional strong assumptions, β will not be the unique minimizer. The reason is that we know that, in general, the truncated regression model will not be identified with exponentially distributed errors. As a result, assumptions must be added that rule out the exponential distribution. In Appendix 1 (section 7.2), we show that β is the unique minimizer of $E\left[R\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)\right]$, and more generally of $E\left[R_{\Xi}\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)\right]$, provided that the errors have a log–concave probability distribution.

3 Estimation

The arguments leading to identification of the parameters of interest, β, above were based on comparing two observations for the same individual and we showed that β could be expressed as the unique minimizer of an expectation of the form $E\left[Q\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)\right]$ for some function Q. This suggests estimating β by minimizing a sample analog of this such as

$$\widehat{\beta} = \arg\min_{b} \frac{1}{n} \sum_{i=1}^{n} \binom{T_i}{2}^{-1} \sum_{1 \leq s < t \leq T_i} Q\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)$$

In this aggregation, observations get different weight depending on the number of observations for a given individual. Alternatively, one could also use objective functions of the type

$$\arg\min_{b} \frac{1}{n} \sum_{i=1}^{n} \sum_{1 \leq s < t \leq T_i} w_{ist} Q\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)$$

where the w_{ist}'s are exogenous weights. In particular, with unbalanced panels, one might want w_{ist} to depend on T_i, the number of time periods for individual i. For example, one can think of the usual fixed effects estimator in a linear regression model as minimizing

$$\sum_{i=1}^{n} \sum_{1 \leq s < t \leq T_i} \frac{1}{T_i} \left(y_{is} - y_{it} - (x_{is} - x_{it})' b \right)^2,$$

so a simple natural choice for w_{ist} could be $\frac{1}{T_i}$.

For two–sided censoring, the resulting estimator is[4]

$$\widehat{\beta} = \arg\min_b \sum_{i=1}^{n} \sum_{1 \leq s < t \leq T_i} w_{ist} U \left(y_{it}, y_{is}, (x_{it} - x_{is})' b \right)$$

or more generally

$$\widehat{\beta} = \arg\min_b \sum_{i=1}^{n} \sum_{1 \leq s < t \leq T_i} w_{ist} U_{\Xi} \left(y_{it}, y_{is}, (x_{it} - x_{is})' b \right) \tag{15}$$

where the functions U and U_{Ξ} are defined in Section 2.1. Standard arguments[5] yield

Theorem 1 *Consider a random sample of size n from* $\left(T_i, \{y_{it}, x_{it}\}_{t=1}^{T_i} \right)$. *If*

1. $$y_{it} = \begin{cases} 0 & \text{if} \quad x_{it}'\beta + \varepsilon_{it} < 0 \\ x_{it}'\beta + \varepsilon_{it} & \text{if} \quad 0 \leq x_{it}'\beta + \varepsilon_{it} \leq 1 \\ 1 & \text{if} \quad x_{it}'\beta + \varepsilon_{it} > 1 \end{cases},$$

2. $(\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{iT_i})$ *is continuously distributed conditional on* $\left(T_i, \{x_{it}\}_{t=1}^{T_i} \right)$ *with a density that is continuous and positive everywhere,*

3. *the sequence* $\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{iT_i}$ *is stationary conditional on* $\left(T_i, \{x_{it}\}_{t=1}^{T_i} \right)$, *and for any* $s, t \leq T_i$ *there exists a random variable,* ν_i^{st}, *such that* ε_{is} *and* ε_{it} *are independent conditional on* ν_i^{st},

4. *the matrix*

$$E\left[(x_{is} - x_{it}) (x_{is} - x_{it})' \middle| -1 < (x_{is} - x_{it})' \beta < 1 \right]$$

has full rank

[4] A Stata-program for calculating this estimator can be found at `www.princeton.edu/~honore/stata`.

[5] Consistency follows from Theorem 4.1.1 of Amemiya (1985) and asymptotoc normality from, for example, Theorem 3.3 of Pakes and Pollard (1989).

then

$$\sqrt{n}\left(\widehat{\beta}-\beta\right)\xrightarrow{d}N\left(0,\Gamma^{-1}V\Gamma^{-1}\right)$$

where $\widehat{\beta}$ *is defined in (15) and*

$$\Gamma=\left.\frac{dE\left[\sum_{s<t}w_{i,t-s}\xi\left(u\left(y_{it},y_{is},\left(x_{it}-x_{is}\right)'b\right)\right)\left(x_{it}-x_{is}\right)\right]}{db'}\right|_{b=\beta}$$

and

$$V=E\left[v_iv_i'\right]$$

with

$$v_i=\sum_{s<t}w_{i,t-s}\xi\left(u\left(y_{it},y_{is},\left(x_{it}-x_{is}\right)'\beta\right)\right)\left(x_{it}-x_{is}\right).$$

These assumptions are consistent with a "fixed effects" model in which $\varepsilon_{it}=\alpha_i+\widetilde{\varepsilon}_{it}$ with α_i unrestricted and the sequence $\{\widetilde{\varepsilon}_{it}\}_{t=1}^{T_i}$ independent and identically distributed. The assumptions also allow for some correlation in the $\widetilde{\varepsilon}_{it}$'s. For example if $(\widetilde{\varepsilon}_{is},\widetilde{\varepsilon}_{it})$ is bivariate normal with the same variance, then they can be written as $\widetilde{\varepsilon}_{is}=Z_{is}+Q_i$ and $\widetilde{\varepsilon}_{it}=Z_{it}+Q_i$ where Z_{it}, Z_{is} and Q_i are independent normals. So $\widetilde{\varepsilon}_{is}$ and $\widetilde{\varepsilon}_{it}$ are independent conditional on Q_i. The assumption that $(\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{iT_i})$ is continuously distributed is necessary if one wants to allow Ξ to be non–differentiable(i.e., $\Xi(d)=|d|$). Without it, the derivative in the expression for Γ might not exist.

When $\Xi(d)=d^2$, condition 3 can be reduced to assuming that the sequence $\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{iT_i}$ is stationary conditional on $\left(T_i,\{x_{it}\}_{t=1}^{T_i}\right)$, and condition 2 is not necessary. In that case the terms in the asymptotic variance reduce to

$$\begin{aligned}\Gamma=E\Bigg[&\sum_{s<t}w_{i,t-s}1\left\{-1<\left(x_{is}-x_{it}\right)'\beta<1\right\}\\&\Big(1\left\{-1<\left(x_{is}-x_{it}\right)'\beta<y_{is}-1\right\}-1\left\{0<\left(x_{is}-x_{it}\right)'\beta<y_{is}\right\}\\&-1\left\{-y_{it}<\left(x_{is}-x_{it}\right)'\beta<0\right\}+1\left\{1-y_{it}<\left(x_{is}-x_{it}\right)'\beta<1\right\}\Big)\left(x_{is}-x_{it}\right)\left(x_{is}-x_{it}\right)'\Bigg]\end{aligned}$$

and

$$V=E\left[v_iv_i'\right]$$

with

$$v_i=\sum_{s<t}w_{i,t-s}u\left(y_{is},\left(x_{is}-x_{it}\right)'\beta\right)\left(x_{is}-x_{it}\right)$$

Following standard arguments, these are consistently estimated by

$$\widehat{\Gamma} = \frac{1}{n}\sum_{i=1}^{n}\left[\sum_{s<t} w_{i,t-s} 1\left\{-1 < (x_{is}-x_{it})'\widehat{\beta} < 1\right\}\right.$$
$$\left(1\left\{-1 < (x_{is}-x_{it})'\widehat{\beta} < y_{is}-1\right\} - 1\left\{0 < (x_{is}-x_{it})'\widehat{\beta} < y_{is}\right\}\right.$$
$$\left.\left. - 1\left\{-y_{it} < (x_{is}-x_{it})'\widehat{\beta} < 0\right\} + 1\left\{1-y_{it} < (x_{is}-x_{it})'\widehat{\beta} < 1\right\}\right)(x_{is}-x_{it})(x_{is}-x_{it})'\right]$$

and

$$\widehat{V} = \frac{1}{n}\sum_{i=1}^{n}\widehat{v}_i\widehat{v}_i'$$

with

$$\widehat{v}_i = \sum_{s<t} w_{i,t-s}\ u\left(y_{is}, (x_{is}-x_{it})'\widehat{\beta}\right)(x_{is}-x_{it})$$

For two–sided truncation the resulting estimator is

$$\widehat{\beta} = \arg\min_b \sum_{i=1}^{n}\sum_{1\le s<t\le T_i} w_{ist} R\left(y_{it}, y_{is}, (x_{it}-x_{is})'b\right) \tag{16}$$

or more generally

$$\widehat{\beta} = \arg\min_b \sum_{i=1}^{n}\sum_{1\le s<t\le T_i} w_{ist} R_\Xi\left(y_{it}, y_{is}, (x_{it}-x_{is})'b\right) \tag{17}$$

where the functions R and R_Ξ are defined in Section 2.2.

We have

Theorem 2 *Consider a random sample of size n from* $\left(T_i, \{y_{it}, x_{it}\}_{t=1}^{T_i}\right)$. *If*

1. y_{it} *is drawn from the distribution of* $x_{it}'\beta + \varepsilon_{it}$ *conditional on* $0 \le x_{it}'\beta + \varepsilon_{it} \le 1$

2. $(\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{iT_i})$ *is continuously distributed conditional on* $\left(T_i, \{x_{it}\}_{t=1}^{T_i}\right)$ *with a density that is continuous and positive everywhere*

3. *the sequence* $\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{iT_i}$ *is stationary conditional on* $\left(T_i, \{x_{it}\}_{t=1}^{T_i}\right)$, *and for any* $s, t \le T_i$ *there exists a random variable,*ν_i^{st}, *such that* ε_{is} *and* ε_{it} *are independent and have log–concave density conditional on* ν_i^{st},

4. *the matrix*

$$E\left[(x_{is}-x_{it})(x_{is}-x_{it})'\middle| -1 < (x_{is}-x_{it})'\beta < 1\right]$$

has full rank

then

$$\sqrt{n}\left(\widehat{\beta}-\beta\right) \xrightarrow{d} N\left(0,\Gamma^{-1}V\Gamma^{-1}\right)$$

where $\widehat{\beta}$ *is defined in (17) and*

$$\Gamma = \left.\frac{dE\left[\sum_{s<t} w_{i,t-s}\xi\left(r\left(y_{it},y_{is},\left(x_{it}-x_{is}\right)'b\right)\right)\left(x_{it}-x_{is}\right)\right]}{db'}\right|_{b=\beta}$$

and

$$V = E\left[v_i v_i'\right]$$

with

$$v_i = \sum_{s<t} w_{i,t-s}\xi\left(r\left(y_{it},y_{is},\left(x_{it}-x_{is}\right)'\beta\right)\right)\left(x_{it}-x_{is}\right)$$

4 Extensions

4.1 Mixed Censored/Truncation

Having considered models with two–sided censoring or truncation, it is natural to also consider a regression model with censoring from one side and truncation from the other:

$$\begin{aligned} y_{it}^* &= x_{it}'\beta+\varepsilon_{it} \\ \left(y_{it},x_{it}\right) &= \left(\min\left\{y_{it}^*,U_{it}\right\},x_{it}\right) \text{ conditional on } L_{it}\leq y_{it} \end{aligned} \tag{18}$$

To simplify the notation, we again focus on the case where $L_{it}=0$ and $U_{it}=1$. In this case the moment condition based on the same logic as above is

$$\begin{aligned} 0 &= E\left[1\left\{y_{it}-x_{it}'\beta>-x_{is}'\beta\right\}1\left\{y_{is}-x_{is}'\beta>-x_{it}'\beta\right\}\right. \\ &\quad \left.\left(\min\left\{y_{it}-x_{it}'\beta,1-x_{is}'\beta\right\}-\min\left\{y_{is}-x_{is}'\beta,1-x_{it}'\beta\right\}\right)\middle|\, x_{it},x_{is}\right] \\ &= E\left[1\left\{y_{it}>\left(x_{it}-x_{is}\right)'\beta\right\}1\left\{y_{is}>-\left(x_{it}-x_{is}\right)'\beta\right\}\right. \\ &\quad \left.\left(\min\left\{y_{it},1+\left(x_{it}-x_{is}\right)'\beta\right\}-\min\left\{y_{is},1-\left(x_{it}-x_{is}\right)'\beta\right\}\right)-\left(x_{it}-x_{is}\right)'\beta\middle|\, x_{it},x_{is}\right] \\ &= E\left[t\left(y_{it},y_{is}\left(x_{it}-x_{is}\right)'\beta\right)\middle|\, x_{it},x_{is}\right]. \end{aligned}$$

where we have assumed that ε_{it} and ε_{is} are independent and identically distributed conditional on (x_{it},x_{is}).

This implies the unconditional moment condition

$$E\left[t\left(y_{it},y_{is}\left(x_{it}-x_{is}\right)'\beta\right)\left(x_{it}-x_{is}\right)\right]=0$$

where

$$t(y_1, y_2, d) = \begin{cases} 0 & \text{for} \quad d < \widetilde{c}_1 \\ 1 - y_2 & \text{for} \quad \widetilde{c}_1 < d < c_2 \\ (y_1 - y_2 - d) & \text{for} \quad c_2 < d < c_3 \\ y_1 - 1 & \text{for} \quad c_3 < d < \widetilde{c}_4 \\ 0 & \text{for} \quad \widetilde{c}_4 < d \end{cases}$$

where $\widetilde{c}_1 = -y_2$, $c_2 = \max\{y_1 - 1, -y_2\}$, $c_3 = \min\{1 - y_2, y_1\}$ and $\widetilde{c}_4 = y_1$. Note that c_2 and c_3 are defined as before, but $\widetilde{c}_1$ and $\widetilde{c}_4$ differ from c_1 and c_4.

Let

$$T(y_1, y_2, d) = \begin{cases} -2(1 - y_2)\widetilde{c}_1 + 2(1 - y_2)c_2 + (y_1 - y_2 - c_2)^2 & \text{for} \quad d < \widetilde{c}_1 \\ -2(1 - y_2)d + 2(1 - y_2)c_2 + (y_1 - y_2 - c_2)^2 & \text{for} \quad \widetilde{c}_1 < d < c_2 \\ (y_1 - y_2 - d)^2 & \text{for} \quad c_2 < d < c_3 \\ -2(y_1 - 1)d + 2(y_1 - 1)c_3 + (y_1 - y_2 - c_3)^2 & \text{for} \quad c_3 < d < \widetilde{c}_4 \\ -2(y_1 - 1)\widetilde{c}_4 + 2(y_1 - 1)c_3 + (y_1 - y_2 - c_3)^2 & \text{for} \quad \widetilde{c}_4 < d \end{cases}$$

We then define the estimator of β by minimizing

$$\sum_i \sum_{t<s} w_{its} T\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)$$

The function T and its derivative are depicted in Figures 8 and 9 for a pair of uncensored observations and for a pair with one censored and one uncensored observation, respectively.

Figure 10 shows the moment condition and the expected value of the objective function when $y_1^* \sim N(0.5, 1)$ and $y_2^* \sim N(0.4, 1)$ and both are truncated from below at 0 and censored from above at 1.

As before, we also have

$$E\left[\xi\left(t\left(y_{it}, y_{is}(x_{it} - x_{is})'\beta\right)\right)(x_{it} - x_{is})\right] = 0$$

where Ξ is convex and symmetric, and $\xi(\cdot) = \Xi'(\cdot)$ (when it exists)

Let

$$T_\Xi(y_1, y_2, d) = \begin{cases} -\xi(1 - y_2)\widetilde{c}_1 + \xi(1 - y_2)c_2 + \Xi(y_1 - y_2 - c_2) & \text{for} \quad d < \widetilde{c}_1 \\ -\xi(1 - y_2)d + \xi(1 - y_2)c_2 + \Xi(y_1 - y_2 - c_2) & \text{for} \quad \widetilde{c}_1 < d < c_2 \\ \Xi(y_1 - y_2 - d) & \text{for} \quad c_2 < d < c_3 \\ -\xi(y_1 - 1)d + \xi(y_1 - 1)c_3 + \Xi(y_1 - y_2 - c_3) & \text{for} \quad c_3 < d < \widetilde{c}_4 \\ -\xi(y_1 - 1)\widetilde{c}_4 + \xi(y_1 - 1)c_3 + \Xi(y_1 - y_2 - c_3) & \text{for} \quad \widetilde{c}_4 < d \end{cases}$$

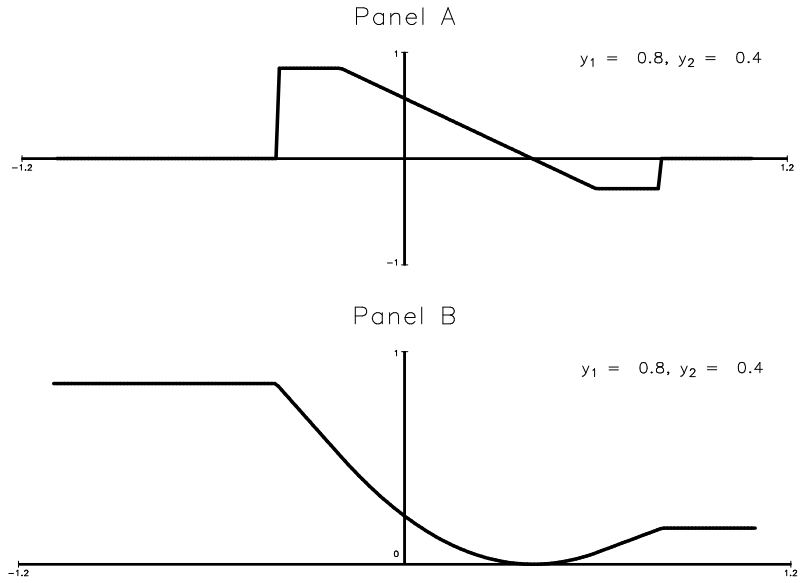

Figure 8: The Functions $t(y_1, y_2, \cdot)$ and $T(y_1, y_2, \cdot)$. Neither Observation Censored.

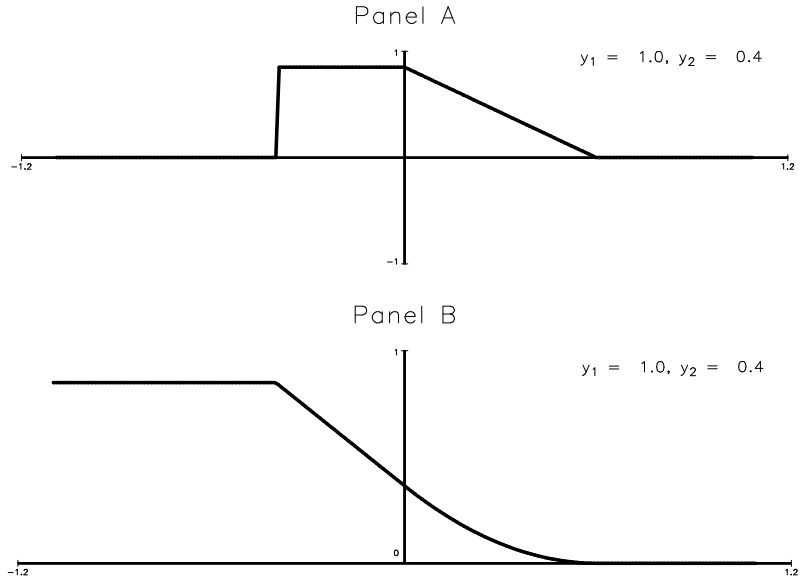

Figure 9: The Functions $t(y_1, y_2, \cdot)$ and $T(y_1, y_2, \cdot)$. One observation censored.

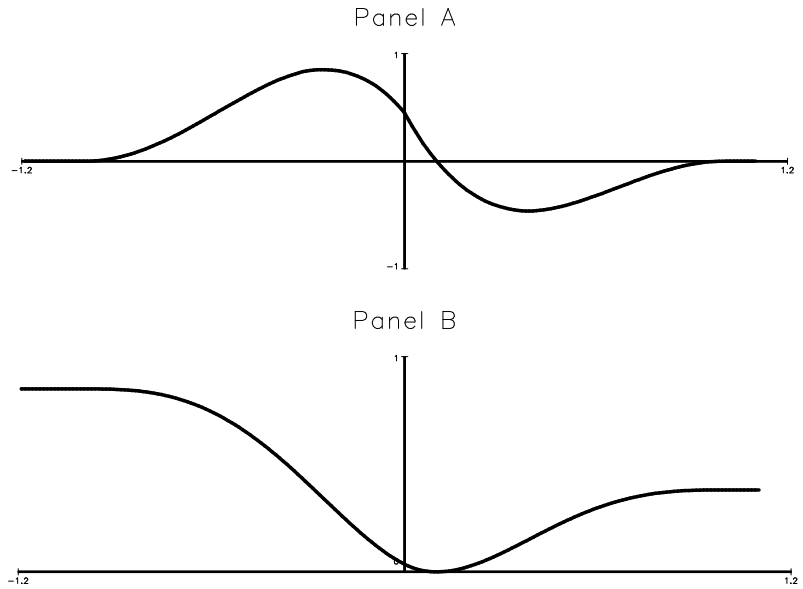

Figure 10: The Functions $E\left[t\left(y_1, y_2, \cdot\right)\right]$ and $E\left[T\left(y_1, y_2, \cdot\right)\right]$.

We then define the estimator of β by minimizing

$$\sum_i \sum_{t<s} w_{its} T_\Xi \left(y_{it}, y_{is}, \left(x_{it} - x_{is}\right)' b\right).$$

4.2 Pairwise Difference Versions

If we can estimate a panel data with 2 observations per unit, then we can apply the same idea to any two observations in a cross section, treating the constant in the cross sectional model as an individual–specific effect. This idea was explicitly used in Honoré and Powell (1994) to construct estimators for the parameters of cross sectional (one–sided) censored and truncated regression models based on the panel data estimator in Honoré (1992). Among others, this also characterizes the relationship between the estimators in Manski (1987) and Han (1987) and between the estimators in Kyriazidou (1997) and Powell (1987). The same idea can be applied to the models with two–sided censoring and truncation considered here, and the asymptotic properties follow from the arguments used in Honoré and Powell (1994).

4.3 General Censoring Points.

It is easy to generalize the results above to the case where the truncation and censoring points are not all 0 and 1. For example, consider the model with two–sided censoring

$$y_{it} = \begin{cases} L_i & \text{if} \quad y_{it}^* < L_i \\ x_{it}'\beta + \varepsilon_{it} & \text{if} \quad L_i \leq y_{it}^* \leq U_i \\ U_i & \text{if} \quad y_{it}^* > U_i \end{cases}$$

then

$$\frac{y_{it} - L_i}{U_i - L_i} = \begin{cases} 0 & \text{if} \quad y_{it}^* < L_i \\ \left(\frac{x_{it}}{U_i - L_i}\right)'\beta + \frac{\varepsilon_{it} - L_i}{U_i - L_i} & \text{if} \quad L_i \leq y_{it}^* \leq U_i \\ 1 & \text{if} \quad y_{it}^* > U_i \end{cases}$$

and we could estimate β by

$$\widehat{\beta} = \arg\min_b \sum_i \sum_{t<s} w_{its} U_\Xi \left(\frac{y_{it} - L_i}{U_i - L_i}, \frac{y_{is} - L_i}{U_i - L_i}, \left(\frac{x_{it} - x_{is}}{U_i - L_i}\right)' b\right) \tag{19}$$

This simple approach does not work when the censoring points are time–varying, because then $\frac{\varepsilon_{it} - L_{it}}{U_{it} - L_{it}}$ is not stationary.

In order to proceed, we need explicit expressions for the difference in these "re–censored" residuals. We first note that only pairs for which the support of the re–censored residuals overlap can play a role in the moment conditions leading to the objective function. These pairs are characterized by

$$L_{it} - U_{is} < x_{it}'\beta - x_{is}'\beta < U_{it} - L_{is}$$

and for such pairs, the difference in the artificially censored residuals for individual i in periods t and s is

$$\begin{aligned} & \text{mami}\left\{L_{is} - x_{is}'\beta, y_{it} - x_{it}'\beta, U_{is} - x_{is}'\beta\right\} - \text{mami}\left\{L_{it} - x_{it}'\beta, y_{is} - x_{is}'\beta, U_{it} - x_{it}'\beta\right\} \\ = \quad & \text{mami}\left\{L_{is}, y_{it} - \left(x_{it}'\beta - x_{is}'\beta\right), U_{is}\right\} - \text{mami}\left\{L_{it}, y_{is} + \left(x_{it}'\beta - x_{is}'\beta\right), U_{it}\right\} + \left(x_{it}'\beta - x_{is}'\beta\right) \end{aligned}$$

where we use the notation $\text{mami}\{a, x, b\} = \max\{a, \min\{x, b\}\}$.

If we define

$$k(L, U, y, d) = \begin{cases} U & \text{for} \quad d < y - U \\ y - d & y - U < d < y - L \\ L & \text{for} \quad d > y - L \end{cases}$$

and

$$u(y_1, y_2, d, L_1, L_2, U_1, U_2) = 1\{L_1 - U_2 < d < U_1 - L_2\}\left(k(L_2, U_2, y_1, d) - k(L_1, U_1, y_2, -d) + d\right)$$

then

$$E\left[u\left(y_{it},y_{is},\left(x_{it}'\beta-x_{is}'\beta\right),L_{it},L_{is},U_{it},U_{is}\right)\middle|\left(x_{it},x_{is}\right)\right]=0$$

and hence

$$E\left[u\left(y_{it},y_{is},\left(x_{it}'\beta-x_{is}'\beta\right),L_{it},L_{is},U_{it},U_{is}\right)\left(x_{it}-x_{is}\right)\right]=0.$$

These will be the moment conditions that lead to the estimator in this case.

Also define

$$K\left(L,U,y,d\right)=\begin{cases}2yU-2dU-U^{2} & \text{for} \quad d<y-U\\ \left(y-d\right)^{2} & y-U<d<y-L\\ 2yL-2dL-L^{2} & \text{for} \quad d>y-L\end{cases}$$

$$S\left(y_{1},y_{2},d,L_{1},L_{2},U_{1},U_{2}\right)=K\left(L_{2},U_{2},y_{1},d\right)+K\left(L_{1},U_{1},y_{2},-d\right)-d^{2}$$

and

$$V\left(y_{1},y_{2},d,L_{1},L_{2},U_{1},U_{2}\right)=\begin{cases}S\left(y_{1},y_{2},L_{1}-U_{2},L_{1},L_{2},U_{1},U_{2}\right) & \text{for} \quad d<L_{1}-U_{2}\\ S\left(y_{1},y_{2},d,L_{1},L_{2},U_{1},U_{2}\right) & \text{for} \quad L_{1}-U_{2}<d<U_{1}-L_{2}\\ S\left(y_{1},y_{2},U_{1}-L_{2},L_{1},L_{2},U_{1},U_{2}\right) & \text{for} \quad d>U_{1}-L_{2}\end{cases}$$

and the estimator for β is then defined by

$$\arg\min_{b}\sum_{i=1}^{n}\sum_{t<s}w_{its}V\left(y_{it},y_{is},\left(x_{it}-x_{is}\right)'b,L_{it},L_{is},U_{it},U_{is}\right)$$

A version of this can be developed for a general loss function.

All of these extensions assume that the censoring and truncation points are exogenous in the sense that one must make assumptions on the error terms conditional on them. In a recent paper, Khan, Ponomareva, and Tamer (2011) consider a (one–sided) censored regression model with endogenous censoring. Their approach only leads to partial identification, but it would be interesting to generalize it to more general versions of the models considered here.

5 Empirical Application

In this section we apply the estimator in Section 2.1 to analyze the portfolio-reshuffling effect of a tax reform that increased the after-tax capital income on bonds relative to stocks in Denmark in 1987. We use a panel data set constructed from administrative records covering two years before and after the reform to estimate a portfolio share equation for bonds as a function of marginal tax

rates on capital income. The analysis presented here follows the literature on taxation and portfolio structure, e.g., Feldstein (1976), Hubbard (1985), King and Leape (1998), Samwick (2000), Poterba and Samwick (2002), Poterba (2002) and Alan, Crossley, Atalay, and Jeon (2010). These papers analyze (repeated) cross sections of households.[6] Here the analysis is extended by using panel data and controlling for time–invariant correlated heterogeneity, i.e., fixed effects. Controlling for correlated unobserved fixed factors is likely to be important in this context, since the portfolio composition of a household is likely to be influenced by time–invariant factors such as risk aversion and time discounting.

In the next subsection we give a brief overview over the tax reform. After this, we introduce the data and present the results.

5.1 The Tax Reform

The tax reform, announced in 1985 and implemented in 1987, broke the link between the marginal tax rates on earned income and capital income. Before the reform, all income was taxed at the same marginal tax rate. With the reform the tax rate on positive capital income for high-income households was decreased from 73 percent to 56 percent. The reform thereby increased the after-tax return on interest-bearing assets and therefore encouraged households to shift their portfolios toward such assets. The reform also changed the tax value of interest deductions from 73 to about 50 percent, and this substantially increased the cost of debt, primarily mortgages, for leveraged high-income households. For such households the reform effectively brought a negative wealth shock, giving them a strong incentive to lower their debt burden.[7]

The exact changes, however, differed across municipalities. The Danish income tax system is built around a proportional local government tax and a progressive tax collected by the central government. While the progressive schedule is the same for everybody in Denmark, the local

[6] Bakija (2000) uses the limited panel module of the American Survey of Consumer Finances (SCF) to study portfolio changes around the 1988 tax reform. However, his data set is very small (984 households) and unrepresentative due to the well-known attrition problem in the SCF panel module; see Kennickell and Woodburn (1997). More important in this context, the estimators applied do not exploit the full potential of the panel data in handling unobserved heterogeneity. Ioannides (1992) also employs the 1983-1986 SCF panel module but does not control for unobserved heterogeneity.

[7] Alan and Leth-Petersen (2006) document that the reduced value of the interest deduction led households to liquidate financial assets to lower their mortgage debt. This was possible because pre-payment of mortgage debt is not restricted in Denmark.

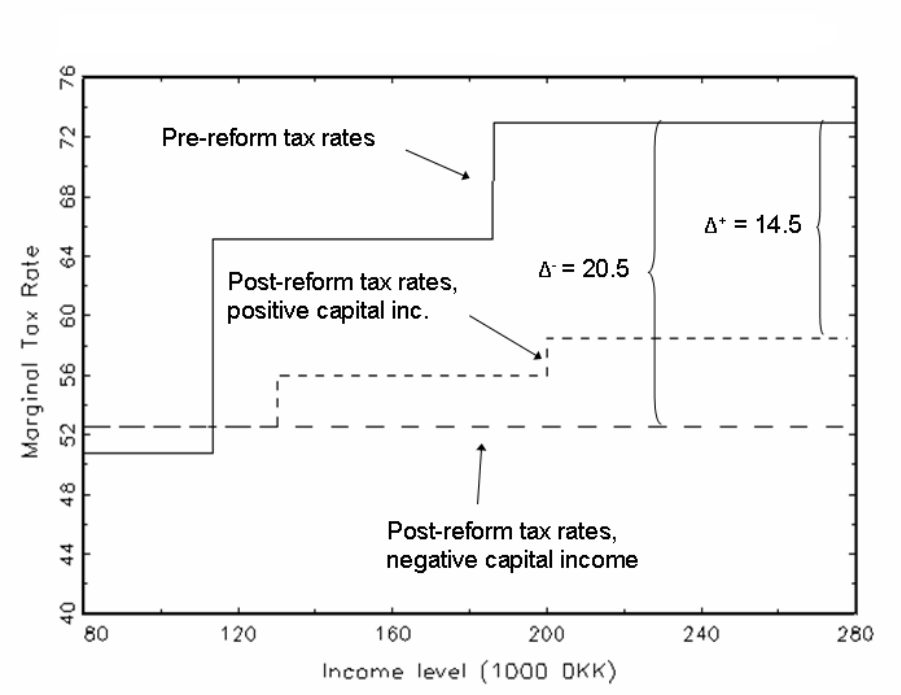

Figure 11: Marginal Tax Rate for High–Tax Municipality.

government tax rates vary across municipalities. A tax ceiling, however, insured that the marginal tax rate could be at the maximum 73 percent. After the reform the tax ceiling on earned income was reduced to 68 percent in the highest bracket[8] and 56 percent in the middle bracket. Capital income was now taxed at the same rate independently of the level of earned income. The marginal tax rates across tax brackets before and after the reform are summarized in Table 1 (see Appendix 2)

The application of a tax ceiling together with the heterogeneous local government tax rates implies that the reform had differential effects on people living in different municipalities. Figures 11 and 12 illustrate the changes in marginal tax rates due to the reform for a high-tax and a low-tax municipality, respectively.

For a high-income person living in the municipality with the high local government tax, the marginal tax rate on positive net capital income falls by 14.5 percentage points and the marginal tax rate on negative net capital income falls by 20.5 percentage points. For a similar person living in the municipality with the low local government tax rate, the marginal tax rate on positive

[8] Approximately 20 percent of the population belong in the top bracket.

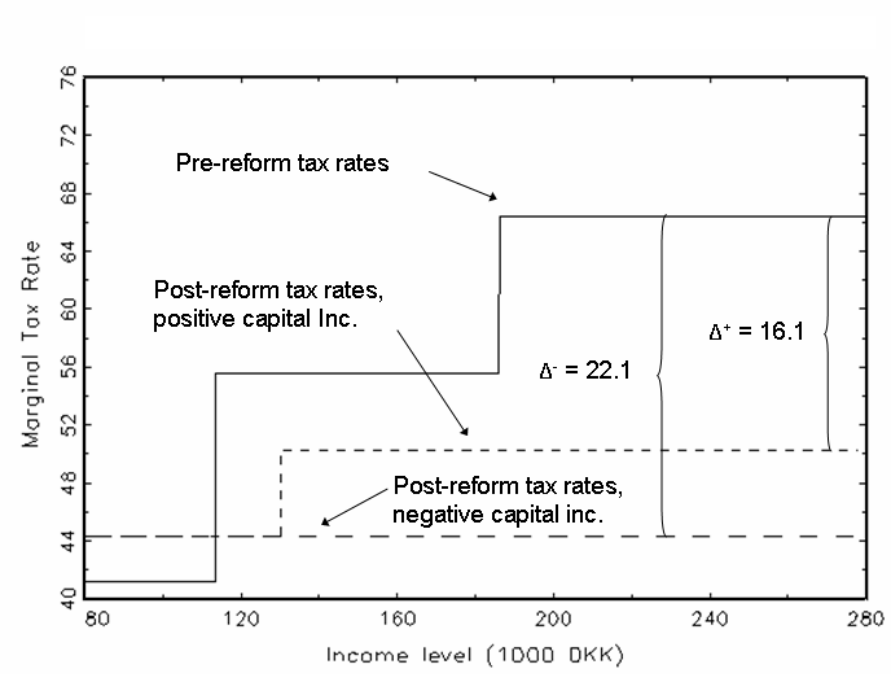

Figure 12: Marginal Tax Rate for Low–Tax Municipality.

net capital income falls by 16.1 percentage points and negative net capital income falls by 22.1 percentage points. It is these differences in changes of marginal tax rates that we will exploit for identifying the effect of changes in marginal tax rates on the portfolio allocation when using the fixed effects estimator.

The marginal tax rates on capital income refer to income received in the form of dividends on stocks and interest payments from interest bearing accounts and bonds. Both before and after the reform, realized capital gains/losses associated with trading assets were generally not taxed. The exemption from this rule is capital gains from corporate stocks held for less than three years. Such capital gains are taxed as earnings. Dividend payments were low relative to interest received from bonds.[9] This suggests that lowering the marginal tax rate on positive capital income affected bonds and stocks differentially, favoring mainly income from bonds. In the empirical analysis we therefore focus on reshuffling between bonds and stocks.

[9]The median household in the sample holding stocks received dividends corresponding to 2 percent of the value of the stocks. The median household in the sample holding bonds received interest payments from these corresponding to 10 percent of the value of the bonds.

5.2 Data

The data set is drawn from a random sample of 10 percent of the Danish population observed in the years 1984 to 1988. Information on portfolio allocations, income, wealth and demographics is collected and merged from different public administrative registers for all adult members of the household that the sampled person belongs to. Portfolio and income information is obtained from the income tax register. The portfolio information exists because Denmark had a wealth tax that required all wealth holdings to be reported to the tax authorities. This information allows us to break the wealth of each household into holdings of stocks and bonds. "Stocks" includes all holdings of publicly and privately traded stocks, and "bonds" includes government and corporate bonds. The holdings of stocks and bonds are self–reported through the tax return and then audited by the tax authority.

5.2.1 Sample selection

For our analysis we exclude observations if one of the household members is self-employed, since register data are not likely to contain a good measure of own business wealth and because taxable income is quite volatile for those individuals. Sampled individuals younger than 18 or older than 60 are dropped as are students and individuals living together with his/her parents or living in a common household, i.e., a household with more than one family. To keep the focus on the importance of tax incentives, we include only stable couples, i.e., couples where the partner is the same in 1984 through 1988. On the same grounds we also exclude couples moving in the sample period. For the purpose of the analysis we require that households entering the sample be observed in all years in the period 1984-1988 so that we have a balanced panel.

Our objective is to investigate whether households reshuffle their portfolios in response to a change in tax incentives. As in most industrialized countries many Danish households have fairly undiversified portfolios. Since the decrease in the value of interest deductions generated a large negative wealth shock, clearly, these households are not likely to engage in portfolio reshuffling and hence cannot give us a clean answer regarding portfolio readjustments. We therefore construct a sub-sample of households holding positive amounts of stocks or bonds of at least 5,000 DKK in 1984. We also require households to hold a positive amount of either stocks or bonds throughout the rest of the observation period. This selection is introduced because we want to focus on households with a potential to reshuffle between stocks and bonds. Also renters are deselected because there

are few renters with diversified portfolios.[10] The final subsample includes 8,577 households.[11]

5.3 Results

In this section we investigate if households reshuffled their portfolio of bonds and stocks as a response to the changes in relative after-tax returns on assets brought about by the 1987 tax reform. To do, this we employ the estimator presented above and estimate a portfolio share equation where the fraction of bonds in financial wealth, defined as the sum of bonds and stocks, is regressed on the marginal tax rate on positive capital income and some control variables.

The distinguishing feature of our data set is the panel dimension. This facilitates estimating portfolio share equations allowing for correlated unobserved time–invariant heterogeneity. This is important because we believe that unobserved time–invariant factors, such as risk aversion and time preferences, are correlated with wealth. High risk aversion may, for example, lead to a higher portfolio share of safe assets, such as bonds, for a given level of wealth.

Before the reform, capital income and earnings were lumped together and taxed according to a progressive tax scheme. This implies that households choose their tax bracket when choosing their portfolios and that the marginal tax rate on capital income is likely to be an endogenous regressor. We address this by calculating the marginal tax rate on capital income based on the household's income in 1984, the year before the reform was announced, but using current year rules. In this way the individual level tax bracket is allocated based on information that was predetermined relative to the portfolio response to the reform.

We regress portfolio shares on the marginal tax rate on positive capital income, the log of total financial assets, i.e., assets held in stocks and bonds, and a set of year dummies. Tax rate changes vary across municipalities, but most of the change in tax rates is common across municipalities. Year dummies control for the effect of this common part, thereby also removing the major part of the wealth effect brought about by the reform. Effectively, by introducing year dummies, the coefficients on marginal tax rates are identified by differences in changes of marginal tax rates. Year dummies may also pick up common effects related to fluctuations in assets. Financial assets

[10] For assessing portfolio reshuffling renters could have been included. We have chosen to leave them out of this analysis because there are only a few renters (898) with positive financial wealth of at least 5000 DKK in 1984. Moreover, renters generally do not provide a good comparison group for homeowners, since different preference parameters may govern their behavior.

[11] See Alan and Leth-Petersen (2006) for a more detailed analysis.

control for any remaining wealth effect that might be present.[12]

Table 2 of Appendix 2 presents the parameter estimates from estimating random effects Tobit and fixed effects censored regression models for the portfolio share of bonds in financial wealth. The estimated parameter on the marginal tax rate is negative.[13] If year dummies and financial assets pick up the wealth effect related to the reform, in particular the effect of the reduction in the value of interest deduction that led households to liquidate financial assets, then this is exactly what economic theory predicts. Households should substitute from stocks toward bonds, whose relative after-tax return increased, and this is what the results indicate.

Considering the corresponding random effects estimates, we can see that the parameter estimates on financial assets and on year dummies are quite different, and the test of equality of all the parameters in the random effects and fixed effects specifications rejects.

6 Concluding Remarks

This paper constructs estimators for panel data regression models with individual specific heterogeneity and two–sided censoring and truncation. Following Powell (1986) the estimation strategy is based on moment conditions constructed from re–censored or re–truncated residuals. While these moment conditions do not identify the parameter of interest, they can be used to motivate objective functions that do. This part is the main methodological contribution of the paper. We apply one of the estimators to study the effect of a Danish tax reform on household portfolio choice. We find that a random effects specification can be rejected in favor of the "fixed" effects specification studied here, although both models yield the same sign of the key parameter that one would anticipate from economic theory. The estimators are fairly easy to implement and a link to a program that calculates the leading estimator is provided at `http://www.princeton.edu/~honore/stata.`

[12] An alternative identification strategy could be based on comparing the behavior of households in different tax brackets. Households in the lowest tax bracket faced only very small changes in marginal tax rates on capital income, and households in the middle tax bracket faced different changes in marginal tax rates than households in the highest tax bracket. In our case this is not a natural approach to follow. High– and low–income people are different in terms of wealth levels and portfolio composition and possibly different with respect to preference parameters such as the discount rate and the level of risk aversion. Households in lower tax brackets therefore do not represent a natural control group for high–income households.

[13] As explained in Honoré (2008), the parameter estimates for both the random effects and the fixed effects models can be converted to marginal effects by multiplying them by the fraction of observations that are not censored.

References

Abrevaya, J. (1999): "Rank estimation of a transformation model with observed truncation," *Econometrics Journal*, 2(2), 292–305.

Alan, S., T. Crossley, K. Atalay, and S.-H. Jeon (2010): "New Evidence on Taxes and Portfolio Choice," *Journal of Public Economics*, 94, 813–823.

Alan, S., and S. Leth-Petersen (2006): "Tax Incentives and Household Portfolios: A Panel Data Analysis," Center for Applied Microeconometrics, University of Copenhagen, Working paper number 2006-13.

Amemiya, T. (1985): *Advanced Econometrics*. Harvard University Press.

Andrews, M., T. Schank, and R. Simmons (2005): "Does Worksharing Work? Some Empirical Evidence from the IAB-Establishment Panel," *Scottish Journal of Political Economy*, 52(2).

Arellano, M., and B. E. Honoré (2001): "Panel Data Models. Some Recent Developments," *Handbook of Econometrics*, 5, Elsevier Science, Amsterdam.

Bakija, J. (2000): "The Effect of Taxes on Portfolio Choice: Evidence from Panel Data Spanning the Tax Reform Act of 1986," Williams College.

Charness, G., G. Frechette, and J. Kagel (2004): "How Robust Is Laboratory Gift Exchange?," *Experimental Economics*, 7(2), 189–205.

Chen, S. (forthcoming): "Nonparametric Identification and Estimation of Truncated Regression Models," *Review of Economic Studies*.

de Figueriredo, J., and E. Tiller (2001): "The Structure and Conduct of Corporate Lobbying: How Firms Lobby the Federal Communications Commission," *Journal of Economics and Management Strategy*, 10(1), 91–122.

Fehr, E., E. Kirchler, A. Weichbold, and S. Gachter (1998): "When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets," *Journal of Labor Economics*, 16(2), 324–351.

Feldstein, M. (1976): "Personal Taxation and Portfolio Composition: An Econometric Analysis," *Econometrica*, 44(4), 631–650.

Fenn, G., and N. Liang (2001): "Corporate Payout Policy and Managerial Stock Incentives," *Journal of Financial Economics*, 60(1), 45–72.

Ferrie, J. P., and K. Rolf (2011): "Socioeconomic Status in Childhood and Health after Age 70: A New Longitudinal Analysis for the U.S., 1895-2005," *SSRN eLibrary*.

Gifford, K., and J. Bernard (2005): "Influencing Consumer Purchase Likelihood of Organic Food," *International Journal of Consumer Studies*, pp. 1–9.

Han, A. (1987): "Nonparametric Analysis of a Generalized Regression Model: The Maximum Rank Correlation Estimator," *Journal of Econometrics*, 15.

Honoré, B. E. (1992): "Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects," *Econometrica*, 60, 533–565.

——— (2008): "On Marginal Effects in Semiparametric Censored Regression Models," Princeton University.

Honoré, B. E., and L. Hu (2004): "Estimation of Cross Sectional and Panel Data Censored Regression Models with Endogeneity," *Journal of Econometrics*, 122(2), 293–316.

Honoré, B. E., and J. L. Powell (1994): "Pairwise Difference Estimators of Censored and Truncated Regression Models," *Journal of Econometrics*, 64, 241–78.

Honoré, B. E., and J. L. Powell (1994): "Pairwise Difference Estimators of Censored and Truncated Regression Models," *Journal of Econometrics*, 64(1-2), 241–78.

Houston, J., and M. Ryngaert (1997): "Equity Issuance and Adverse Selection: A Direct Test Using Conditional Stock Offers," *Journal of Finance*, 52(1), 197–219.

Hu, L. (2002): "Estimation of a Censored Dynamic Panel Data Model," *Econometrica*, 70(6), pp. 2499–2517.

Huang, M., and R. Hauser (1998): "Trends in Black-White Test Score Differences: WORDSUM Vocabulary Test," in *The Rising Curve: Long-IQ and Related Measures*, ed. by U. Neisser, pp. 303–332. American Psychological Association, Washington, D.C.

——— (2001): "Convergent Trends in Black-White Verbal Test Score Differentials in the U.S.: Period and Cohort Perspectives," *EurAmerica*, 31(2), 185–230.

Hubbard, R. G. (1985): "Personal Taxation, Pension Wealth, and Portfolio Composition," *The Review of Economics and Statistics*, 67(1), 53–60.

Ioannides, Y. (1992): "Dynamics of the Composition of Household Asset Portfolios and Life Cycle," *Applied Financial Economics*, 2(3), 145–159.

Kennickell, A., and L. Woodburn (1997): "Weighting Design for the 1983-89 SCF Panel," Washington DC: Federal Reserve Board of Governors.

Khan, S., M. Ponomareva, and E. Tamer (2011): "Identification of Panel Data Models with Endogenous Censoring," MPRA Paper 30373, University Library of Munich, Germany.

King, M., and J. Leape (1998): "Wealth and Portfolio Composition: Theory and Evidence," *Journal of Public Economics*, 69(2), 155–193.

Kyriazidou, E. (1997): "Estimation of a Panel Data Sample Selection Model," *Econometrica*, 65, 1335–1364.

Lafontaine, F. (1993): "Contractual Arrangements as Signalling Devices: Evidence from Franchising," *Journal of Law, Economics, and Organization*, 9(2), 256–289.

Manski, C. (1987): "Semiparametric Analysis of Random Effects Linear Models from Binary Panel Data," *Econometrica*, 55, 357–62.

McMillan, J., and C. Woodruff (1999): "Interfirm Relationships and Informal Credit in Vietnam," *Quarterly Journal of Economics*, 114(4), 1285–1320.

Nickerson, J., and B. Silverman (2003): "Why Aren't All Truck Drivers Owner-Operators? Asset Ownership and the Employment Relation in Interstate For-Hire Trucking," *Journal of Economics and Management Strategy*, 12(1), 91–118.

Officer, M. (2004): "Collars and Renegotiation in Mergers and Acquisitions," *Journal of Finance*, 59(6), 2719–2743.

Pakes, A., and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057.

Petersen, M., and R. Rajan (1994): "The Benefits of Lending Relationships: Evidence from Small Business Data," *Journal of Finance*, 49(1), 3–37.

——— (1995): "The Effect of Credit Market Competition on Lending Relationships," *Quarterly Journal of Economics*, 110(2), 407–443.

Poterba, J., and A. Samwick (2002): "Taxation and Household Portfolio Composition: US Evidence from the 1980s and 1990s," *Journal of Public Economics*, 87(1), 5–38.

Poterba, J. M. (2002): "Taxation and Portfolio Structure: Issues and Implications," in *Household Portfolios*, ed. by L. Guiso, M. Haliassos, and T. Jappelli. MIT Press.

Powell, J. L. (1986): "Symmetrically Trimmed Least Squares Estimation for Tobit Models," *Econometrica*, 54(6), 1435–60.

——— (1987): "Semiparametric Estimation of Bivariate Latent Models," Working Paper no. 8704, Social Systems Research Institute, University of Wisconsin–Madison.

Samwick, A. (2000): "Portfolio Responses to Taxation: Evidence from the End of the Rainbow," in *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, ed. by J. Slemrod. Harvard University Press, Cambridge, MA.

7 Appendix 1: Proofs and Derivations

7.1 Two–Sided Censoring

This section provides justification for the statements about the estimators for the model with two–sided censoring. We first verify that $u\left(y_{it}, y_{is}, x_{it}'\beta - x_{is}'\beta\right)$ in equation (8) does indeed yield the difference in the re-censored residuals defined in (6) and (7). Write $d = x_{it}'\beta - x_{is}'\beta$, and consider first the case where $d > 0$. In this case, the difference in the re-censored residuals is

$$\max\left\{y_{it} - d, 0\right\} - \min\left\{y_{is}, 1 - d\right\}$$

which is most easily analyzed by considering a number of cases for d between 0 and 1. As mentioned earlier, the contribution to the moment condition should be 0 when d is outside the interval between -1 and 1,which is consistent with the definition of u.

There are four cases based on combinations of whether $y_{it} - d \geq 0$ and $y_{is} \geq 1 - d$:[14]

- Case 1 ($y_{it} - d \geq 0$ and $y_{is} \geq 1 - d$): In this case, $\max\left\{y_{it} - d, 0\right\} - \min\left\{y_{is}, 1 - d\right\} = (y_{it} - d) - (1 - d) = y_{it} - 1$.
- Case 2 ($y_{it} - d \leq 0$ and $y_{is} \geq 1 - d$): In this case, $\max\left\{y_{it} - d, 0\right\} - \min\left\{y_{is}, 1 - d\right\} = 0 - (1 - d) = d - 1$.
- Case 3 ($y_{it} - d \geq 0$ and $y_{is} \leq 1 - d$): In this case, $\max\left\{y_{it} - d, 0\right\} - \min\left\{y_{is}, 1 - d\right\} = (y_{it} - d) - y_{is} = y_{it} - y_{is} - d$.
- Case 4 ($y_{it} - d \leq 0$ and $y_{is} \leq 1 - d$): In this case, $\max\left\{y_{it} - d, 0\right\} - \min\left\{y_{is}, 1 - d\right\} = 0 - y_{is} = -y_{is}$.

Case 3 corresponds to values of d close to (or at) 0. Specifically, the region for Case 3 is $(0, \min\left\{y_{it}, 1 - y_{is}\right\}) = (0, c_3)$. Noting that $c_2 \leq 0$, it is clear that $\max\left\{y_{it} - d, 0\right\} - \min\left\{y_{is}, 1 - d\right\} = y_{it} - y_{is} - d$ is consistent with the definition of u in equation (8).

Case 2 corresponds to values of d close to (or at) 1, specifically the region $(\max\left\{y_{it}, 1 - y_{is}\right\}, 1) = (c_4, 1)$, and it is again clear that u delivers $\max\left\{y_{it} - d, 0\right\} - \min\left\{y_{is}, 1 - d\right\}$.

The region that defines the other two cases is $(\min\left\{y_{it}, 1 - y_{is}\right\}, \max\left\{y_{it}, 1 - y_{is}\right\}) = (c_3, c_4)$. Cases 1 and 4 give different expressions depending on whether y_{it} or $1 - y_{is}$ is larger, but these

[14] Since both the difference in the re-censored residuals and u are continuous in d, it is not necessary to distinguish between closed and open intervals in the following discussion.

expressions correspond exactly to the two cases for the max in the definition of u in (8) over the interval (c_3, c_4).

The case $d < 0$ is dealt with in exactly the same manner.

We now turn to the question of why

$$E\left[U\left(y_{it}, y_{is}, (x_{it} - x_{is})' b\right)\right]$$

is uniquely minimized at the true β. This is the key for consistency of the proposed estimator.

The result follows from the following lemma

Lemma 3 *Suppose*

$$y_{i1} = \text{mami}\left\{0, \delta + \varepsilon_{i1}, 1\right\}$$

and

$$y_{i2} = \text{mami}\left\{0, \varepsilon_{i2}, 1\right\}$$

where ε_{i1} and ε_{i2} are identically distributed random variables with support on the whole real line. Then

$$\arg \max_{d \in [-1,1]} E\left[U\left(y_{i1}, y_{i2}, d\right)\right] = \begin{cases} -1 & \text{if} \quad \delta \leq -1 \\ \delta & \text{if} \quad -1 < \delta < 1 \\ 1 & \text{if} \quad \delta \geq 1 \end{cases}$$

PROOF: For $1 \geq d \geq 0$

$$\begin{aligned} E\left[u\left(y_{i1}, y_{i2}, d\right)\right] &= E\left[\max\left\{y_{i1} - d, 0\right\} - \min\left\{y_{i2}, 1 - d\right\}\right] \\ &= E\left[\max\left\{\text{mami}\left\{0, \delta + \varepsilon_{i1}, 1\right\} - d, 0\right\} - \min\left\{\text{mami}\left\{0, \varepsilon_{i2}, 1\right\}, 1 - d\right\}\right] \\ &= E\left[\max\left\{\text{mami}\left\{0 - d, \delta - d + \varepsilon_{i1}, 1 - d\right\}, 0\right\} - \text{mami}\left\{0, \varepsilon_{i2}, 1 - d\right\}\right] \\ &= E\left[\text{mami}\left\{0, \delta - d + \varepsilon_{i1}, 1 - d\right\} - \text{mami}\left\{0, \varepsilon_{i2}, 1 - d\right\}\right] \end{aligned}$$

If ε_{i1} (and ε_{i2}) have full support, then this is negative for $d > \delta$ and positive for $d < \delta$.

For $-1 \leq d \leq 0$

$$\begin{aligned} E\left[u\left(y_{i1}, y_{i2}, d\right)\right] &= E\left[\min\left\{y_{i1}, 1 + d\right\} - \max\left\{y_{i2} + d, 0\right\}\right] \\ &= E\left[\min\left\{\text{mami}\left\{0, \delta + \varepsilon_{i1}, 1\right\}, 1 + d\right\} - \max\left\{\text{mami}\left\{0, \varepsilon_{i2}, 1\right\} + d, 0\right\}\right] \\ &= E\left[\text{mami}\left\{0, \delta + \varepsilon_{i1}, 1 + d\right\} - \max\left\{\text{mami}\left\{d, \varepsilon_{i2} + d, 1 + d\right\}, 0\right\}\right] \\ &= E\left[\text{mami}\left\{0, \delta + \varepsilon_{i1}, 1 + d\right\} - \text{mami}\left\{0, \varepsilon_{i2} + d, 1 + d\right\}\right] \end{aligned}$$

If ε_{i1} (and ε_{i2}) have full support, then this is negative for $d > \delta$ and positive for $d < \delta$.

Since

$$E\left[U'\left(y_{i1}, y_{i2}, d\right)\right] = E\left[u\left(y_{i1}, y_{i2}, d\right)\right]$$

the argument above shows that

$$\arg \max_{d \in [-1,1]} E\left[U\left(y_{i1}, y_{i2}, d\right)\right] = \begin{cases} -1 & \text{if} \quad \delta \leq -1 \\ \delta & \text{if} \quad -1 < \delta < 1 \\ 1 & \text{if} \quad \delta \geq 1 \end{cases} \qquad \blacksquare$$

Corollary 4 *Consider the model*

$$\begin{aligned} y_{it}^* &= x_{it}'\beta + \varepsilon_{it} \\ \\ y_{it} &= \begin{cases} 0 & \textit{if} \quad y_{it}^* < 0 \\ y_{it}^* & \textit{if} \quad 0 \leq y_{it}^* \leq 1 \\ 1 & \textit{if} \quad y_{it}^* > 1 \end{cases} \end{aligned}$$

for $t = 1, 2$. If ε_{it} is stationary conditional on (x_{i1}, x_{i2}) with support on the whole real line, then the set of solutions to

$$\max_b E\left[U\left(y_{i1}, y_{i2}, \text{mami}\left\{-1, \left(x_{i1} - x_{i2}\right)' b, 1\right\}\right)\right]$$

is

$$\left\{b : P\left(\text{mami}\left\{-1, \left(x_{i1} - x_{i2}\right)' b, 1\right\} = \text{mami}\left\{-1, \left(x_{i1} - x_{i2}\right)' \beta, 1\right\}\right) = 1\right\}$$

The Corollary above requires that the errors are stationary conditional on the regressors. This is much more general than the usual assumption that the individual—specific effect and the contemporaneous errors interact additively. To see that $E\left[U_\Xi\left(y_{it}, y_{is}, \left(x_{it} - x_{is}\right)' b\right)\right]$ in (12) is uniquely minimized, it is convenient to assume that ε_{it} and ε_{is} are independent and identically distributed conditional on some individual—specific effect, v_i. With this assumption, the argument for why $E\left[U_\Xi\left(y_{it}, y_{is}, \left(x_{it} - x_{is}\right)' b\right)\right]$ is uniquely minimized follows essentially the same logic as above. Specifically, when $1 \geq d \geq 0$

$$\xi\left(u\left(y_{it}, y_{is}, d\right)\right) = \xi\left(\text{mami}\left\{0, \delta - d + \varepsilon_{it}, 1 - d\right\} - \text{mami}\left\{0, \varepsilon_{is}, 1 - d\right\}\right)$$

Let $\delta = \left(x_{it} - x_{is}\right)' \beta$.

If ε_{it} and ε_{is} are independent and identically distributed conditional on v_i, then

$$E\left[\xi\left(u\left(y_{it},y_{is},\delta\right)\right)\middle|x_{it},x_{is}\right]=E\left[E\left[\xi\left(u\left(y_{it},y_{is},\delta\right)\right)\middle|\nu_i,x_{it},x_{is}\right]\middle|x_{it},x_{is}\right]$$
$$=E\left[E\left[\xi\left(\text{mami}\left\{0,\varepsilon_{it},1-\delta\right\}-\text{mami}\left\{0,\varepsilon_{is},1-\delta\right\}\right)\middle|\nu_i,x_{it},x_{is}\right]\middle|x_{it},x_{is}\right]=0$$

because $\text{mami}\left\{0,\varepsilon_{it},1-\delta\right\}-\text{mami}\left\{0,\varepsilon_{is},1-\delta\right\}$ is symmetrically distributed conditional on ν_i, and $\xi\left(\cdot\right)$ is an odd function.

For $d>\delta$

$$\text{mami}\left\{0,\delta-d+\varepsilon_{it},1-d\right\}-\text{mami}\left\{0,\varepsilon_{is},1-d\right\}$$
$$\leq\text{mami}\left\{0,\varepsilon_{it},1-d\right\}-\text{mami}\left\{0,\varepsilon_{is},1-d\right\}$$

with probability 1 (conditional on ν_i), and since ξ is increasing

$$E\left[\xi\left(\text{mami}\left\{0,\delta-d+\varepsilon_{it},1-d\right\}-\text{mami}\left\{0,\varepsilon_{is},1-d\right\}\right)\middle|\nu_i,x_{it},x_{is}\right]$$
$$\leq E\left[\xi\left(\text{mami}\left\{0,\varepsilon_{it},1-d\right\}-\text{mami}\left\{0,\varepsilon_{is},1-d\right\}\right)\middle|\nu_i,x_{it},x_{is}\right]=0$$

The line of argument is the same when $d<\delta$. Strict inequalities follow from a full support assumption on ε_{i1} (and ε_{i2}).

We therefore have that $E\left[U_\Xi\left(y_{it},y_{is},d\right)\middle|x_{it},x_{is}\right]$ is decreasing to the left of $\left(x_{it}-x_{is}\right)'\beta$ and increasing to the right. Hence it is minimized at $d=\left(x_{it}-x_{is}\right)'\beta$. Subject to a rank condition, this implies that $E\left[U_\Xi\left(y_{it},y_{is},\left(x_{it}-x_{is}\right)'b\right)\middle|x_{it},x_{is}\right]$ is minimized at $b=\beta$.

7.2 Two–Sided Truncation

The following Lemma (combined with the obvious rank–condition) establishes that minimization of $E\left[R_\Xi\left(y_{i1},y_{i2},\left(x_{i1}-x_{2i}\right)'b\right)\right]$ will identify β if the distribution of ε is log–concave. This is the assumption that was made in a number of other papers (including Honoré (1992), Honoré and Powell (1994) and Abrevaya (1999); see also the discussion in Chen (forthcoming)). In Section 2.2, we only consider the case with (two–sided) truncation at 0 and 1. It is just as easy to prove identification for general individual– and time–specific truncation points. In the following we therefore denote the truncation points by L and U.

Lemma 5 *Let (L,U) be a vector of random variables such that $L<U$ with probability 1. Assume that ε is independent of (L,U) and has a continuous, log–concave distribution with support on the whole real line. Let $y_{it}=\delta_{it}+\varepsilon_{it}$ for some real number, δ_{it}, and consider two draws (y_{i1},L_{i1},U_{i1}) and*

(y_{i2}, L_{i2}, U_{i2}) *from the distribution of* (y, L, U) *conditional on* $L < y < U$*. then* $E\left[R_{\Xi}\left(y_{i1}, y_{i2}, d\right)\right]$ *is uniquely minimized at* $d = \delta_{i1} - \delta_{i2}$*.*

7.3 Proof of Lemma 5.

Let E denote expectation conditional on truncation and E^* in population.

The moment condition can then be written as

$$\begin{aligned}
& E\left[\xi\left(\triangle y_i - \triangle d_i\right) 1\left\{L_{i2} - d_{i2} \leq y_{i1} - d_{i1} \leq U_{i2} - d_{i2}\right\} 1\left\{L_{i1} - d_{i1} \leq y_{i2} - d_{i2} \leq U_{i1} - d_{i1}\right\}\right] \\
= \; & E\left[\xi\left(\triangle y_i - \triangle d_i\right) 1\left\{L_{i2} + \triangle d_i \leq y_{i1} \leq U_{i2} + \triangle d_i\right\} 1\left\{L_{i1} - \triangle d_i \leq y_{i2} \leq U_{i1} - \triangle d_i\right\}\right] \\
= \; & E[\xi\left(\triangle \varepsilon_i + \triangle \delta_i - \triangle d_i\right) 1\left\{L_{i2} - \left(\triangle \delta_i - \triangle d_i\right) \leq \varepsilon_{i1} + \delta_{i2} \leq U_{i2} - \left(\triangle \delta_i - \triangle d_i\right)\right\} \\
& 1\left\{L_{i1} + \left(\triangle \delta_i - \triangle d_i\right) \leq \varepsilon_{i2} + \delta_{i1} \leq U_{i1} + \left(\triangle \delta_i - \triangle d_i\right)\right\}]
\end{aligned}$$

where $\triangle a_i = a_{i1} - a_{i2}$. Letting $\kappa_i = \left(\triangle \delta_i - \triangle d_i\right)$

$$\begin{aligned}
= \; & E[\xi\left(\triangle \varepsilon_i + \kappa_i\right) 1\left\{L_{i2} - \delta_{i2} - \kappa_i \leq \varepsilon_{i1} \leq U_{i2} - \delta_{i2} - \kappa_i\right\} 1\left\{L_{i1} - \delta_{i1} + \kappa_i \leq \varepsilon_{i2} \leq U_{i1} - \delta_{i1} + \kappa_i\right\}] \\
= \; & E^*[\xi\left(\triangle \varepsilon_i + \kappa_i\right) 1\left\{L_{i2} - \delta_{i2} - \kappa_i \leq \varepsilon_{i1} \leq U_{i2} - \delta_{i2} - \kappa_i\right\} 1\left\{L_{i1} - \delta_{i1} + \kappa_i \leq \varepsilon_{i2} \leq U_{i1} - \delta_{i1} + \kappa_i\right\} \\
& 1\left\{L_{i1} \leq y_{i1} \leq U_{i1}\right\} 1\left\{L_{i2} \leq y_{i2} \leq U_{i2}\right\}] \\
& \frac{1}{P\left(L_{i1} \leq y_{i1} \leq U_{i1}, L_{i2} \leq y_{i2} \leq U_{i2}\right)}.
\end{aligned}$$

It suffices to show that this is nonpositive for $\kappa_i < 0$, strictly negative in a neighborhood to the left of 0, 0 for $\kappa_i = 0$, nonnegative for $\kappa_i > 0$, and strictly positive in a neighborhood to the right of 0. Now consider the term

$$\begin{aligned}
& E^*[\xi\left(\triangle \varepsilon_i + \kappa_i\right) 1\left\{L_{i2} - \delta_{i2} - \kappa_i \leq \varepsilon_{i1} \leq b_j - \delta_{i2} - \kappa_i\right\} 1\left\{L_{i1} - \delta_{i1} + \kappa_i \leq \varepsilon_{i2} \leq U_{i1} - \delta_{i1} + \kappa_i\right\} \\
& 1\left\{L_{i1} - \delta_{i1} \leq \varepsilon_{i1} \leq U_{i1} - \delta_{i1}\right\} 1\left\{L_{i2} - \delta_{i2} \leq \varepsilon_{i2} \leq b_j - \delta_{i2}\right\}].
\end{aligned}$$

Defining $w_{i1} = \frac{1}{2}\left(\varepsilon_{i1} - \varepsilon_{i2}\right)$ and $w_{i2} = \frac{1}{2}\left(\varepsilon_{i1} + \varepsilon_{i2}\right)$ (so $\varepsilon_{i1} = \left(w_{i1} + w_{i2}\right)$ and $\varepsilon_{i2} = w_{i2} - w_{i1}$ and $\triangle \varepsilon_i = 2w_{i1}$)

$$
\begin{aligned}
& E^*[\xi\left(\triangle\varepsilon_i+\kappa_i\right)1\left\{L_{i2}-\delta_{i2}-\kappa_i\leq\varepsilon_{i1}\leq U_{i2}-\delta_{i2}-\kappa_i\right\}1\left\{L_{i1}-\delta_{i1}+\kappa_i\leq\varepsilon_{i2}\leq U_{i1}-\delta_{i1}+\kappa_i\right\} \\
& 1\left\{L_{i1}-\delta_{i1}\leq\varepsilon_{i1}\leq U_{i1}-\delta_{i1}\right\}1\left\{L_{i2}-\delta_{i2}\leq\varepsilon_{i2}\leq U_{i2}-\delta_{i2}\right\}] \\
= \; & E^*[\xi\left(2w_{i1}+\kappa_i\right)1\left\{L_{i2}-\delta_{i2}-\kappa_i\leq w_{i1}+w_{i2}\leq U_{i2}-\delta_{i2}-\kappa_i\right\} \\
& 1\left\{L_{i1}-\delta_{i1}+\kappa_i\leq w_{i2}-w_{i1}\leq U_{i1}-\delta_{i1}+\kappa_i\right\} \\
& 1\left\{L_{i1}-\delta_{i1}\leq w_{i1}+w_{i2}\leq U_{i1}-\delta_{i1}\right\} \\
& 1\left\{L_{i2}-\delta_{i2}\leq w_{i2}-w_{i1}\leq U_{i2}-\delta_{i2}\right\}]
\end{aligned}
$$

$$
\begin{aligned}
= \; & E^*[\xi\left(2w_{i1}+\kappa_i\right) \\
& 1\left\{L_{i2}-\delta_{i2}-w_{i2}-\frac{1}{2}\kappa_i\leq w_{i1}+\frac{1}{2}\kappa_i\leq U_{i2}-\delta_{i2}-w_{i2}-\frac{1}{2}\kappa_i\right\} \\
& 1\left\{-U_{i1}+\delta_{i1}+w_{i2}-\frac{1}{2}\kappa_i\leq w_{i1}+\frac{1}{2}\kappa_i\leq -L_{i1}+\delta_{i1}+w_{i2}-\frac{1}{2}\kappa_i\right\} \\
& 1\left\{L_{i1}-\delta_{i1}-w_{i2}+\frac{1}{2}\kappa_i\leq w_{i1}+\frac{1}{2}\kappa_i\leq U_{i1}-\delta_{i1}-w_{i2}+\frac{1}{2}\kappa_i\right\} \\
& 1\left\{-U_{i2}+\delta_{i2}+w_{i2}+\frac{1}{2}\kappa_i\leq w_{i1}+\frac{1}{2}\kappa_i\leq -L_{i2}+\delta_{i2}+w_{i2}+\frac{1}{2}\kappa_i\right\}]
\end{aligned}
$$

$$
\begin{aligned}
= \; & E^*[\xi\left(2w_{i1}+\kappa_i\right) \\
& 1\{\max\left\{L_{i2}-\delta_{i2}-w_{i2}-\frac{1}{2}\kappa_i,-U_{i1}+\delta_{i1}+w_{i2}-\frac{1}{2}\kappa_i,L_{i1}-\delta_{i1}-w_{i2}+\frac{1}{2}\kappa_i,-U_{i2}+\delta_{i2}+w_{i2}+\frac{1}{2}\kappa_i\right\} \\
\leq \; & w_{i1}+\frac{1}{2}\kappa_i\leq \\
& \min\left\{U_{i2}-\delta_{i2}-w_{i2}-\frac{1}{2}\kappa_i,-L_{i1}+\delta_{i1}+w_{i2}-\frac{1}{2}\kappa_i,U_{i1}-\delta_{i1}-w_{i2}+\frac{1}{2}\kappa_i,-L_{i2}+\delta_{i2}+w_{i2}+\frac{1}{2}\kappa_i\right\}\}] \\
= \; & E^*\left[\xi\left(2w_{i1}+\kappa_i\right)1\left\{-c_i\leq w_{i1}+\frac{1}{2}\kappa_i\leq c_i\right\}\right] \\
= \; & E^*\left[E^*\left[\xi\left(2w_{i1}+\kappa_i\right)1\left\{-c_i\leq w_{i1}+\frac{1}{2}\kappa_i\leq c_i\right\}\middle| w_{i2},L_{i1},U_{i1},L_{i2},U_{i2}\right]\right]
\end{aligned}
$$

where $c_i=\min\left\{U_{i2}-\delta_{i2}-w_{i2}-\frac{1}{2}\kappa_i,-L_{i1}+\delta_{i1}+w_{i2}-\frac{1}{2}\kappa_i,U_{i1}-\delta_{i1}-w_{i2}+\frac{1}{2}\kappa_i,-L_{i2}+\delta_{i2}+w_{i2}+\frac{1}{2}\kappa_i\right\}$. Strict log–concavity of ε_{it}, implies that w_{i1} is strictly unimodal and symmetric conditional on w_{i2}.

It therefore follows that

$$E^*\left[E^*\left[\xi\left(2w_{i1}+\kappa_i\right)1\left\{-c_i\le w_{i1}+\frac{1}{2}\kappa_i\le c_i\right\}\middle| w_{i2},L_{i1},U_{i1},L_{i2},U_{i2}\right]\right]$$

$$\begin{cases}\le 0 & \text{if} \quad \kappa_i<0\\ <0 & \text{if} \quad \kappa_i<0, \quad P\left(-c_i\le w_{i1}+\frac{1}{2}\kappa_i\le c_i\right)>0\\ =0 & \text{if} \quad \kappa_i=0\\ \ge 0 & \text{if} \quad \kappa_i>0\\ >0 & \text{if} \quad \kappa_i>0, \quad P\left(-c_i\le w_{i1}+\frac{1}{2}\kappa_i\le c_i\right)>0.\end{cases}$$

Since the ε's (and hence the w_i's) are continuous, the condition that $P\left(-c_i\le w_{i1}+\frac{1}{2}\kappa_i\le c_i\right)>0$ will be satisfied if $P\left(c_i>0\right)>0$.

We will next show that for $\kappa_i\in(0,k)$, $P\left(c_i>0\right)>0$ for some $k>0$. Note that

$$\begin{aligned}1\left\{c_i>0\right\} &= 1\left\{-L_{i1}+\delta_{i1}+w_{i2}-\frac{1}{2}\kappa_i>0\right\}\cdot 1\left\{U_{i1}-\delta_{i1}-w_{i2}+\frac{1}{2}\kappa_i>0\right\}\\ &\quad\cdot 1\left\{-L_{i2}+\delta_{i2}+w_{i2}+\frac{1}{2}\kappa_i>0\right\}\cdot 1\left\{U_{i2}-\delta_{i2}-w_{i2}-\frac{1}{2}\kappa_i>0\right\}\\ &= 1\left\{w_{i2}>L_{i1}-\delta_{i1}+\frac{1}{2}\kappa_i\right\}\cdot 1\left\{U_{i1}-\delta_{i1}+\frac{1}{2}\kappa_i>w_{i2}\right\}\\ &\quad\cdot 1\left\{w_{i2}>L_{i2}-\delta_{i2}-\frac{1}{2}\kappa_i\right\}\cdot 1\left\{U_{i2}-\delta_{i2}-\frac{1}{2}\kappa_i>w_{i2}\right\}\\ &= 1\left\{U_{i1}-\delta_{i1}+\frac{1}{2}\kappa_i>w_{i2}>L_{i1}-\delta_{i1}+\frac{1}{2}\kappa_i\right\}\cdot 1\left\{U_{i2}-\delta_{i2}-\frac{1}{2}\kappa_i>w_{i2}>L_{i2}-\delta_{i2}-\frac{1}{2}\kappa_i\right\}.\end{aligned}$$

This will have positive probability provided that

$$P\left(L_{i1}-\delta_{i1}+\frac{1}{2}\kappa_i<U_{i2}-\delta_{i2}-\frac{1}{2}\kappa\right)>0$$

which follows from Lemma 6.

Lemma 6 *If (U_1,V_1) and (U_2,V_2) are two independent draws of a random vector (U,V) with $P\left(U<V\right)=1$, then there exists a $k>0$ such that for $0<\kappa<k$, $P\left(U_1<V_2-\kappa\right)>0$.*

Proof. Since $V-U>0$ with probability 1, there exists an $m>0$ such that $P\left(V-U>m\right)\ge\frac{1}{2}$. Now consider the space $\{(u,v):v-u>m\}$. This can be divided into a countable number of regions, $\mathcal{A}_k$, such that for $(u_1,v_1),(u_2,v_2)\in\mathcal{A}_k$, $|u_1-u_2|<\frac{m}{2}$ and $|v_1-v_2|<\frac{m}{2}$. Since $P\left(V-U>m\right)=\frac{1}{2}$, at least one of these regions, $\mathcal{A}_k$, must have positive probability. Hence there is positive probability that (U_1,V_1) and (U_2,V_2) in the statement of the lemma both belong to this $\mathcal{A}_k$, in which case $U_1<V_1-m<V_2-\frac{m}{2}$. ■

8 Appendix 2: Empirical Results

Table 1: Marginal tax rates before and after implementation of the 1987 tax reform

Before Reform		After Reform			
Tax bracket	Earnings + Cap inc	Tax bracket	Earnings	inc < 0	inc. $> 0^{(1)}$
0-113	$M + 19.75$	$0 - 130$	$M + 22.00$	$M + 22.00$	$M + 22.00$
113-186	$M + 34.15$	$130 - 200$	$M + 28.00$	$M + 22.00$	$M + 28.00$
186-	$M + 44.95$	$200-$	$M + 40.00$	$M + 22.00$	$M + 28.00$
Tax ceiling	73	Tax ceiling	$68.00/56.00^{(2)}$		56.00

Note: M is the local government tax rate. Threshold values for the tax brackets are given in 1000 DKK. Thresholds are adjusted yearly. Threshold values used in the table are for 1986 (before the reform) and 1987 (after the reform). The marginal tax rates refer to personal income (as opposed to household income).

(1) The tax brackets for positive net capital income refer to the sum of earnings and positive net capital income.

After the reform positive capital income is taxed progressively up to the first threshold, 130,000 DKK. For a married couple the progression threshold is 260,000 based on the sum of their joint positive net capital income and earnings.

(2) The 68 percent tax ceiling applies only to the top bracket.

Table 2:. Random and Fixed Effects Censored Regression Estimates of the Portfolio Share of Bonds in Financial Wealth.

	Fixed Effects	Random Effects
MTR capital income	−0.130	−0.206
	(0.071)	(0.045)
Ln(Financial Assets)	0.177	0.085
	(0.008)	(0.003)
D85	−0.214	−0.157
	(0.006)	(0.006)
D86	−0.314	−0.248
	(0.009)	(0.006)
D87	−0.318	−0.285
	(0.013)	(0.008)
D88	−0.383	−0.331
	(0.013)	(0.008)
Constant	—	0.047
		(0.046)
# households/observations	8,577 / 42,885	
# left/right censored obs	7,529 / 15,655	
Test of Parameter Equality (d.f.)	167 (6)	

Standard errors in parenthesis

Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

A Leverage-based Model of Speculative Bubbles **WP-08-01**
Gadi Barlevy

Displacement, Asymmetric Information and Heterogeneous Human Capital **WP-08-02**
Luojia Hu and Christopher Taber

BankCaR (Bank Capital-at-Risk): A credit risk model for US commercial bank charge-offs **WP-08-03**
Jon Frye and Eduard Pelz

Bank Lending, Financing Constraints and SME Investment **WP-08-04**
Santiago Carbó-Valverde, Francisco Rodríguez-Fernández, and Gregory F. Udell

Global Inflation **WP-08-05**
Matteo Ciccarelli and Benoît Mojon

Scale and the Origins of Structural Change **WP-08-06**
Francisco J. Buera and Joseph P. Kaboski

Inventories, Lumpy Trade, and Large Devaluations **WP-08-07**
George Alessandria, Joseph P. Kaboski, and Virgiliu Midrigan

School Vouchers and Student Achievement: Recent Evidence, Remaining Questions **WP-08-08**
Cecilia Elena Rouse and Lisa Barrow

Does It Pay to Read Your Junk Mail? Evidence of the Effect of Advertising on Home Equity Credit Choices **WP-08-09**
Sumit Agarwal and Brent W. Ambrose

The Choice between Arm's-Length and Relationship Debt: Evidence from eLoans **WP-08-10**
Sumit Agarwal and Robert Hauswald

Consumer Choice and Merchant Acceptance of Payment Media **WP-08-11**
Wilko Bolt and Sujit Chakravorti

Investment Shocks and Business Cycles **WP-08-12**
Alejandro Justiniano, Giorgio E. Primiceri, and Andrea Tambalotti

New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard **WP-08-13**
Thomas Klier and Joshua Linn

Realized Volatility **WP-08-14**
Torben G. Andersen and Luca Benzoni

Revenue Bubbles and Structural Deficits: What's a state to do? **WP-08-15**
Richard Mattoon and Leslie McGranahan

Working Paper Series *(continued)*

The role of lenders in the home price boom **WP-08-16**
Richard J. Rosen

Bank Crises and Investor Confidence **WP-08-17**
Una Okonkwo Osili and Anna Paulson

Life Expectancy and Old Age Savings **WP-08-18**
Mariacristina De Nardi, Eric French, and John Bailey Jones

Remittance Behavior among New U.S. Immigrants **WP-08-19**
Katherine Meckel

Birth Cohort and the Black-White Achievement Gap:
The Roles of Access and Health Soon After Birth **WP-08-20**
Kenneth Y. Chay, Jonathan Guryan, and Bhashkar Mazumder

Public Investment and Budget Rules for State vs. Local Governments **WP-08-21**
Marco Bassetto

Why Has Home Ownership Fallen Among the Young? **WP-09-01**
Jonas D.M. Fisher and Martin Gervais

Why do the Elderly Save? The Role of Medical Expenses **WP-09-02**
Mariacristina De Nardi, Eric French, and John Bailey Jones

Using Stock Returns to Identify Government Spending Shocks **WP-09-03**
Jonas D.M. Fisher and Ryan Peters

Stochastic Volatility **WP-09-04**
Torben G. Andersen and Luca Benzoni

The Effect of Disability Insurance Receipt on Labor Supply **WP-09-05**
Eric French and Jae Song

CEO Overconfidence and Dividend Policy **WP-09-06**
Sanjay Deshmukh, Anand M. Goel, and Keith M. Howe

Do Financial Counseling Mandates Improve Mortgage Choice and Performance? **WP-09-07**
Evidence from a Legislative Experiment
Sumit Agarwal,Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff

Perverse Incentives at the Banks? Evidence from a Natural Experiment **WP-09-08**
Sumit Agarwal and Faye H. Wang

Pay for Percentile **WP-09-09**
Gadi Barlevy and Derek Neal

The Life and Times of Nicolas Dutot **WP-09-10**
François R. Velde

Regulating Two-Sided Markets: An Empirical Investigation **WP-09-11**
Santiago Carbó Valverde, Sujit Chakravorti, and Francisco Rodriguez Fernandez

Working Paper Series *(continued)*

The Case of the Undying Debt **WP-09-12**
François R. Velde

Paying for Performance: The Education Impacts of a Community College Scholarship Program for Low-income Adults **WP-09-13**
Lisa Barrow, Lashawn Richburg-Hayes, Cecilia Elena Rouse, and Thomas Brock

Establishments Dynamics, Vacancies and Unemployment: A Neoclassical Synthesis **WP-09-14**
Marcelo Veracierto

The Price of Gasoline and the Demand for Fuel Economy:
Evidence from Monthly New Vehicles Sales Data **WP-09-15**
Thomas Klier and Joshua Linn

Estimation of a Transformation Model with Truncation,
Interval Observation and Time-Varying Covariates **WP-09-16**
Bo E. Honoré and Luojia Hu

Self-Enforcing Trade Agreements: Evidence from Antidumping Policy **WP-09-17**
Chad P. Bown and Meredith A. Crowley

Too much right can make a wrong: Setting the stage for the financial crisis **WP-09-18**
Richard J. Rosen

Can Structural Small Open Economy Models Account
for the Influence of Foreign Disturbances? **WP-09-19**
Alejandro Justiniano and Bruce Preston

Liquidity Constraints of the Middle Class **WP-09-20**
Jeffrey R. Campbell and Zvi Hercowitz

Monetary Policy and Uncertainty in an Empirical Small Open Economy Model **WP-09-21**
Alejandro Justiniano and Bruce Preston

Firm boundaries and buyer-supplier match in market transaction:
IT system procurement of U.S. credit unions **WP-09-22**
Yukako Ono and Junichi Suzuki

Health and the Savings of Insured Versus Uninsured, Working-Age Households in the U.S. **WP-09-23**
Maude Toussaint-Comeau and Jonathan Hartley

The Economics of "Radiator Springs:" Industry Dynamics, Sunk Costs, and
Spatial Demand Shifts **WP-09-24**
Jeffrey R. Campbell and Thomas N. Hubbard

On the Relationship between Mobility, Population Growth, and
Capital Spending in the United States **WP-09-25**
Marco Bassetto and Leslie McGranahan

The Impact of Rosenwald Schools on Black Achievement **WP-09-26**
Daniel Aaronson and Bhashkar Mazumder

Working Paper Series *(continued)*

Comment on "Letting Different Views about Business Cycles Compete" **WP-10-01**
Jonas D.M. Fisher

Macroeconomic Implications of Agglomeration **WP-10-02**
Morris A. Davis, Jonas D.M. Fisher and Toni M. Whited

Accounting for non-annuitization **WP-10-03**
Svetlana Pashchenko

Robustness and Macroeconomic Policy **WP-10-04**
Gadi Barlevy

Benefits of Relationship Banking: Evidence from Consumer Credit Markets **WP-10-05**
Sumit Agarwal, Souphala Chomsisengphet, Chunlin Liu, and Nicholas S. Souleles

The Effect of Sales Tax Holidays on Household Consumption Patterns **WP-10-06**
Nathan Marwell and Leslie McGranahan

Gathering Insights on the Forest from the Trees: A New Metric for Financial Conditions **WP-10-07**
Scott Brave and R. Andrew Butters

Identification of Models of the Labor Market **WP-10-08**
Eric French and Christopher Taber

Public Pensions and Labor Supply Over the Life Cycle **WP-10-09**
Eric French and John Jones

Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash **WP-10-10**
Luca Benzoni, Pierre Collin-Dufresne, and Robert S. Goldstein

Prenatal Sex Selection and Girls' Well-Being: Evidence from India **WP-10-11**
Luojia Hu and Analía Schlosser

Mortgage Choices and Housing Speculation **WP-10-12**
Gadi Barlevy and Jonas D.M. Fisher

Did Adhering to the Gold Standard Reduce the Cost of Capital? **WP-10-13**
Ron Alquist and Benjamin Chabot

Introduction to the *Macroeconomic Dynamics*:
Special issues on money, credit, and liquidity **WP-10-14**
Ed Nosal, Christopher Waller, and Randall Wright

Summer Workshop on Money, Banking, Payments and Finance: An Overview **WP-10-15**
Ed Nosal and Randall Wright

Cognitive Abilities and Household Financial Decision Making **WP-10-16**
Sumit Agarwal and Bhashkar Mazumder

Complex Mortgages **WP-10-17**
Gene Amromin, Jennifer Huang, Clemens Sialm, and Edward Zhong

Working Paper Series *(continued)*

The Role of Housing in Labor Reallocation **WP-10-18**
Morris Davis, Jonas Fisher, and Marcelo Veracierto

Why Do Banks Reward their Customers to Use their Credit Cards? **WP-10-19**
Sumit Agarwal, Sujit Chakravorti, and Anna Lunn

The impact of the originate-to-distribute model on banks
before and during the financial crisis **WP-10-20**
Richard J. Rosen

Simple Markov-Perfect Industry Dynamics **WP-10-21**
Jaap H. Abbring, Jeffrey R. Campbell, and Nan Yang

Commodity Money with Frequent Search **WP-10-22**
Ezra Oberfield and Nicholas Trachter

Corporate Average Fuel Economy Standards and the Market for New Vehicles **WP-11-01**
Thomas Klier and Joshua Linn

The Role of Securitization in Mortgage Renegotiation **WP-11-02**
Sumit Agarwal, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff

Market-Based Loss Mitigation Practices for Troubled Mortgages
Following the Financial Crisis **WP-11-03**
Sumit Agarwal, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff

Federal Reserve Policies and Financial Market Conditions During the Crisis **WP-11-04**
Scott A. Brave and Hesna Genay

The Financial Labor Supply Accelerator **WP-11-05**
Jeffrey R. Campbell and Zvi Hercowitz

Survival and long-run dynamics with heterogeneous beliefs under recursive preferences **WP-11-06**
Jaroslav Borovička

A Leverage-based Model of Speculative Bubbles (Revised) **WP-11-07**
Gadi Barlevy

Estimation of Panel Data Regression Models with Two-Sided Censoring or Truncation **WP-11-08**
Sule Alan, Bo E. Honoré, Luojia Hu, and Søren Leth–Petersen