Estimation of dynamic panel threshold model using Stata

Myung Hwan Seo
Department of Economics and Institute of Economic Research
Seoul National University
Seoul, Korea
myunghseo@snu.ac.kr

Sueyoul Kim
Department of Economics
University of Maryland
College Park, MD
shelkim@econ.umd.edu

Young-Joo Kim
Department of Economics
Hongik University
Seoul, Korea
y.j.kim@hongik.ac.kr

Abstract. In this article, we develop a command, `xthenreg`, that implements the first-differenced generalized method of moments estimation of the dynamic panel threshold model that Seo and Shin (2016, *Journal of Econometrics* 195: 169–186) proposed. Furthermore, we derive the asymptotic variance formula for a kink-constrained generalized method of moments estimator of the dynamic threshold model and provide an estimation algorithm. We also propose a fast bootstrap algorithm to implement the bootstrap for the linearity test. We illustrate the use of `xthenreg` through a Monte Carlo simulation and an economic application.

Keywords: st0573, xthenreg, dynamic panel threshold model

1 Introduction

The panel model with threshold effects in Hansen (1999) has been widely used in empirical research. Hansen’s fixed-effects estimator has been applied to applications on the investment decision of firms under financial constraints, the relation between a fiscal deficit and economic growth (Adam and Bevan 2005), the relation between inflation and growth (Khan and Ssnhadji 2001), and others. The threshold effect in the model allows for the asymmetric effect of the exogenous variables, depending on whether the threshold variable is above or below the unknown threshold. The threshold variable is typically dictated by the economic model. For instance, in the investment decision problem, the size of the firm is often considered a candidate threshold variable. Wang (2015) developed a command, `xthreg`, that computes Hansen’s estimator.

Hansen’s (1999) model is static, and his fixed-effects estimator requires the covariates to be strongly exogenous for the estimator to be consistent. However, strong exogeneity can be restrictive in many real applications. Thus, the model has been extended to the dynamic panel model with a potentially endogenous threshold variable as proposed by Seo and Shin (2016). Their model allows lagged dependent variables and endogenous
covariates. Indeed, various applications of Hansen’s fixed-effects estimation can benefit from dynamic modeling. For instance, the investment decision depends on the previous period’s investment, and the panel threshold autoregressive model is another example of dynamic models.

We developed commands for the first-differenced generalized method of moments (GMM) estimators and the associated asymptotic variance estimator that are proposed by Seo and Shin (2016) as well as linearity testing for the presence of a threshold effect. While the previous command xthreg computes the fixed-effect estimator and thus is not consistent under this general setting, our command xthenreg produces consistent and asymptotically normal estimates.

We also propose a computationally more attractive bootstrap algorithm to implement the linearity test than the nonparametric independent and identically distributed bootstrap that was originally proposed by Seo and Shin (2016). Furthermore, we present a constrained GMM estimator that reflects the kink restriction that has become more popular in recent years—see, for example, Zhang, Zhou, and Jiang (2017)—along with its asymptotic variance formula and a consistent estimator.

This article is organized as follows: Section 2 introduces the dynamic threshold panel model and the first-differenced GMM estimator. It also presents the asymptotic variance formula for a kink-constrained estimator and a bootstrap algorithm for the linearity test. Section 3 explains the command xthenreg. Sections 4 and 5 illustrate its use through Monte Carlo simulations and an application. Section 6 concludes.

2 Model

The dynamic panel threshold model is given by

\[ y_{it} = x'_i t \beta + (1, x'_i t) \delta 1 \{q_{it} > \gamma \} + \mu_i + \varepsilon_{it} \quad i = 1, \ldots, n; \ t = 1, \ldots, T \]

where \( x_{it} \) may include lagged dependent variables and \( q_{it} \) is the threshold variable. We assume \( T \) is fixed while the sample size \( n \) grows to infinity. Thus, we remove the incidental parameter \( \mu_i \) with the first-difference transformation and estimate the unknown parameters \( \theta = (\beta', \delta', \gamma)' \) through the GMM. The following describes the GMM method from Seo and Shin (2016).

Specifically, set an \( l \)-dimensional vector of instrument variables, \( (z_{i0}', \ldots, z_{iT}')' \), from the lagged variables and exogenous variables, where \( 2 < t_0 \leq T \). Next, construct the sample moment

\[ \bar{g}_n(\theta) = g_{1n} - g_{2n}(\gamma) (\beta', \delta')' = \frac{1}{n} \sum_{i=1}^{n} g_{1i} - \frac{1}{n} \sum_{i=1}^{n} g_{2i}(\gamma) (\beta', \delta')' \]

where

\[ g_{1i} = \begin{pmatrix} z_{i_{t_0}} \Delta y_{i_{t_0}} \\ \vdots \\ z_{i_T} \Delta y_{i_T} \end{pmatrix}, \ g_{2i}(\gamma) = \begin{pmatrix} z_{i_{t_0}} (\Delta x_{i_{t_0}}', 1_{i_{t_0}}(\gamma)' X_{i_{t_0}}) \\ \vdots \\ z_{i_T} (\Delta x_{i_T}', 1_{i_T}(\gamma)' X_{i_T}) \end{pmatrix} \]
with $\Delta$ signifying the first-difference operator and
\[
X_{it} = \begin{pmatrix} 1, x'_{it} \\ 1, x'_{i,t-1} \end{pmatrix}
\]
and $1_{it}(\gamma) = \begin{pmatrix} 1 \{ q_{it} > \gamma \} \\ -1 \{ q_{it-1} > \gamma \} \end{pmatrix}$

Then, introduce the GMM criterion function with a weight matrix $W_n$,
\[
J_n(\theta) = J_n(\theta)' W_n J_n(\theta)
\]
which is minimized to produce a GMM estimate, $\hat{\theta}$.

The minimization is done by the grid search because for each fixed $\gamma$ the model becomes a linear panel with a fixed effect, which yields the closed-form solution
\[
\{ \hat{\beta}(\gamma)', \hat{\delta}(\gamma) ' \} = \left\{ g_{2n}(\gamma)' W_n g_{2n}(\gamma) \right\}^{-1} g_{2n}(\gamma)' W_n g_{1n}
\]
and the criterion function $J_n(\theta)$ is a step function over $\gamma$ with at most $nT$ jumps. However, note that this algorithm is different from splitting the sample in two and applying the linear GMM for each partitioned sample.

For the weight matrix, either $W_n = I_l$ or
\[
W_n = \begin{pmatrix}
\frac{2}{n} \sum_{i=1}^{n} z_{it0} z'_{it0} & \frac{-1}{n} \sum_{i=1}^{n} z_{it0} z'_{it0+1} & 0 & \cdots \\
\frac{-1}{n} \sum_{i=1}^{n} z_{it0+1} z'_{it0} & \frac{2}{n} \sum_{i=1}^{n} z_{it0+1} z'_{it0+1} & \cdots \\
0 & \cdots & \cdots & \cdots \\
\vdots & \cdots & \cdots & \cdots \\
\frac{-1}{n} \sum_{i=1}^{n} z_{i'T-1} z'_{i'T} & \frac{2}{n} \sum_{i=1}^{n} z_{i'T} z'_{i'T} & \cdots \\
\end{pmatrix}^{-1}
\]
was proposed in the first step, and it is updated to
\[
W_n = \left( \frac{1}{n} \sum_{i=1}^{n} \hat{g}_i \hat{g}_i' - \frac{1}{n^2} \sum_{i=1}^{n} \hat{g}_i \sum_{i=1}^{n} \hat{g}_i' \right)^{-1}
\]
where $\hat{g}_i = (\hat{\Delta} \varepsilon_{it0} z'_{it0}, \ldots, \hat{\Delta} \varepsilon_{iT} z'_{iT})'$ and $\hat{\Delta} \varepsilon_{it}$ is the residual from the first-step estimation.

Seo and Shin (2016) showed that under suitable regularity conditions,\(^1\) the GMM estimator is asymptotically normal. Specifically,
\[
\sqrt{n} \begin{pmatrix}
\hat{\beta} - \beta_0 \\
\hat{\delta} - \delta_n \\
\end{pmatrix} \xrightarrow{d} \mathcal{N} \left\{ 0, (G' \Omega^{-1} G)^{-1} \right\}
\]
where $G = \{ G_\beta, G_\delta(\gamma_0), G_\gamma(\gamma_0) \}$ with
\[
G_\beta = \begin{bmatrix}
- E \left( z_{it0} \Delta x'_{it0} \right) \\
\vdots \\
- E \left( z_{iT} \Delta x'_{iT} \right)
\end{bmatrix}_{l \times k_1}, \quad G_\delta(\gamma) = \begin{bmatrix}
- E \left( z_{it0} 1_{it0} (\gamma)' X_{it0} \right) \\
\vdots \\
- E \left( z_{iT} 1_{iT} (\gamma)' X_{iT} \right)
\end{bmatrix}_{l \times (k_1 + 1)}
\]
\[
G_\gamma(\gamma_0) = \begin{bmatrix}
- E \left( z_{it0} 1_{it0} (\gamma_0)' X_{it0} \right) \\
\vdots \\
- E \left( z_{iT} 1_{iT} (\gamma_0)' X_{iT} \right)
\end{bmatrix}_{l \times (k_1 + 1)}
\]
\footnote{One of the conditions allows for $\delta_0$ to be both fixed and shrinking toward zero at $n^{-\alpha}$.}
and
\[
\begin{bmatrix}
E_{t_0-1} \{ \z_{it_0} (1, \x'_{it_0-1}) | \gamma \} \ p_{t_0-1} (\gamma) - E_{t_0} \{ \z_{it_0} (1, \x'_{it_0}) | \gamma \} \ p_{t_0} (\gamma) \\
\vdots \\
E_{T-1} \{ \z_{iT} (1, \x'_{IT-1}) | \gamma \} \ p_{T-1} (\gamma) - E_{T} \{ \z_{iT} (1, \x'_{IT}) | \gamma \} \ p_{T} (\gamma)
\end{bmatrix} \delta_0
\]
where \( E_\cdot (\cdot | \gamma) \) denotes the conditional expectation given \( q_{it} = \gamma \) and \( p_\cdot (\cdot) \) denotes the density of \( q_{it} \).

The estimation of the asymptotic variance is standard; that is,
\[
\hat{\Omega} (\theta) = \frac{1}{n} \sum_{i=1}^{n} g_i (\theta) g_i (\theta)' - \frac{1}{n} \sum_{i=1}^{n} g_i (\theta) \frac{1}{n} \sum_{i=1}^{n} g_i (\theta)
\]
where \( g_i (\theta) = g_{it} + g_{2it} (\gamma) (\beta^i, \delta^i)' \) and
\[
\hat{G}_\beta = \begin{bmatrix}
- \frac{1}{n} \sum_{i=1}^{n} \z_{it_0} \Delta x'_{it_0} \\
\vdots \\
- \frac{1}{n} \sum_{i=1}^{n} \z_{iT} \Delta x'_{IT}
\end{bmatrix},
\hat{G}_\delta (\gamma) = \begin{bmatrix}
- \frac{1}{n} \sum_{i=1}^{n} \z_{it_0} \mathbb{1}_{it_0} (\gamma)' \x_{it_0} \\
\vdots \\
- \frac{1}{n} \sum_{i=1}^{n} \z_{iT} \mathbb{1}_{iT} (\gamma)' \x_{iT}
\end{bmatrix}
\]
\[
\hat{G}_\gamma (\theta) = \begin{bmatrix}
\frac{1}{nh} \sum_{i=1}^{n} \z_{it_0} \left\{ \left( 1, \x'_{it_0-1} \right)' K \left( \frac{2q_{it_0-1} - h}{h} \right) - \left( 1, \x'_{it_0} \right)' K \left( \frac{2q_{it_0} - h}{h} \right) \right\} \delta \\
\vdots \\
\frac{1}{nh} \sum_{i=1}^{n} \z_{iT} \left\{ \left( 1, \x'_{IT-1} \right)' K \left( \frac{2q_{IT-1} - h}{h} \right) - \left( 1, \x'_{IT} \right)' K \left( \frac{2q_{IT} - h}{h} \right) \right\} \delta
\end{bmatrix}
\]
which is the Nadaraya–Watson kernel estimator for some kernel \( K \) and bandwidth \( h \), for example, the Gaussian kernel and Silverman’s rule of thumb. We plug in \( \theta = \hat{\theta} \).

### 2.1 Kink model

Although the threshold model typically implies the presence of a discontinuity of the regression function, it may mean the presence of a kink, not a jump, if \( (1, \x'_{it}) \delta = \kappa (q_{it} - \gamma) \) for some \( \kappa \). It happens when one element of \( \x_{it} \) is \( q_{it} \) with the coefficient \( \kappa \) and the first element of \( \delta \) equals to \( -\gamma \kappa \). Under these restrictions, the model becomes
\[
y_{it} = \x'_{it} \beta + \kappa (q_{it} - \gamma) \mathbb{1} \{ q_{it} > \gamma \} + \alpha_i + \varepsilon_{it} \quad i = 1, \ldots, n; \ t = 1, \ldots, T
\]

Even when the true model is a kink model, it is shown that the asymptotic distribution of the GMM estimator in the preceding section is valid. This contrasts with the least-squares estimator for the linear regression, for which Hidalgo, Lee, and Seo (2019) have shown that the cube root phenomenon appears.
The asymptotic distribution of the constrained GMM estimator of \((\beta, \kappa, \gamma)\) that imposes the kink restriction can also be derived for the same reason as Seo and Shin (2016). Specifically, the asymptotic variance is given by redefining \(G = (G_\beta, G_\kappa, G_\gamma)\), where \(G_\beta\) is the same as above and

\[
G_\kappa = \left( \begin{array}{c} \mathbb{E}z_{i0} [(q_{i0} - \gamma_0) 1\{q_{i0} > \gamma_0\} - (q_{i0-1} - \gamma_0) 1\{q_{i0-1} > \gamma_0\}] \\ \vdots \\ \mathbb{E}z_{iT} [(q_{iT} - \gamma_0) 1\{q_{iT} > \gamma_0\} - (q_{iT-1} - \gamma_0) 1\{q_{iT-1} > \gamma_0\}] \end{array} \right)
\]

\[
G_\gamma = \kappa_0 \left( \begin{array}{c} \mathbb{E}z_{i0} (1\{q_{i0-1} > \gamma_0\} - 1\{q_{i0} > \gamma_0\}) \\ \vdots \\ \mathbb{E}z_{iT} (1\{q_{iT-1} > \gamma_0\} - 1\{q_{iT} > \gamma_0\}) \end{array} \right)
\]

The estimation of these terms is analogous to that of \(G_\delta\) and \(G_\gamma\) in the preceding section.

### 2.2 Bootstrap test of linearity

This section proposes a fast bootstrap algorithm to test for the presence of the threshold effect, that is, the null hypothesis

\[ H_0: \delta_0 = 0 \quad \text{for any } \gamma \in \Gamma \]

where \(\Gamma\) denotes the parameter space for \(\gamma\), against the alternative hypothesis

\[ H_1: \delta_0 \neq 0 \quad \text{for some } \gamma \in \Gamma \]

A standard approach is to use a supremum-type statistic to take care of the loss of identification under the null; that is,

\[
\sup W = \sup_{\gamma \in \Gamma} W_n (\gamma)
\]

where \(W_n (\gamma)\) is the standard Wald statistic for each fixed \(\gamma\); that is,

\[
W_n (\gamma) = n\hat{\delta} (\gamma)' \hat{\Sigma}_\delta (\gamma)^{-1} \hat{\delta} (\gamma)
\]

where \(\hat{\delta} (\gamma)\) is the GMM estimator of \(\delta\) for a given \(\gamma\) and

\[
\hat{\Sigma}_\delta (\gamma) = R \left\{ \hat{V}_s (\gamma)' \hat{V}_s (\gamma) \right\}^{-1} R'
\]

is a consistent asymptotic variance estimator, where \(R = (0_{(k_1 + 1) \times k_1}, I_{k_1+1})\), and \(\hat{V}_s (\gamma) = \hat{\Omega}^{-1/2} [\hat{G}_\delta, \hat{G}_\delta (\hat{\theta} (\gamma))]\).

Because the asymptotic distribution is not pivotal, we propose a bootstrap algorithm, which is faster than the independent and identically distributed bootstrap proposed in Seo and Shin (2016). Specifically:
Estimation of dynamic panel threshold model using Stata

1. Draw \{\eta_i\}_i=1^n independently from the standard normal.

2. Recall the definition of \(\hat{\delta}(\gamma)\) in (1), and compute \(\hat{\delta}(\gamma)^*\) by replacing \(\Delta y_{it}\) with \(\Delta y_{it}^* = \Delta \tilde{\varepsilon}_{it} \eta_i\), where \(\Delta \tilde{\varepsilon}_{it} = \Delta y_{it} - \Delta x_{it}' \hat{\beta} - \hat{\delta}' X_{it} \hat{1}_{a(t)}(\gamma)\) is the residual from the original sample.

3. Compute a bootstrap statistic \(W_n^*(\gamma) = n \hat{\delta}(\gamma)^* \widehat{\Sigma}_n(\gamma)^{-1} \hat{\delta}(\gamma)^*\) and its supremum over \(\Gamma\) to get \(\text{sup} W^*\).

4. Repeat steps 1–3 \(B\) times, and compute the empirical proportion of \(\text{sup} W^*\) bigger than \(\text{sup} W\).

3 The xthenreg command

3.1 Syntax

\texttt{xthenreg depvar indepvars [if] [in] [, static kink endogenous(varlist) inst(varlist) grid num(integer) trim rate(real) h0(real) boost(integer)]}

where \texttt{depvar} is the dependent variable and \texttt{indepvars} are the independent variables. Detailed instructions for users are as follows:

1. \texttt{xtset} your data, and sort your variables by i) panel variable and ii) time variable.

2. Your inputs should be \texttt{y q x1 x2 ...}, where \(q\) is the threshold variable and \(x1 x2 ...\) are other independent variables.

3. When there are endogenous independent variables, set the \texttt{endogenous()} option. For example, if \(x1\) is exogenous and \(x2\) is endogenous, the input must be \texttt{y q x1, endogenous(x2)}. If you run the dynamic model (default), \(L.y\) is automatically contained in independent variables.

4. Default instruments are the independent variable itself for an exogenous independent variable and deeper lags (as in the Arellano–Bond estimator) for an endogenous independent variable such as \(L.y\). You can use additional instrumental variables with the \texttt{inst()} option. They must be time variant and contain no missing value.

5. The threshold variable \(q\) can also appear in covariates. But in this case, one has to write \(q\) twice. For instance, suppose the model is

\[ y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 q_t + (\delta_0 + \delta_1 y_{t-1} + \delta_2 q_t) \mathbb{1}_{\{q_t > \gamma\}} + \varepsilon_t \]

then \texttt{xthenreg y q q} would produce proper estimate results.

6. Strongly balanced panel data are required, and the \texttt{moremata} package is also required because \texttt{xthenreg} uses the \texttt{mm_quantile()} function.
3.2 Options

**static** specifies a static model; the default model is dynamic. Unlike the dynamic model, the static model does not automatically include \( L.y \) as an independent variable.

**kink** specifies a kink model.

**endogenous(varlist)** specifies endogenous independent variables. The endogenous variables must be excluded from the list of independent variables before the comma. For example, if \( x1 \) and \( x2 \) are independent variables and \( x2 \) is endogenous, you must use a command like `xthenreg y q x1, endogenous(x2)`, not `xthenreg y q x1 x2, endogenous(x2)`.

**inst(varlist)** specifies the list of additional instrumental variables.

**grid_num(integer)** determines the number of grid points to estimate the threshold \( \gamma \). The default is `grid_num(100)`.

**trim_rate(real)** determines the trim rate when constructing a grid for estimating \( r \). It must be a positive real number smaller than 1. For example, if the trim rate is 0.2, the start and end points of the grid are the 0.1 quantile and 0.9 quantile of observed variable \( q \), respectively. The default is `trim_rate(0.4)`.

**h_0(real)** determines a parameter for Silverman’s rule of thumb, which is used to determine the bandwidth of the Nadaraya–Watson kernel for covariance matrix estimation. It must be a positive real number. The default is `h_0(1.5)`.

**boost(integer)** is the number of bootstrap iterations for the linearity test. The default is `boost(0)`. If you give a positive integer for this option, the bootstrapping linearity test in Seo and Shin (2016) is executed. However, note that this test requires a lot of computation time.
Estimation of dynamic panel threshold model using Stata

3.3 Stored results

xthenreg stores the following in e():

Scalars
- e(N) number of units of panel data
- e(T) time length of panel data
- e(boots_p) p-value for bootstrap linearity test; -1 if the test is not used
- e(grid) number of grid points used
- e(trim) trim rate for grid search
- e(bs) number of bootstrap iterations

Macros
- e(depvar) name of dependent variable
- e(indepvar) name of independent variable or variables
- e(properties) b V
- e(zx) name of instrumental variables
- e(qx) name of threshold variable

Matrices
- e(b) estimates of coefficients
- e(V) estimate of covariance matrix
- e(CI) 95% asymptotic confidence interval for b

4 Monte Carlo experiments

In this section, we illustrate the finite sample performance of the bootstrap linearity test. Some estimation simulations were performed in Seo and Shin (2016). Here the model under $H_0$ is linear; that is, $\delta = 0$. We consider the following data-generating process. Specific coefficient values differ across simulations.

$$y_{i,t} = \beta_1 y_{i,t-1} + \beta_2 x_{i,t} + (\delta_0 + \delta_1 y_{i,t-1} + \delta_2 x_{i,t})I(x_{i,t} > 0) + \epsilon_{i,t}$$

We summarize more details of our simulation design in the following table:

| Parameter | Definition                  | Value |
|-----------|----------------------------|-------|
| N         | Cross-sectional sample size | 500   |
| T         | Time periods               | 12    |
| #grid     | Number of grid points      | 100   |
| #iter     | Number of iterations       | 500   |
| $\alpha$  | Significance level         | 0.05  |

Moreover, $x_{i,t}$ and $\epsilon_{i,t}$ were drawn independently from the centered normal distribution with a standard deviation of 1 and 0.25, respectively. For each iteration, we calculate the bootstrapped $supW^*$ once following, for example, Giacomini, Politis, and White (2013). Consequently, we obtain 500 simulated $supW$ and $supW^*$ statistics. With these, we compute the bootstrap critical value, which is the empirical $(1 - \alpha) 100$-percentile of those 500 $supW^*$ statistics, and the rejection probability for the given $\alpha$, which is the proportion of 500 $supW$ statistics bigger than the bootstrap critical value.
4.1 Test size

Here we impose \((\beta_1, \beta_2, \delta_0, \delta_1, \delta_2) = (0.5, 0.8, 0.0, 0.0, 0.0)\) so that \(H_0: \delta = 0\) holds. This implies the true underlying model is linear. The simulated rejection probability was 0.066, which is reasonably close to the true \(\alpha = 0.05\). Empirical distributions of \(\text{supW}\) and \(\text{supW}^*\) are as follows:

![Distribution of \(\text{sup W}\) under \(H_0\), i.e. \(\delta = 0\)](image)

Figure 1. Original and bootstrapped sup-Wald statistics under \(H_0\)

4.2 Test power

Here we tested three sets of coefficient choices, maintaining \(H_1: \delta \neq 0\) holds:

| Parameters \((\beta_1, \beta_2, \delta_0, \delta_1, \delta_2)\) | Values \((0.5, 0.8, 0.0, -0.5, 0.0)\) | Rejection probability |
|---|---|---|
| \((0.5, 0.0, 0.0, -0.5, 0.0)\) | 0.54 |
| \((0.5, 0.0, 0.0, -0.9, 0.0)\) | 0.96 |

We observe that our test has significant power to reject \(H_0\) when \(H_1\) is true, especially when the true \(\delta\) is sufficiently far from zero.
Estimation of dynamic panel threshold model using Stata

5 Application

We apply our method to evaluate the effect of obesity on work hours. Obesity is measured with body mass index (BMI), which is someone’s weight in kilograms divided by height in meters squared. Individuals with a BMI between 25 and 30 are considered to be overweight, and individuals with a BMI of 30 or higher are considered to be obese. Using data from the British Cohort Study, which can be accessed through the UK Data Service, and the methods described in the earlier section, we examine how BMI is associated with work hours. For more detailed discussion, see Kim (2019).

In this example, we consider work hours and BMI of male workers using the following model with a kink in BMI:

\[
y_{it} = \beta_0 + x_{it}\beta_1 + q_{it}\beta_2 + \delta(q_{it} - \gamma)1\{q_{it} \geq \gamma\} + \alpha_i + \epsilon_{it}
\]

We present work hours as \(y_{it}\) for an individual \(i\) for a period \(t\), \(x_{it}\) as family size, and \(q_{it}\) as BMI. We have two-period panel data (\(t = 1, 2\)) and take the first difference as follows to remove \(\alpha_i\), the individual time-invariant characteristics that are associated with work hours:

\[
\Delta y_{i2} = \Delta x_{i2}\beta_1 + \Delta q_{i2}\beta_2 + \delta(q_{i2} - \gamma)1\{q_{i2} \geq \gamma\} - (q_{i1} - \gamma)\delta1\{q_{i1} \geq \gamma\} + \Delta \epsilon_{i2}
\]

To implement GMM estimation, we use instrumental variables birthweight (bweight), a worker’s own childhood BMI (bmic), and that worker’s parents’ BMIs (bmim, bmid) for BMI variables of \(\Delta q_{i2}, q_{i2}, q_{i1}\) in the first-differenced model.

After loading the data, we first need to declare that the data are panel. The default model for xthenreg is a dynamic model. Because we consider a static model, not a dynamic model, we use the static option. We also impose a kink in the model by using the kink option:

```
. use hour
. xtset ilabel time
    panel variable: ilabel (strongly balanced)
    time variable: time, 1 to 2
    delta: 1 unit
. xthenreg hour bmi hsize, endogenous(bmi) inst(bweight bm ic bmim bmid hsize) > kink static
```

2. See https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200001.
The information preceding the table is as follows: N is the total number of unique subjects, T is the number of time periods, and the Number of moment conditions is provided based on the choice of instruments.

Because we change the set of included and excluded instruments using the inst() option, the number of moment conditions varies accordingly.

```
xthenreg hour bmi hsize, endogenous(bmi) inst(bweight bmic bmim bmid)
> kink static
N = 768, T = 2
Panel Var. = ilabel
Time Var. = time
Number of moment conditions = 6

|        | Coef.  | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|--------|--------|-----------|-------|--------|---------------------|
| hour   |        |           |       |        |                     |
| hsize_b| 0.4964864 | 0.2774487 | 1.79  | 0.074  | -0.0473031 1.040276 |
| bmi_b  | -0.7783025 | 1.147452  | -0.68 | 0.498  | -3.027267 1.470662 |
| kink_slope | 2.527627 | 2.29367   | 1.10  | 0.270  | -1.967894 7.023139 |
| r      | 28.9816 | 6.002713  | 4.83  | 0.000  | 17.2165 40.7467   |
```

We can fit the model with a restriction on the sample:

```
xthenreg hour bmi hsize if region==1, endogenous(bmi)
> inst(bweight bmic bmim bmid hsize) kink static
5 sample(s) are ignored further due to missing values
N = 637, T = 2
Panel Var. = ilabel
Time Var. = time
Number of moment conditions = 7

|        | Coef.  | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|--------|--------|-----------|-------|--------|---------------------|
| hour   |        |           |       |        |                     |
| hsize_b| 0.5270078 | 0.3784874 | 1.39  | 0.164  | -.2148138 1.268829 |
| bmi_b  | -0.3205365 | 1.841899  | -0.17 | 0.862  | -3.930593 3.28952  |
| kink_slope | 2.444602 | 7.610258  | 0.32  | 0.748  | -12.47123 17.36043 |
| r      | 29.14014 | 17.36101  | 1.68  | 0.093  | -4.886813 63.16709 |
```

Next, we consider discontinuity in BMI effect without imposing a kink in the model.

\[
y_{it} = \beta_0 + x_{it}\beta_1 + q_{it}\beta_2 + (\delta_0 + x_{it}\delta_1 + q_{it}\delta_2)1\{q_{it} > \gamma\} + \alpha_i + \varepsilon_{it}
\]

By taking the first difference, we obtain the following model and fit it only with the static option.

\[
\Delta y_{i2} = \Delta x_{i2}\beta_1 + \Delta q_{i2}\beta_2 + (\delta_0 + x_{i2}\delta_1 + q_{i2}\delta_2)1\{q_{i2} > \gamma\} - (\delta_0 + x_{i1}\delta_1 + q_{i1}\delta_2)1\{q_{i1} > \gamma\} + \Delta \varepsilon_{i2}
\]
Estimation of dynamic panel threshold model using Stata

```
xthenreg hour bmi hsize, endogenous(bmi) inst(bweight bcmic bmim bmid hsize)
> static
N = 768, T = 2
Panel Var. = ilabel
Time Var. = time
Number of moment conditions = 7

|       | Coef.  | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|-------|--------|-----------|------|------|---------------------|
| hour  |        |           |      |      |                     |
| hsize_b | -2.093733 | 16.48417 | -0.13 | 0.899 | -34.40211 30.21464  |
| bmi_b  |  6.069513 | 19.04971 |  0.32 | 0.750 | -31.26724 43.40626 |
| cons_d | 106.3078  | 626.0473  |  0.17 | 0.865 | -3120.722 1333.338 |
| hsize_d|  4.898759 | 16.62541  |  0.29 | 0.768 | -27.68644 37.48395 |
| bmi_d  | -5.569615 | 27.32028  | -0.20 | 0.838 | -59.11639 47.97716 |
| r      |  25.64312 | 14.89532  |  1.72 | 0.085 | -3.551176 54.83741 |
```

6 Conclusion

In this article, we presented a set of algorithms to facilitate inference for the dynamic threshold panel model with two regimes. An interesting future research avenue is to develop an algorithm to determine the number of regimes. The commonly used sequential approach in, for example, Bai and Perron (1998), where the linearity test is applied to each subsample until the test cannot reject the linearity null, is less appealing for the first-differenced GMM estimation because of the lack of the oracle property in the threshold estimate. The penalized approaches in Lee, Seo, and Shin (2016) or Lee et al. (2018) may prove useful.

7 Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2017S1A5A8019707) and 2018 Hongik University Research Fund. Seo acknowledges financial support from the Center for National Competitiveness in the Institute of Economic Research.

8 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 19-3
. net install st0573 (to install program files, if available)
. net get st0573 (to install ancillary files, if available)
```
9 References

Adam, C. S., and D. L. Bevan. 2005. Fiscal deficits and growth in developing countries. *Journal of Public Economics* 89: 571–597.

Bai, J., and P. Perron. 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* 66: 47–78.

Giacomini, R., D. N. Politis, and H. White. 2013. A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory* 29: 567–589.

Hansen, B. E. 1999. Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics* 93: 345–368.

Hidalgo, J., J. Lee, and M. H. Seo. 2019. Robust inference for threshold regression models. *Journal of Econometrics* 210: 291–309.

Khan, M. S., and A. S. Ssnhadji. 2001. Threshold effects in the relationship between inflation and growth. *IMF Staff Papers* 48: 1–21.

Kim, Y.-J. 2019. The effect of weight on work hours. Unpublished manuscript.

Lee, S., Y. Liao, M. H. Seo, and Y. Shin. 2018. Oracle estimation of a change point in high-dimensional quantile regression. *Journal of the American Statistical Association* 113: 1184–1194.

Lee, S., M. H. Seo, and Y. Shin. 2016. The lasso for high dimensional regression with a possible change point. *Journal of the Royal Statistical Society, Series B* 78: 193–210.

Seo, M. H., and Y. Shin. 2016. Dynamic panels with threshold effect and endogeneity. *Journal of Econometrics* 195: 169–186.

Wang, Q. 2015. Fixed-effect panel threshold model using Stata. *Stata Journal* 15: 121–134.

Zhang, Y., Q. Zhou, and L. Jiang. 2017. Panel kink regression with an unknown threshold. *Economics Letters* 157: 116–121.

About the authors

Myung Hwan Seo is an associate professor of economics at Seoul National University in Seoul, Korea.

Sueyoul Kim is a PhD student in the Department of Economics at the University of Maryland.

Young-Joo Kim is an assistant professor of economics at Hongik University in Seoul, Korea.