Supplemental material for "*Regular and reverse Midastar models: Threshold autoregression with mixed frequency data*"

Kaiji Motegi[*] – Kobe University

John W. Dennis[†] – Institute for Defense Analyses

Seok Young Hong[‡] – Nanyang Technological University

This draft: February 18, 2025

In this supplemental material, we present full Monte Carlo simulations for the regular Midastar model. Since this whole document is concerned on the regular Midastar, we drop the word "regular" throughout. To make this document self-contained, we review and extend the set-up of the main paper. Let y be a low frequency target variable, and let x^* be a high frequency threshold variable. We observe $\{y_t\}_{t \in \mathbb{L}}$ and $\{x^*_t\}_{t \in \mathbb{H}}$, where $\mathbb{L} = \{1, 2, \dots, n\}$ is the set of low frequency time periods, $\mathbb{H}_t = \{t - 1 + 1/m, t - 1 + 2/m, \dots, t\}$ is the set of high frequency time periods within each $t \in \mathbb{L}$, and $\mathbb{H} = \cup_{t \in \mathbb{L}} \mathbb{H}_t$ is the entire set of high frequency time periods. The ratio of sampling frequencies is $m \in \{3, 12\}$.

In Section 1, the data generating process (DGP) is specified. In Section 2, we examine the performance of the profiling estimation of parameters. In Section 3, we inspect the empirical size and power of the wild-bootstrap tests for the no-threshold-effect hypothesis. In Section 4, we aggregate the high frequency threshold variable x^* into the low frequency level to gauge the impacts of temporal aggregation. In Section 5, we report and discuss rejection

---

[*] *Corresponding author.* Graduate School of Economics, Kobe University. Address: 2-1 Rokkodai-cho, Nada, Kobe, Hyogo 657-8501 Japan. Email: `motegi@econ.kobe-u.ac.jp`

[†] Institute for Defense Analyses (IDA). Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of IDA. Email: `jay.dennis@alumni.unc.edu`

[‡] School of Social Sciences & School of Physical and Mathematical Sciences, Nanyang Technological University. Email: `seokyoung.hong@ntu.edu.sg`

frequencies of the Diebold-Mariano tests for comparing the out-of-sample predictive accuracy of the models with the disaggregated versus aggregated threshold variable.

We use the following notation throughout: $\mathbb{R}$ is the set of real numbers, $\mathbb{N}$ is the set of natural numbers, $\lfloor a \rfloor$ is the largest integer not larger than $a \in \mathbb{R}$, the Euclidean norm of any k-dimensional vector $\boldsymbol{c} \in \mathbb{R}^k$ is denoted as $\|\boldsymbol{c}\| = (\boldsymbol{c}^\top \boldsymbol{c})^{1/2}$, $\#A$ is the number of elements of set A, and convergence in probability is denoted by $\xrightarrow{p}$.

1 Data generating process

Suppose that the DGP of the low frequency target variable y is the Midastar process with lag length $p = 1$:

$$y_t = \begin{cases} \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t & \text{if } \ x^*_{t-\frac{d_0}{m}} < \mu_0, \\ \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t & \text{if } \ x^*_{t-\frac{d_0}{m}} \geq \mu_0, \qquad t \in \mathbb{L}, \end{cases} \tag{1}$$

where the regression parameters in regime 1 are $\boldsymbol{\beta}_{10} = (\alpha_{10}, \phi_{10})^\top = (0, 0.2)^\top$, the delay parameter is $d_0 = 1$, the threshold parameter is $\mu_0 = 0$, and the error term is $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. The regression parameters in regime 2 are assumed to have a local drift toward $\boldsymbol{\beta}_{10}$:

$$\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}, \qquad \phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0},$$

where $\lambda_0 \neq 0$ and $\delta_0 \geq 0$. Let $\boldsymbol{\beta}_{20n} = (\alpha_{20n}, \phi_{20n})^\top$, then $\boldsymbol{\beta}_{20n} - \boldsymbol{\beta}_{10} = O(n^{-\delta_0})$. As elaborated by Andrews and Cheng (2012), δ_0 is a key quantity that determines the *identification category* of the nuisance parameter $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$. The smaller value of δ_0 implies the slower convergence of $\boldsymbol{\beta}_{20n}$ to $\boldsymbol{\beta}_{10}$ and hence the greater identifiability of $\boldsymbol{\gamma}_0$. We consider five values for δ_0:

Case #1 (non-identification): $\delta_0 \to \infty$, in which case $\boldsymbol{\gamma}_0$ is *unidentified* as $\boldsymbol{\beta}_{20n} = \boldsymbol{\beta}_{10}$ for any $n \in \mathbb{N}$.

Case #2 (weak identification): $\delta_0 = 0.75$, in which case $\boldsymbol{\gamma}_0$ is *weakly identified* as $\boldsymbol{\beta}_{20n} \to \boldsymbol{\beta}_{10}$ as $n \to \infty$ at a sufficiently fast rate.

Case #3 (weak identification, boundary case): $\delta_0 = 0.5$, which is the boundary case of weak identification.

Case #4 (semi-strong identification): $\delta_0 = 0.25$, in which case $\boldsymbol{\gamma}_0$ is *semi-strongly identified* as $\boldsymbol{\beta}_{20n} \to \boldsymbol{\beta}_{10}$ at a sufficiently slow rate.

Case #5 (strong identification): $\delta_0 = 0$, in which case $\boldsymbol{\gamma}_0$ is *strongly identified* as $\boldsymbol{\beta}_{20n} = \boldsymbol{\beta}_{10} + \lambda_0 \boldsymbol{\iota}_2$ for any $n \in \mathbb{N}$, where $\lambda_0 \neq 0$ and $\boldsymbol{\iota}_2 = (1,1)^\top$.

In the main paper, we considered only Case #1 (non-identification) and Case #5 (strong identification) out of the five cases to save space; Case #1 here matches Case 1 in the main paper, and Case #5 here matches Case 2 in the main paper. In this section, we additionally consider Cases #2–#4 to better understand the small-sample and large-sample behavior of our proposed methods.

The low frequency sample size is chosen from $n \in \{125, 250, 500, 1000, 2000\}$. In the main paper, we considered $n \in \{125, 250, 500, 1000\}$. We add the larger sample size $n = 2000$ to fully observe subtle differences across Cases #1–#5. For Case #1, the value of λ_0 is irrelevant since $\delta_0 \to \infty$. We set $\lambda_0 \in \{13.08, 3.91, 1.17, 0.35\}$ for Cases #2–#5, respectively. Given these values of λ_0, $\boldsymbol{\beta}_{20,125} \approx (0.35, 0.55)^\top$ for these four cases, which makes comparison easier. As n increases, $\boldsymbol{\beta}_{20n}$ converges to $\boldsymbol{\beta}_{10} = (0, 0.2)^\top$ at rates $n^{0.75}$, $n^{0.5}$, and $n^{0.25}$ in Cases #2–#4, respectively; $\boldsymbol{\beta}_{20n}$ is constant at $(0.35, 0.55)$ in Case #5.

These parameterizations are summarized in Table 1. Let $\Delta_{0n} = \|\boldsymbol{\beta}_{10} - \boldsymbol{\beta}_{20n}\|$ be the Euclidean norm between the true regression parameters in regimes 1 and 2. It serves as an overall measure of heterogeneity between the two regimes. In Case #1, $\Delta_{0n} = 0$ for all n. We have that $\Delta_{0,2000} \in \{0.062, 0.124, 0.247\}$ in Cases #2–#4, respectively. In Case #5, $\Delta_{0n} = 0.495$ for all n.

Suppose that the DGP of the high frequency threshold variable x^* is AR(1):

$$x_t^* = \psi_0 x_{t-\frac{1}{m}}^* + \nu_t^*, \qquad \nu_t^* \overset{i.i.d.}{\sim} \mathcal{N}(0,1), \qquad t \in \mathbb{H}.$$

Assume that $\{\epsilon_t\}_{t \in \mathbb{L}}$ and $\{\nu_t^*\}_{t \in \mathbb{H}}$ are mutually independent sequences. True values of the AR(1) coefficient of x^* is chosen from $\psi_0 \in \{0.3, 0.6, 0.9\}$. We considered $\psi_0 \in \{0.3, 0.9\}$ in the main paper, and add the intermediate case $\psi_0 = 0.6$ here for completeness.

2 Profiling estimation of parameters

The Midastar model with lag length $p = 1$ is specified as

$$y_t = \begin{cases} \alpha_1 + \phi_1 y_{t-1} + u_t & \text{if } \ x^*_{t-\frac{d}{m}} < \mu, \\ \alpha_2 + \phi_2 y_{t-1} + u_t & \text{if } \ x^*_{t-\frac{d}{m}} \geq \mu, \qquad t \in \mathbb{L}. \end{cases} \tag{2}$$

The choice space of the delay parameter d is $D = \{1, 2, 3\}$. Note that the value of d is in terms of high frequency time periods. When $d = m = 3$, model (2) coincides with the single-frequency TAR model with delay being one low frequency time period.

To construct the choice space of μ, it is of use to aggregate the m high frequency observations of x^* within each low frequency time period. We propose to take a simple average:

$$x_t = \frac{1}{m} \sum_{k=1}^{m} x^*_{t-1+\frac{k}{m}}, \qquad t \in \mathbb{L}. \tag{3}$$

Let $x_{[1]} \leq \cdots \leq x_{[n]}$ be a sorted version of $\{x_t\}_{t \in \mathbb{L}}$. In general, the space of μ is specified as

$$\mathcal{X}_{\kappa,n} = \left\{ x_{[\lfloor 0.5(1-\kappa)n \rfloor]}, \dots, x_{[\lfloor \{1-0.5(1-\kappa)\}n \rfloor]} \right\}, \tag{4}$$

where $\kappa \in [0, 1)$ signifies the fraction of $\#\mathcal{X}_{\kappa,n}$ to n. Each of the two regimes accounts for at least $50(1-\kappa)\%$ of the whole sample on average. Choosing a too large value for κ (e.g., $\kappa = 0.9$) may cause an identification problem in small samples. Following a well-known suggestion of Andrews (1993), we pick $\kappa = 0.7$ so that each regime accounts for at least 15% of the entire sample on average. The space of $\boldsymbol{\gamma} = (d, \mu)^\top$ is denoted as $\Gamma_{\kappa,n} = D \times \mathcal{X}_{\kappa,n}$; when there is no risk of confusion, we use a short-hand notation Γ instead of $\Gamma_{\kappa,n}$.

We fit model (2) to each of $J = 1000$ Monte Carlo samples generated from DGP (1), and estimate the regression parameters $\boldsymbol{\beta} = (\alpha_1, \phi_1, \alpha_2, \phi_2)^\top$ and the nuisance parameter $\boldsymbol{\gamma}$ via profiling as described in Section 3.1 of the main paper. The resulting estimator is denoted as $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$, and they are stacked as $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\boldsymbol{\gamma}}^\top)^\top$.

In Tables 2-11, we report the bias and standard deviation of $\hat{\boldsymbol{\theta}}$. The results with $m = 3$ and Cases #1–#5 are presented in Tables 2-6, respectively. Similarly, the results with $m = 12$ and Cases #1–#5 are presented in Tables 7-11. When computing the bias, the true value of the parameter is defined as $\boldsymbol{\theta}_{0n} = (\boldsymbol{\beta}_{0n}^\top, \boldsymbol{\gamma}_0^\top)^\top$, where $\boldsymbol{\beta}_{0n} = (\boldsymbol{\beta}_{10}^\top, \boldsymbol{\beta}_{20n}^\top)^\top$.

The bias of $\hat{\boldsymbol{\beta}}$ is negligibly small for all cases considered, suggesting that $\hat{\boldsymbol{\beta}}$ is an unbiased estimator of $\boldsymbol{\beta}_{0n}$. Further, the standard deviation of $\hat{\boldsymbol{\beta}}$ vanishes as the sample size n grows. The standard deviation of $\hat{\phi}_2$ in Case #1 with $(m, \psi_0) = (3, 0.3)$, for example, is $\{0.207, 0.152, 0.109, 0.081, 0.052\}$ for $n \in \{125, 250, 500, 1000, 2000\}$, respectively (Table 2). These results are in line with the implication of Theorem 1.(i) of the main paper that $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{0n} \xrightarrow{p} \mathbf{0}$ irrespective of identification categories.

The stronger threshold effect facilitates estimation of the regression parameter $\boldsymbol{\beta}$. Fixing $(m, \psi_0, n) = (3, 0.3, 250)$, the standard deviation of $\hat{\phi}_2$ is $\{0.152,\ 0.128,\ 0.119,\ 0.104,\ 0.093\}$ under Cases #1–#5, respectively (Tables 2-6). Another finding is that the ratio of sampling frequencies, m, does not affect the small sample performance of $\hat{\boldsymbol{\beta}}$. Fixing $(m, \psi_0, n) = (12, 0.3, 250)$, the standard deviation of $\hat{\phi}_2$ is $\{0.131, 0.113, 0.113, 0.094, 0.086\}$ under Cases #1–#5 (Tables 7-11). For each identification category, the standard deviations are almost the same between $m \in \{3, 12\}$. The robustness to a large m is a practical advantage of the Midastar model over standard Mixed Data Sampling (MIDAS) regression models. The number of regression parameters in Midastar does not depend on m, since we do not regress y onto x^* directly. This indirect specification makes Midastar free of parameter proliferation.

The persistence of x^* does not have substantial impacts on the small sample performance of $\hat{\boldsymbol{\beta}}$. Take $(m, n) = (3, 250)$ as an example, then the standard deviation of $\hat{\phi}_2$ under Case #1 is $\{0.152, 0.155, 0.166\}$ for $\psi_0 \in \{0.3, 0.6, 0.9\}$, respectively (Table 2). Similarly, the standard deviation of $\hat{\phi}_2$ under Case #5 is $\{0.093, 0.095, 0.088\}$ (Table 6).

The estimated nuisance parameter $\hat{\boldsymbol{\gamma}}$ does not converge to $\boldsymbol{\gamma}_0$ under Cases #1–#3 but it does converge under Cases #4–#5. Taking Case #1 with $(m, \psi_0) = (3, 0.3)$ as an instance, the bias of $\hat{\mu}$ is $\{0.026, 0.013, -0.007\}$ and the standard deviation of $\hat{\mu}$ is $\{0.501, 0.488, 0.500\}$ for $n \in \{500, 1000, 2000\}$, respectively (Table 2). The standard deviation is almost constant as n grows, suggesting that $\hat{\boldsymbol{\gamma}} = O_p(1)$ under Case #1. The standard deviations are constant in Cases #2–#3 too, although they are uniformly smaller than those in Case #1 (Tables 3-4). By contrast, the standard deviation of $\hat{\mu}$ is $\{0.204, 0.138, 0.102\}$ in Case #4 and $\{0.092, 0.045, 0.024\}$ in Case #5, monotonically approaching 0 as n grows (Tables 5-6).

Intuitively, the consistency of $\hat{\boldsymbol{\gamma}}$ should hold if there exist sufficiently strong threshold effects (i.e., if $\boldsymbol{\beta}_{20n}$ converges to $\boldsymbol{\beta}_{10}$ at a sufficiently slow rate). Our simulation results are indicative of non-convergence under non-identification and weak identification (i.e., $\hat{\boldsymbol{\gamma}} = O_p(1)$ for $\delta_0 \in [0.5, \infty)$), and indicative of convergence under semi-strong and strong identification (i.e., $\hat{\boldsymbol{\gamma}} \xrightarrow{p} \boldsymbol{\gamma}_0$ for $\delta_0 \in [0, 0.5)$). We observe similar patterns when the ratio of sampling frequencies is $m = 12$ (Tables 7-11). These results are in line with the implication of Theorem 1.(ii) of the main paper that $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{0n})$ is *not* asymptotically normal under non-identification and weak identification.

The greater persistence in x^* has negative impacts on the small sample performance of $\hat{\boldsymbol{\gamma}}$. This result is intuitive because the stronger persistence in a threshold variable should make signal on $\boldsymbol{\gamma}$ weaker. The standard deviation of $\hat{\mu}$ under Case #5 with $(m, n) = (3, 250)$ is $\{0.191, 0.244, 0.538\}$ for $\psi_0 \in \{0.3, 0.6, 0.9\}$, respectively (Table 6). In summary, the simulation results in Tables 2-11 are all reasonable and we can conclude that the profiling estimation with the Midastar model performs well in small and large samples.

3 Testing the no-threshold-effect hypothesis

Consider testing the no-threshold-effect hypothesis $H_0^* : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, where $\boldsymbol{\beta}_r = (\alpha_r, \phi_r)^\top$ is the regression parameters in regime $r \in \{1, 2\}$. As described in Section 3.2 of the main paper, we adopt the wild-bootstrap tests of Hansen (1996) since $\boldsymbol{\gamma}_0$ is unidentified under H_0^*. When computing actual and bootstrap test statistics, we use the heteroscedasticity-robust covariance matrix estimator, although the true error term ϵ is homoscedastic. For comparison, the sup-Wald, ave-Wald, exp-Wald, sup-LM, ave-LM, and exp-LM tests are all implemented. Rejection frequencies of each test are computed across $J = 1000$ Monte Carlo samples, where the nominal size is $a = 0.05$ and the number of bootstrap samples is $B = 500$.

Rejection frequencies for Cases #1–#5 are reported in Tables 12-16, respectively. Under Case #1, the empirical size of the six tests all converges to the nominal size as $n \to \infty$, confirming Theorem 2.(i) of the main paper. The Wald tests reject the correct null hypothesis too often when $n \leq 250$, although the over-rejections vanish as n grows. Taking the ave-Wald test with $(\psi_0, m) = (0.3, 3)$ as an example, the empirical size is $\{0.091, 0.060, 0.053\}$ for $n \in \{125, 250, 500\}$, respectively (Table 12). The sup-Wald and exp-Wald tests lead to even worse over-rejections than the ave-Wald test. The LM tests, by contrast, achieve accurate empirical size for all sample sizes considered. The empirical size of the ave-LM test with $(\psi_0, m) = (0.3, 3)$, for instance, is $\{0.045, 0.042, 0.047\}$. The sup-LM and exp-LM tests are comparable with the ave-LM test in terms of empirical size. Hence, it is advised to use the LM tests instead of the Wald tests to better control the type-I error rate in small samples.

Fixing the sample size $n \geq 250$, the empirical power of each test is lowest in Case #2 (weak identification) and highest in Case #5 (strong identification). This is intuitive since threshold effects become stronger as we proceed from #2 to #5. Besides, the empirical power of each test decreases as $n \to \infty$ under Case #2, stays constant under #3; and approaches 1 under #4–#5. The decreasing power in Case #2 reflects the sufficiently fast convergence of $\boldsymbol{\beta}_{20n}$ to $\boldsymbol{\beta}_{10}$, while the increasing power in #4–#5 reflects the sufficiently slow convergence or non-convergence of $\boldsymbol{\beta}_{20n}$. These results are in line with Theorem 2.(ii) of the main paper. The contrast between #2–#3 and #4–#5 also agrees with our previous observation that $\hat{\boldsymbol{\gamma}}$ does not converge to $\boldsymbol{\gamma}_0$ when $\delta_0 \geq 0.5$ but it does converge when $\delta_0 < 0.5$ (Section 2). Focusing on $(\psi_0, m) = (0.3, 3)$, the empirical power of the ave-LM test associated with $n \in \{500, 1000, 2000\}$ is $\{0.245, 0.157, 0.145\}$ for Case #2, $\{0.456, 0.481, 0.431\}$ for #3, $\{0.848, 0.952, 0.995\}$ for #4, and $\{0.998, 1.000, 1.000\}$ for #5 (Tables 13-16).

We can draw a few more implications from Tables 12-16. First, the sup-LM, ave-LM, and exp-LM tests are roughly as powerful as each other in Cases #2–#5. Second, when we raise m from 3 to 12, the rejection frequencies of each test are almost unchanged. Robustness to the ratio of sampling frequencies is a desirable feature of the Midastar model, as emphasized in Section 2. Third, when we raise ψ_0 from 0.3 to 0.6 and further to 0.9, the empirical size of each test is invariant and the empirical power improves slightly. The empirical power of the ave-LM test with $(m, n) = (3, 250)$ in Case #3 is $\{0.470, 0.503, 0.567\}$ for $\psi_0 \in \{0.3, 0.6, 0.9\}$, respectively (Table 14). Overall, the simulation results in Tables 12-16 are all reasonable and we can conclude that the bootstrap LM tests achieve desired size and power properties in small and large samples.

4 Consequence of temporal aggregation

In this section, we study the impacts of temporal aggregation on the finite sample performance of the profiling estimation and the wild-bootstrap tests for the no-threshold-effect hypothesis. To this end, we aggregate the high frequency threshold variable x^* into the low frequency level according to (3).[1] We then fit a single-frequency TAR model to $\{y_t, x_t\}_{t \in \mathbb{L}}$:

$$y_t = \begin{cases} \alpha_1 + \phi_1 y_{t-1} + u_t & \text{if } \ x_{t-d} < \mu, \\ \alpha_2 + \phi_2 y_{t-1} + u_t & \text{if } \ x_{t-d} \geq \mu, \qquad t \in \mathbb{L}. \end{cases} \tag{5}$$

The choice space of the delay parameter is $D = \{1, 2, 3\}$. Note that the value of d is in terms of low frequency time periods. We construct the choice space of the threshold parameter

---

[1] We also considered the stock aggregation in place of averaging: $x_t = x_t^*$ for all $t \in \mathbb{L}$. Simulation results associated with averaging and stock aggregation are qualitatively similar to each other, hence we focus on averaging to conserve space. The omitted results of the stock aggregation are available upon request.

according to (4) with $\kappa = 0.7$. Based on model (5), we perform the profiling estimation and the bootstrap tests.

The whole simulation study of this section can be divided into six scenarios depending on the combination of $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $m \in \{3, 12\}$. Recall that ψ_0 determines the persistence of the high frequency threshold variable x^* and that m signifies the ratio of sampling frequencies. When ψ_0 is larger or m is smaller, temporal aggregation is expected to incur the smaller loss of information and hence smaller adverse impacts on statistical inference. Thus, the adverse effects of temporal aggregation are expected to be smallest for $(\psi_0, m) = (0.9, 3)$ and largest for $(\psi_0, m) = (0.3, 12)$. There are four intermediate scenarios between these two extreme scenarios, and we align them as follows.

Scenario 1 $(\psi_0, m) = (0.9, 3)$, under which the adverse impacts of temporal aggregation on statistical inference is expected to be smallest.

Scenario 2 $(\psi_0, m) = (0.9, 12)$.

Scenario 3 $(\psi_0, m) = (0.6, 3)$.

Scenario 4 $(\psi_0, m) = (0.6, 12)$.

Scenario 5 $(\psi_0, m) = (0.3, 3)$.

Scenario 6 $(\psi_0, m) = (0.3, 12)$, under which the adverse impacts of temporal aggregation on statistical inference is expected to be largest.

For each scenario, there are five cases depending on identification categories. In Case #1 (non-identification), the threshold variable is irrelevant under DGP (1), and hence the aggregated TAR model (5) is correctly specified although its two-regime structure is redundant. In Case #5 (strong identification), there are threshold effects which do not diminish asymptotically. A regime change at time $t \in \mathbb{H}$ is triggered by $x^*_{t-1/m}$, and it cannot be captured by averaged x. Hence, under Case #5, model (5) is misspecified relative to DGP (1) even in large samples. Cases #2–#4 are intermediate cases between #1 and #5. We inspect the performance of profiling estimator in Section 4.1, the empirical size of the bootstrap tests for the no-threshold-effect hypothesis in Section 4.2, and their empirical power in Section 4.3.

4.1 Profiling estimation

We report the bias and standard deviation of the profiling estimator in Tables 17-26. We consider $m = 3$ and Cases #1–#5 in Tables 17-21, respectively. Similarly, we consider $m = 12$ and Cases #1–#5 in Tables 22-26. In Scenario 1 (i.e., $\psi_0 = 0.9$ and $m = 3$), the bias of $\hat{\boldsymbol{\beta}}_1 = (\hat{\alpha}_1, \hat{\phi}_1)^\top$ is positive, while the bias of $\hat{\boldsymbol{\beta}}_2 = (\hat{\alpha}_2, \hat{\phi}_2)^\top$ is negative. The positive bias of $\hat{\boldsymbol{\beta}}_1$ and the negative bias of $\hat{\boldsymbol{\beta}}_2$ have similar magnitudes to each other, and both of them are negligibly small for all identification categories. Focusing on $n = 250$, the bias of $\hat{\phi}_1$ is $\{0.004, 0.001, 0.013, 0.024\}$ and the bias of $\hat{\phi}_2$ is $\{-0.045, -0.057, -0.053, -0.065\}$ for Cases #2–#5, respectively (Tables 18-21). The standard deviations of $(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2)$ diminish as n grows. The standard deviation of $\hat{\phi}_1$ under Case #5, for example, is $\{0.097, 0.069, 0.047\}$ for $n \in \{500, 1000, 2000\}$, respectively (Table 21).

We move from Scenario 1 to Scenario 2 by raising m from 3 to 12, while keeping $\psi_0 = 0.9$. Both the positive bias of $\hat{\boldsymbol{\beta}}_1$ and the negative bias of $\hat{\boldsymbol{\beta}}_2$ increase as threshold effects become stronger. The bias of $(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2)$ shrinks as n grows under Cases #1–#4, but not under Case #5. The bias of $\hat{\phi}_1$ associated with $n \in \{500, 1000, 2000\}$ is $\{0.044, 0.024, 0.017\}$ in Case #2, $\{0.070, 0.045, 0.034\}$ in Case #3, $\{0.108, 0.091, 0.074\}$ in Case #4, and $\{0.141, 0.148, 0.143\}$ in Case #5 (Tables 23-26). The standard deviation of $\hat{\phi}_1$ shrinks as n grows for all five cases; for example, it is $\{0.128, 0.090, 0.061\}$ in Case #5 for $n \in \{500, 1000, 2000\}$ (Table 26). We observe similar results for $\hat{\phi}_2$ with the sign of the bias being negative.

Results for Scenarios 3-6 are roughly similar to the results for Scenario 2. We restrict our attention to Scenario 6 (i.e., $\psi_0 = 0.3$ and $m = 12$). The bias in Scenario 6 is slightly larger than the bias in Scenario 2 for each parameter, sample size, and identification category, reflecting the fact that $\psi_0 = 0.3$ in Scenario 6 and $\psi_0 = 0.9$ in Scenario 2. The bias of $\{\hat{\phi}_1, \hat{\phi}_2\}$ in Case #5 with $n = 2000$ is $\{0.176, -0.174\}$, respectively (Table 26). Recall from Table 1 that $\phi_{10} = 0.2$ and $\phi_{20} = 0.55$ in Case #5. In view of Table 26, $\hat{\phi}_1$ converges to $0.2 + 0.176 = 0.376$ and $\hat{\phi}_2$ converges to $0.55 - 0.174 = 0.376$. Hence, the aggregated TAR model fails to capture the spread between $\phi_{10} = 0.2$ and $\phi_{20} = 0.55$, a well-known consequence of temporal aggregation called *spurious non-threshold effects*.

Summarizing Tables 17-26, the bias of the profiling estimator is small only when ψ_0 is large *and* m is small (Scenario 1). The bias is large if ψ_0 is small *or* m is large (Scenarios 2-6). The bias vanishes as the sample size increases under Cases #1–#4, but not under #5. Moreover, $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ are biased in a way that the two regimes are hard to distinguish (i.e., probability limits of $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ are close to each other). It is well known in the literature that temporal aggregation tends to weaken non-linearities in original series, leading to spurious non-threshold effects (e.g., Brännäs and Ohlsson, 1999, Granger and Lee, 1999, Paya and Peel, 2006). Our simulation results are in line with these studies.

4.2 Bootstrap tests: Size

Based on the aggregated TAR model (5), we perform the wild-bootstrap tests for the no-threshold-effect hypothesis $H_0^* : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$. The sup-LM, ave-LM, and exp-LM tests are compared. The Wald tests are omitted since we know from Section 3 that they produce size distortions in small samples. When computing the actual and bootstrap test statistics, the heteroscedasticity-robust covariance matrix estimator is used. Rejection frequencies of each test are computed across $J = 1000$ Monte Carlo samples, where the nominal size is $a = 0.05$ and the number of bootstrap samples is $B = 500$.

We report the rejection frequencies of the sup-LM, ave-LM, and exp-LM tests in Tables 27-29, respectively. We first inspect the empirical size of each test by focusing on Case #1 (non-identification). The empirical size of the LM tests is sufficiently close to the nominal size 5% for all cases except for one case. The ave-LM test rejects the true null hypothesis too often for Scenario 1 (i.e., $\psi_0 = 0.9$ and $m = 3$). Remarkably, the over-rejections do not diminish as n grows. The empirical size of the ave-LM test is $\{0.068, 0.104, 0.090, 0.100, 0.088\}$ for $n \in \{125, 250, 500, 1000, 2000\}$, respectively (Table 28). If we aggregate x^* by stock aggregation (i.e., $x_t = x_t^*$ for each $t \in \mathbb{L}$) instead of averaging (3), the ave-LM test still over-rejects the true null hypothesis in Scenario 1 with excess rejection rate being around 4-5%.

This puzzling phenomenon is consistent with what Motegi and Dennis (2024) [MD2024] documented recently via extensive Monte Carlo simulations. Restricting themselves to a single-frequency TAR framework, MD2024 found that the bootstrap ave-Wald and ave-LM tests suffer from over-rejections even in large samples, when the threshold variable x is persistent and the choice space of the delay parameter is large. In Scenario 1, the aggregated threshold variable x is moderately persistent, since the disaggregated x^* is very persistent (i.e., $\psi_0 = 0.9$) and the ratio of sampling frequencies is small (i.e., $m = 3$). Further, the choice space of delay parameter is not small since $D = \{1, 2, 3\}$ by construction. Given the numerical evidence of MD2024, it is not surprising that the empirical size of the ave-LM test exceeds the nominal size by approximately 5% in Scenario 1.

The size distortion of the ave-LM test diminishes once we raise m from 3 to 12 (Scenario

2) or we lower ψ_0 from 0.9 to 0.6 (Scenario 3). These results are in line with MD2024, since the persistence of x is weaker in Scenarios 2-3 than in Scenario 1. The empirical size of the ave-LM test associated with $n \in \{500, 1000, 2000\}$ is $\{0.055, 0.058, 0.068\}$ for Scenario 2 and $\{0.060, 0.056, 0.062\}$ for Scenario 3 (Table 28).

Why did not we observe any over-rejections when we implemented the Midastar-based bootstrap average tests? Under Case #1 with $(\psi_0, m, n) = (0.9, 3, 2000)$, the empirical size is 0.058 for the ave-Wald test and 0.060 for the ave-LM test (Table 12). Recall from (2) that we used $D = \{1, 2, 3\}$, and these delays are in terms of high frequency time periods. If we replace it with $D = \{3, 6, 9\}$, then the Midastar approach is exactly identical to the single-frequency approach with stock aggregation. In fact, the Midastar-based average tests with $D = \{3, 6, 9\}$ produce excess rejection rates of around 4-5%, as expected.

Our simulation evidence adds a new insight to the over-rejection puzzle posed by MD2024, as they did not consider mixed frequency data. MD2024 argued that it is the cardinality of D, not specific candidate values of d, that played a key role in the over-rejection puzzle. Their argument is likely correct in the single-frequency framework, but specific candidate values of d also matter in the mixed frequency framework. Indeed, we have over-rejections for $D = \{3, 6, 9\}$ but not for $D = \{1, 2, 3\}$. This puzzle is a curious and important problem which challenges the asymptotic validitiy of the bootstrap average tests under H_0^*. We leave a further investigation as a future task.

4.3 Bootstrap tests: Power

Reasons for the over-rejection puzzle of the average tests remain unknown in the literature. MD2024 made a practical suggestion that the sup-LM or exp-LM test should be used instead of the ave-LM test to better control the type-I error rate. Following their suggestion, we focus on the exp-LM test and discuss its empirical power. The exp-LM test becomes much less powerful after aggregating x^* into x. In Scenario 1 (i.e., $\psi_0 = 0.9$ and $m = 3$) with $n = 2000$, the empirical power of the TAR-based exp-LM test is $\{0.089, 0.255, 0.808, 1.000\}$ for Cases #2–#5, respectively (Table 29). These values are uniformly smaller than the Midastar-based rejection frequencies $\{0.137, 0.535, 0.998, 1.000\}$ (Tables 13-16).

When we raise m from 3 to 12 (Scenario 2) or lower ψ_0 from 0.9 to 0.6 (Scenario 3), the exp-LM test with aggregated x loses power drastically. These results confirm that the adverse impacts of temporal aggregation on inference are larger when m is larger or ψ_0 is smaller. Fixing $n = 2000$, the empirical power of the exp-LM test associated with Cases

#3–#5 is $\{0.064, 0.095, 0.262\}$ for Scenario 2 and $\{0.058, 0.117, 0.447\}$ for Scenario 3 (Table 29). When we consider Scenarios 4-6, the exp-LM test has no power for all of Cases #2–#5. Thus, we conclude that temporal aggregation causes spurious non-threshold effects, and consequently the bootstrap tests with the aggregated threshold variable often fail to detect threshold effects in the underlying Midastar process.

5 Comparing out-of-sample predictive accuracy

In this section, we compare the out-of-sample forecast performance of the Midastar model (2) and the aggregated TAR model (5). The DGP is the same as before (Section 1). For each model, we perform the rolling window one-step-ahead prediction with the whole sample size $N \in \{100, 250, 500, 1000, 2000\}$. The window size is fixed at $n = 0.8N$, and the number of windows is $T = N - n = 0.2N$. Note that the prediction horizon $h = 1$ is in terms of low frequency time periods.

We perform the asymptotic S_1-test of Diebold and Mariano (1995) [DM1995] as described in Section 3.3 of the main paper. The S_1-test judges if there is a statistically significant difference between the mean squared forecast errors (MSEs) of the two competing models. The null hypothesis of the DM test, denoted as H_0^{eq}, is that the Midastar-based and TAR-based forecasts are equally accurate. We consider three alternative hypotheses: H_1^{eq} states that the two forecasts have different accuracies; H_1^{midas} states that the Midastar forecast is more accurate than the TAR forecast; H_1^{tar} states that the TAR forecast is more accurate than the Midastar forecast.

In Case #1 (non-identification), there are no threshold effects under DGP (1), hence H_0^{eq} is true. In Case #5 (strong identification), there are threshold effects that do not vanish asymptotically, hence H_1^{midas} is true. Hence, we expect that the probability of rejecting H_0^{eq} in favor of H_1^{eq} and H_1^{midas} should increase as we proceed from Case #1 to #5. We also expect that the probability of rejecting H_0^{eq} in favor of H_1^{tar} should converge to 0 as N grows for all cases considered. Besides, the superiority of Midastar relative to TAR should be least salient in Scenario 1 (i.e., $\psi_0 = 0.9$ and $m = 3$) and most salient in Scenario 6 (i.e., $\psi_0 = 0.3$ and $m = 12$). To verify these conjectures, we compute rejection frequencies of the DM test across $J = 1000$ Monte Carlo trials.

Rejection frequencies of the DM test with $\psi_0 \in \{0.3, 0.6, 0.9\}$ are reported in Tables 30-32, respectively. We first focus on the left half of Table 32 to discuss Scenario 1. In Case #1, the rejection frequencies are around 5% for any sample size $N \in \{125, \dots, 2000\}$ and alternative hypotheses $\{H_1^{eq}, H_1^{midas}, H_1^{tar}\}$. These results are reasonable since H_0^{eq} is true under Case #1. The rejection frequencies are all below 10% for Cases #2–#3, indicating that the out-of-sample performance of Midastar versus TAR is comparable under weak identification.

In Case #4 (semi-strong identification), the DM test has some power for H_1^{eq} and H_1^{midas} if the sample size is sufficiently large. The power improves rather slowly as N grows, reflecting the relatively modest threshold effects. The empirical power associated with $N \in \{500, 1000, 2000\}$ is $\{0.088, 0.113, 0.164\}$ for H_1^{eq} and $\{0.136, 0.197, 0.238\}$ for H_1^{midas} (Table 32). It is reasonable that the one-sided test (H_1^{midas}) is more powerful than the two-sided test (H_1^{eq}) for each sample size. When we switch from Case #4 to #5 (strong identification), the DM test becomes much more powerful. Nevertheless, it requires thousands of observations for the power to reach 1. The empirical power is $\{0.177, 0.368, 0.701\}$ for H_1^{eq} and $\{0.286, 0.501, 0.791\}$ for H_1^{midas}. For all Cases #2–#5, the rejection frequencies in favor of H_1^{tar} are always below 5%, indicating that there is a negligibly small risk for the DM test to incorrectly choose TAR over Midastar.

We shift from Scenario 1 to Scenario 2 by raising m from 3 to 12. As expected, the empirical probability of rejecting H_0^{eq} in favor of H_1^{eq} and H_1^{midas} increases due to the larger adverse effect of temporal aggregation on TAR. Taking Case #5 with $N \in \{500, 1000, 2000\}$ as an example, the empirical power of the DM test is $\{0.230, 0.523, 0.833\}$ for H_1^{eq} and $\{0.366, 0.644, 0.890\}$ for H_1^{midas} (Table 32). Rejection frequencies in Scenarios 3-6 are similar to those in Scenario 2, which is consistent with the in-sample results we saw in Section 4.

Summarizing Tables 30-32, the size and power properties of the DM test are all reasonable, but the power is quite low unless there are constant threshold effects (Case #5). It is rather hard to detect the subtle difference between the out-of-sample forecast performance of Midastar and that of TAR when threshold effects are asymptotically vanishing.

References

Andrews, D. W. K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61, 821–856.

Andrews, D. W. K., and X. Cheng (2012): "Estimation and Inference with Weak, Semi-Strong, and Strong Identification," *Econometrica*, 80(5), 2153–2211.

Brännäs, K., and H. Ohlsson (1999): "Asymmetric time series and temporal aggregation," *The Review of Economics and Statistics*, 81, 341–344.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13, 253–263.

Granger, C. W. J., and T.-H. Lee (1999): "The effect of aggregation on nonlinearity," *Econometric Reviews*, 18, 259–269.

Hansen, B. E. (1996): "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica*, 64(2), 413–430.

Motegi, K., and J. W. Dennis (2024): "An over-rejection puzzle of bootstrap average tests for the no-threshold-effect hypothesis," SSRN Working Paper No. 4739462.

Paya, I., and D. A. Peel (2006): "Temporal aggregation of an ESTAR process: Some implications for purchasing power parity adjustment," *Journal of Applied Econometrics*, 21, 655–668.

Table 1: True values of regression parameters in regime 2 ($\alpha_{10} = 0$, $\phi_{10} = 0.2$)

		Case #1	Case #2	Case #3	Case #4	Case #5
n	(δ_0, λ_0)	$(\infty, \text{n.a.})$	$(0.75, 13.08)$	$(0.5, 3.91)$	$(0.25, 1.17)$	$(0, 0.35)$
125	α_{20n}	0	0.350	0.350	0.350	0.350
125	ϕ_{20n}	0.2	0.550	0.550	0.550	0.550
125	Δ_{0n}	0	0.495	0.495	0.495	0.495
250	α_{20n}	0	0.208	0.247	0.294	0.350
250	ϕ_{20n}	0.2	0.408	0.447	0.494	0.550
250	Δ_{0n}	0	0.294	0.350	0.416	0.495
500	α_{20n}	0	0.124	0.175	0.247	0.350
500	ϕ_{20n}	0.2	0.324	0.375	0.447	0.550
500	Δ_{0n}	0	0.175	0.247	0.350	0.495
1000	α_{20n}	0	0.074	0.124	0.208	0.350
1000	ϕ_{20n}	0.2	0.274	0.324	0.408	0.550
1000	Δ_{0n}	0	0.104	0.175	0.294	0.495
2000	α_{20n}	0	0.044	0.087	0.175	0.350
2000	ϕ_{20n}	0.2	0.244	0.287	0.375	0.550
2000	Δ_{0n}	0	0.062	0.124	0.247	0.495

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-1/m} < 0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-1/m} \geq 0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, and $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$. Regression parameters in regime 1 are $\boldsymbol{\beta}_{10} = (\alpha_{10}, \phi_{10})^\top$, whereas those in regime 2 are $\boldsymbol{\beta}_{20n} = (\alpha_{20n}, \phi_{20n})^\top$. Distance between the two regimes is measured as the Euclidean norm $\Delta_{0n} = \|\boldsymbol{\beta}_{10} - \boldsymbol{\beta}_{20n}\|$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: non-identification ($\delta_0 \to \infty$, $\lambda_0 = \text{n.a.}$). Case #2: weak identification ($\delta_0 = 0.75$, $\lambda_0 = 13.08$). Case #3: boundary case of weak identification ($\delta_0 = 0.5$, $\lambda_0 = 3.91$). Case #4: semi-strong identification ($\delta_0 = 0.25$, $\lambda_0 = 1.17$). Case #5: strong identification ($\delta_0 = 0$, $\lambda_0 = 0.35$).

Table 2: Simulation results on the profiling estimation (Midastar, $m = 3$, Case #1)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	-0.001	0.223	0.006	0.158	0.001	0.110	-0.004	0.077	0.001	0.055
0.3	ϕ_1	-0.027	0.214	-0.000	0.151	-0.002	0.101	0.001	0.074	-0.003	0.055
0.3	α_2	-0.008	0.222	-0.002	0.158	0.002	0.114	0.001	0.077	-0.001	0.056
0.3	ϕ_2	-0.012	0.207	-0.011	0.152	-0.006	0.109	-0.002	0.081	-0.001	0.052
0.3	d	1.033	0.810	0.991	0.820	1.003	0.819	1.031	0.825	1.040	0.809
0.3	μ	-0.001	0.513	-0.007	0.499	0.026	0.501	0.013	0.488	-0.007	0.500
0.6	α_1	0.005	0.236	-0.005	0.162	0.004	0.115	-0.001	0.083	-0.000	0.059
0.6	ϕ_1	-0.007	0.233	-0.008	0.161	-0.002	0.115	-0.000	0.082	-0.002	0.058
0.6	α_2	-0.017	0.223	-0.001	0.164	-0.002	0.116	0.001	0.078	0.002	0.056
0.6	ϕ_2	-0.021	0.222	-0.006	0.155	-0.008	0.109	-0.003	0.080	0.001	0.056
0.6	d	0.993	0.813	0.932	0.821	1.014	0.818	0.977	0.831	1.019	0.829
0.6	μ	-0.045	0.709	-0.029	0.703	-0.019	0.708	-0.031	0.704	-0.027	0.699
0.9	α_1	-0.008	0.256	-0.003	0.168	-0.005	0.124	0.001	0.083	-0.002	0.060
0.9	ϕ_1	-0.049	0.225	-0.024	0.168	-0.013	0.115	-0.009	0.084	-0.006	0.059
0.9	α_2	0.001	0.235	-0.005	0.175	-0.005	0.122	0.001	0.085	0.002	0.061
0.9	ϕ_2	-0.037	0.226	-0.016	0.166	-0.008	0.110	-0.005	0.084	0.001	0.057
0.9	d	0.995	0.855	0.984	0.816	0.981	0.836	0.992	0.827	0.974	0.847
0.9	μ	-0.071	1.580	-0.005	1.544	-0.052	1.510	0.042	1.469	-0.036	1.470

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 3: Simulation results on the profiling estimation (Midastar, $m = 3$, Case #2)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.005	0.189	−0.009	0.137	−0.002	0.098	−0.001	0.072	−0.001	0.052
0.3	ϕ_1	−0.017	0.151	−0.006	0.131	−0.005	0.097	−0.005	0.072	0.000	0.052
0.3	α_2	0.013	0.177	0.011	0.135	0.004	0.100	−0.001	0.071	−0.002	0.055
0.3	ϕ_2	−0.015	0.157	−0.012	0.128	0.000	0.094	−0.005	0.072	−0.003	0.051
0.3	d	0.196	0.536	0.350	0.678	0.526	0.774	0.637	0.807	0.805	0.848
0.3	μ	−0.018	0.311	−0.026	0.378	0.002	0.414	−0.037	0.440	−0.007	0.470
0.6	α_1	−0.016	0.173	−0.009	0.132	−0.008	0.107	−0.005	0.077	0.003	0.057
0.6	ϕ_1	−0.023	0.152	−0.014	0.132	−0.014	0.094	−0.006	0.073	−0.004	0.050
0.6	α_2	0.026	0.188	0.025	0.141	0.006	0.102	0.004	0.076	0.001	0.057
0.6	ϕ_2	−0.009	0.152	−0.004	0.123	0.000	0.099	0.001	0.071	0.000	0.054
0.6	d	0.277	0.605	0.401	0.700	0.606	0.789	0.670	0.826	0.750	0.802
0.6	μ	−0.001	0.423	0.013	0.521	0.006	0.586	−0.004	0.613	0.006	0.647
0.9	α_1	−0.029	0.193	−0.021	0.132	−0.019	0.100	−0.012	0.074	−0.004	0.053
0.9	ϕ_1	−0.049	0.181	−0.044	0.132	−0.025	0.101	−0.021	0.074	−0.006	0.054
0.9	α_2	0.037	0.203	0.021	0.140	0.016	0.104	0.009	0.078	0.006	0.058
0.9	ϕ_2	−0.023	0.150	0.003	0.118	0.005	0.096	0.010	0.074	0.004	0.056
0.9	d	0.461	0.724	0.604	0.786	0.765	0.847	0.819	0.837	0.863	0.837
0.9	μ	−0.033	0.966	−0.062	1.077	0.002	1.220	−0.013	1.341	0.062	1.385

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 4: Simulation results on the profiling estimation (Midastar, $m = 3$, Case #3)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	−0.006	0.178	−0.012	0.125	−0.002	0.088	−0.003	0.066	−0.003	0.047
0.3	ϕ_1	−0.020	0.160	−0.012	0.115	−0.008	0.081	−0.006	0.062	−0.005	0.041
0.3	α_2	0.023	0.174	0.007	0.126	0.008	0.092	0.003	0.063	0.001	0.046
0.3	ϕ_2	−0.020	0.164	−0.007	0.119	0.002	0.087	−0.001	0.060	0.003	0.043
0.3	d	0.197	0.531	0.207	0.546	0.243	0.590	0.251	0.583	0.239	0.583
0.3	μ	−0.026	0.328	−0.008	0.320	0.023	0.336	−0.004	0.330	0.003	0.335
0.6	α_1	−0.021	0.186	−0.017	0.128	−0.010	0.087	−0.006	0.059	−0.005	0.046
0.6	ϕ_1	−0.027	0.163	−0.021	0.124	−0.017	0.080	−0.009	0.059	−0.007	0.043
0.6	α_2	0.012	0.182	0.015	0.126	0.009	0.092	0.009	0.062	0.005	0.047
0.6	ϕ_2	−0.019	0.159	−0.004	0.116	0.006	0.083	0.004	0.059	0.004	0.044
0.6	d	0.262	0.599	0.261	0.599	0.267	0.590	0.283	0.622	0.293	0.617
0.6	μ	−0.027	0.466	−0.040	0.453	0.027	0.447	0.011	0.432	0.001	0.479
0.9	α_1	−0.032	0.184	−0.027	0.128	−0.016	0.086	−0.011	0.067	−0.010	0.045
0.9	ϕ_1	−0.042	0.165	−0.036	0.123	−0.021	0.084	−0.015	0.062	−0.012	0.042
0.9	α_2	0.041	0.207	0.023	0.133	0.019	0.092	0.016	0.066	0.008	0.046
0.9	ϕ_2	−0.038	0.164	−0.006	0.107	0.005	0.082	0.006	0.062	0.009	0.043
0.9	d	0.465	0.735	0.499	0.735	0.496	0.750	0.528	0.752	0.547	0.765
0.9	μ	−0.015	0.903	−0.084	0.953	0.020	0.933	−0.010	1.021	−0.007	0.995

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified (boundary case). Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 5: Simulation results on the profiling estimation (Midastar, $m = 3$, Case #4)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	−0.008	0.175	−0.007	0.111	−0.009	0.074	−0.005	0.050	−0.002	0.033
0.3	ϕ_1	−0.008	0.168	−0.019	0.098	−0.009	0.066	−0.007	0.047	−0.003	0.031
0.3	α_2	0.015	0.175	0.014	0.110	0.009	0.075	0.006	0.048	0.003	0.033
0.3	ϕ_2	−0.015	0.160	0.001	0.104	0.001	0.071	0.005	0.042	0.001	0.031
0.3	d	0.222	0.572	0.087	0.363	0.021	0.186	0.010	0.126	0.002	0.063
0.3	μ	−0.004	0.332	−0.007	0.254	−0.001	0.204	−0.005	0.138	0.000	0.102
0.6	α_1	−0.021	0.175	−0.017	0.113	−0.005	0.075	−0.003	0.050	−0.003	0.034
0.6	ϕ_1	−0.030	0.160	−0.020	0.104	−0.011	0.068	−0.008	0.046	−0.003	0.031
0.6	α_2	0.019	0.175	0.015	0.116	0.011	0.077	0.004	0.050	0.004	0.036
0.6	ϕ_2	−0.015	0.154	−0.002	0.100	0.001	0.067	0.004	0.045	−0.000	0.032
0.6	d	0.224	0.553	0.128	0.419	0.068	0.309	0.029	0.219	0.002	0.045
0.6	μ	−0.040	0.438	−0.017	0.350	−0.005	0.279	0.001	0.202	−0.001	0.141
0.9	α_1	−0.035	0.184	−0.021	0.120	−0.008	0.074	−0.008	0.051	−0.004	0.033
0.9	ϕ_1	−0.042	0.159	−0.024	0.106	−0.018	0.069	−0.008	0.050	−0.005	0.032
0.9	α_2	0.046	0.198	0.020	0.125	0.014	0.081	0.007	0.052	0.004	0.035
0.9	ϕ_2	−0.024	0.152	−0.005	0.103	0.003	0.065	−0.001	0.047	0.000	0.032
0.9	d	0.437	0.718	0.302	0.612	0.239	0.564	0.126	0.422	0.074	0.320
0.9	μ	−0.044	0.892	−0.017	0.748	0.008	0.589	−0.001	0.453	−0.003	0.298

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is semi-strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 6: Simulation results on the profiling estimation (Midastar, $m = 3$, Case #5)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	−0.008	0.180	−0.007	0.102	−0.006	0.068	−0.003	0.047	−0.001	0.032
0.3	ϕ_1	−0.024	0.158	−0.010	0.091	−0.007	0.057	−0.003	0.042	−0.002	0.027
0.3	α_2	0.011	0.179	0.009	0.106	0.006	0.068	0.002	0.048	0.002	0.033
0.3	ϕ_2	−0.017	0.159	−0.004	0.093	−0.001	0.060	−0.002	0.041	−0.000	0.028
0.3	d	0.173	0.509	0.025	0.180	0.000	0.000	0.000	0.000	0.000	0.000
0.3	μ	−0.024	0.320	−0.007	0.191	−0.003	0.092	0.002	0.045	0.000	0.024
0.6	α_1	−0.018	0.179	−0.008	0.102	−0.004	0.069	−0.001	0.049	0.000	0.033
0.6	ϕ_1	−0.035	0.154	−0.013	0.093	−0.009	0.061	−0.003	0.042	−0.001	0.028
0.6	α_2	0.020	0.171	0.015	0.108	0.006	0.071	0.003	0.046	0.002	0.033
0.6	ϕ_2	−0.011	0.160	−0.005	0.095	−0.004	0.059	0.000	0.040	0.002	0.028
0.6	d	0.217	0.529	0.036	0.225	0.003	0.071	0.000	0.000	0.000	0.000
0.6	μ	−0.019	0.415	0.000	0.244	−0.006	0.129	−0.000	0.055	−0.000	0.026
0.9	α_1	−0.033	0.185	−0.011	0.106	−0.001	0.069	−0.003	0.046	0.000	0.033
0.9	ϕ_1	−0.053	0.159	−0.023	0.096	−0.008	0.063	−0.003	0.043	−0.001	0.029
0.9	α_2	0.042	0.203	0.022	0.113	0.009	0.076	0.002	0.050	0.003	0.035
0.9	ϕ_2	−0.030	0.152	−0.012	0.088	−0.006	0.056	−0.001	0.039	−0.002	0.028
0.9	d	0.486	0.732	0.197	0.520	0.026	0.193	0.001	0.032	0.000	0.000
0.9	μ	−0.006	0.948	−0.016	0.538	−0.005	0.242	−0.009	0.099	−0.001	0.044

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 7: Simulation results on the profiling estimation (Midastar, $m = 12$, Case #1)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	-0.003	0.198	-0.001	0.137	-0.001	0.097	-0.001	0.068	-0.000	0.051
0.3	ϕ_1	-0.016	0.189	-0.010	0.129	0.003	0.093	-0.003	0.070	-0.002	0.047
0.3	α_2	0.002	0.193	-0.004	0.135	-0.005	0.097	0.001	0.069	0.001	0.046
0.3	ϕ_2	-0.005	0.181	-0.001	0.131	-0.005	0.093	-0.001	0.066	0.002	0.048
0.3	d	1.021	0.821	1.011	0.812	0.958	0.819	0.976	0.816	1.039	0.818
0.3	μ	-0.029	0.281	-0.014	0.276	-0.001	0.279	-0.019	0.281	-0.015	0.275
0.6	α_1	-0.006	0.212	-0.000	0.141	0.007	0.103	0.003	0.072	-0.003	0.050
0.6	ϕ_1	-0.021	0.203	-0.007	0.144	-0.010	0.097	-0.002	0.072	-0.002	0.049
0.6	α_2	-0.007	0.205	-0.004	0.145	-0.000	0.100	-0.001	0.073	0.001	0.051
0.6	ϕ_2	-0.017	0.194	-0.002	0.138	-0.004	0.099	-0.005	0.069	-0.002	0.049
0.6	d	0.928	0.810	0.943	0.831	1.004	0.811	0.974	0.818	0.964	0.809
0.6	μ	-0.027	0.469	-0.004	0.458	-0.021	0.455	0.004	0.453	-0.010	0.441
0.9	α_1	0.000	0.229	-0.003	0.159	0.002	0.112	0.000	0.076	-0.002	0.056
0.9	ϕ_1	-0.006	0.233	-0.008	0.151	-0.011	0.110	-0.000	0.077	-0.002	0.055
0.9	α_2	0.003	0.231	0.003	0.157	0.000	0.112	0.001	0.080	0.003	0.057
0.9	ϕ_2	-0.019	0.223	-0.012	0.153	-0.005	0.109	0.001	0.080	-0.001	0.054
0.9	d	1.004	0.839	1.013	0.834	0.958	0.839	1.004	0.827	0.989	0.842
0.9	μ	-0.068	1.350	-0.019	1.320	0.002	1.285	0.031	1.316	0.003	1.278

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 8: Simulation results on the profiling estimation (Midastar, $m = 12$, Case #2)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	-0.003	0.171	-0.008	0.118	0.001	0.086	-0.002	0.064	-0.001	0.047
0.3	ϕ_1	-0.018	0.140	-0.008	0.112	-0.007	0.082	-0.003	0.064	-0.002	0.046
0.3	α_2	0.011	0.169	0.011	0.123	0.001	0.090	-0.002	0.066	0.001	0.047
0.3	ϕ_2	-0.017	0.144	-0.009	0.113	0.004	0.084	-0.003	0.063	-0.002	0.046
0.3	d	0.167	0.497	0.306	0.639	0.465	0.757	0.614	0.818	0.722	0.821
0.3	μ	-0.018	0.205	-0.008	0.215	-0.008	0.242	-0.003	0.244	-0.018	0.256
0.6	α_1	-0.010	0.163	-0.012	0.122	-0.004	0.091	-0.005	0.067	-0.004	0.049
0.6	ϕ_1	-0.023	0.142	-0.012	0.117	-0.008	0.087	-0.009	0.060	-0.004	0.047
0.6	α_2	0.016	0.160	0.013	0.117	0.008	0.094	0.004	0.066	0.001	0.047
0.6	ϕ_2	-0.005	0.139	-0.006	0.116	0.004	0.085	0.002	0.065	0.002	0.047
0.6	d	0.222	0.548	0.390	0.701	0.524	0.765	0.654	0.803	0.715	0.826
0.6	μ	-0.018	0.311	-0.011	0.362	0.015	0.385	0.011	0.403	-0.011	0.416
0.9	α_1	-0.019	0.171	-0.020	0.130	-0.020	0.099	-0.014	0.070	-0.007	0.053
0.9	ϕ_1	-0.043	0.141	-0.029	0.121	-0.024	0.091	-0.011	0.068	-0.010	0.052
0.9	α_2	0.033	0.169	0.011	0.127	0.017	0.099	0.014	0.070	0.006	0.052
0.9	ϕ_2	-0.004	0.147	0.014	0.116	0.009	0.089	0.008	0.069	0.005	0.050
0.9	d	0.439	0.717	0.572	0.770	0.691	0.805	0.808	0.827	0.852	0.828
0.9	μ	-0.012	0.773	-0.039	0.943	-0.029	1.058	0.029	1.144	-0.035	1.183

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 9: Simulation results on the profiling estimation (Midastar, $m = 12$, Case #3)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	-0.001	0.164	-0.001	0.121	-0.007	0.077	-0.005	0.054	-0.004	0.040
0.3	ϕ_1	-0.010	0.140	-0.012	0.108	-0.008	0.075	-0.002	0.057	-0.005	0.039
0.3	α_2	0.022	0.155	0.003	0.120	0.007	0.078	0.007	0.056	0.002	0.041
0.3	ϕ_2	-0.013	0.143	-0.007	0.113	0.004	0.080	0.001	0.055	0.001	0.040
0.3	d	0.149	0.470	0.206	0.555	0.187	0.512	0.209	0.553	0.220	0.560
0.3	μ	-0.009	0.201	-0.004	0.199	0.006	0.202	0.005	0.206	-0.010	0.204
0.6	α_1	-0.009	0.164	-0.007	0.120	-0.010	0.081	-0.006	0.057	-0.004	0.042
0.6	ϕ_1	-0.030	0.142	-0.021	0.104	-0.010	0.077	-0.006	0.057	-0.005	0.039
0.6	α_2	0.023	0.160	0.013	0.121	0.012	0.078	0.009	0.058	0.007	0.042
0.6	ϕ_2	-0.013	0.141	-0.000	0.101	0.006	0.077	0.002	0.057	0.005	0.040
0.6	d	0.233	0.565	0.229	0.537	0.232	0.559	0.270	0.594	0.301	0.626
0.6	μ	-0.036	0.302	-0.017	0.323	0.015	0.316	-0.003	0.324	0.014	0.336
0.9	α_1	-0.027	0.178	-0.019	0.119	-0.017	0.087	-0.014	0.063	-0.010	0.042
0.9	ϕ_1	-0.041	0.155	-0.024	0.110	-0.024	0.083	-0.013	0.059	-0.010	0.041
0.9	α_2	0.031	0.177	0.032	0.125	0.017	0.088	0.010	0.060	0.010	0.043
0.9	ϕ_2	-0.014	0.151	-0.000	0.106	0.005	0.084	0.008	0.058	0.008	0.041
0.9	d	0.437	0.706	0.442	0.708	0.551	0.768	0.481	0.718	0.525	0.747
0.9	μ	-0.047	0.836	0.002	0.810	-0.008	0.892	-0.038	0.876	0.013	0.901

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified (boundary case). Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 10: Simulation results on the profiling estimation (Midastar, $m = 12$, Case #4)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	−0.014	0.165	−0.010	0.106	−0.008	0.071	−0.005	0.045	−0.004	0.034
0.3	ϕ_1	−0.021	0.141	−0.012	0.096	−0.012	0.064	−0.007	0.047	−0.004	0.032
0.3	α_2	0.009	0.166	0.011	0.107	0.007	0.071	0.004	0.048	0.003	0.033
0.3	ϕ_2	−0.021	0.144	−0.004	0.094	0.005	0.068	0.001	0.046	0.001	0.031
0.3	d	0.169	0.502	0.076	0.352	0.016	0.147	0.009	0.130	0.000	0.000
0.3	μ	−0.017	0.208	−0.007	0.170	−0.008	0.150	−0.003	0.125	−0.003	0.097
0.6	α_1	−0.000	0.164	−0.005	0.108	−0.006	0.074	−0.006	0.050	−0.003	0.033
0.6	ϕ_1	−0.023	0.143	−0.017	0.099	−0.009	0.068	−0.007	0.046	−0.003	0.031
0.6	α_2	0.017	0.166	0.012	0.113	0.012	0.071	0.006	0.051	0.005	0.033
0.6	ϕ_2	−0.000	0.147	0.003	0.098	0.001	0.067	0.003	0.046	0.002	0.033
0.6	d	0.206	0.533	0.119	0.421	0.056	0.281	0.017	0.144	0.000	0.000
0.6	μ	−0.010	0.324	−0.004	0.266	0.005	0.227	0.001	0.180	0.003	0.132
0.9	α_1	−0.023	0.178	−0.017	0.115	−0.011	0.075	−0.004	0.049	−0.003	0.035
0.9	ϕ_1	−0.037	0.150	−0.024	0.105	−0.015	0.069	−0.007	0.046	−0.004	0.033
0.9	α_2	0.033	0.173	0.023	0.110	0.013	0.078	0.007	0.049	0.004	0.033
0.9	ϕ_2	−0.005	0.146	−0.001	0.098	−0.002	0.067	0.004	0.047	0.003	0.032
0.9	d	0.444	0.712	0.329	0.631	0.227	0.549	0.137	0.434	0.056	0.277
0.9	μ	−0.051	0.823	−0.042	0.658	−0.020	0.583	−0.004	0.423	−0.002	0.291

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is semi-strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 11: Simulation results on the profiling estimation (Midastar, $m = 12$, Case #5)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\theta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	−0.012	0.163	−0.008	0.103	−0.003	0.068	0.000	0.047	−0.001	0.033
0.3	ϕ_1	−0.010	0.146	−0.013	0.090	−0.008	0.058	−0.006	0.041	−0.002	0.028
0.3	α_2	0.019	0.158	0.014	0.100	0.007	0.071	0.003	0.045	0.001	0.033
0.3	ϕ_2	−0.015	0.143	−0.004	0.086	0.001	0.059	−0.002	0.040	−0.000	0.028
0.3	d	0.161	0.487	0.021	0.191	0.000	0.000	0.000	0.000	0.000	0.000
0.3	μ	−0.021	0.199	−0.009	0.139	−0.003	0.088	−0.002	0.045	−0.000	0.021
0.6	α_1	−0.014	0.169	−0.006	0.101	−0.007	0.068	−0.000	0.046	−0.001	0.034
0.6	ϕ_1	−0.023	0.144	−0.012	0.091	−0.006	0.059	−0.004	0.040	−0.001	0.028
0.6	α_2	0.007	0.166	0.015	0.100	0.007	0.068	0.004	0.047	0.000	0.033
0.6	ϕ_2	−0.006	0.149	−0.006	0.087	−0.001	0.058	−0.003	0.040	−0.000	0.028
0.6	d	0.220	0.547	0.034	0.221	0.002	0.045	0.000	0.000	0.000	0.000
0.6	μ	−0.015	0.318	0.003	0.208	−0.005	0.112	−0.001	0.051	−0.002	0.024
0.9	α_1	−0.021	0.179	−0.010	0.106	−0.002	0.066	−0.002	0.047	−0.001	0.032
0.9	ϕ_1	−0.027	0.159	−0.018	0.089	−0.008	0.060	−0.002	0.042	−0.003	0.029
0.9	α_2	0.033	0.185	0.013	0.109	0.005	0.071	0.005	0.048	0.001	0.032
0.9	ϕ_2	−0.010	0.154	−0.005	0.092	−0.003	0.059	−0.001	0.040	−0.001	0.028
0.9	d	0.462	0.721	0.176	0.495	0.027	0.180	0.000	0.000	0.000	0.000
0.9	μ	−0.028	0.843	−0.018	0.499	0.008	0.250	0.000	0.097	−0.003	0.048

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 12: Rejection frequencies of the bootstrap tests for the no-threshold-effect hypothesis H_0^* (Midastar, Case #1: non-identification)

		$m = 3$					$m = 12$				
ψ_0	Statistic \ n	125	250	500	1000	2000	125	250	500	1000	2000
0.3	sup-Wald	0.158	0.094	0.069	0.042	0.053	0.109	0.096	0.075	0.059	0.071
0.3	ave-Wald	0.091	0.060	0.053	0.054	0.051	0.070	0.059	0.059	0.061	0.059
0.3	exp-Wald	0.132	0.073	0.058	0.047	0.051	0.093	0.080	0.067	0.056	0.060
0.3	sup-LM	0.040	0.032	0.045	0.036	0.052	0.038	0.048	0.059	0.054	0.063
0.3	ave-LM	0.045	0.042	0.047	0.052	0.049	0.035	0.042	0.048	0.059	0.053
0.3	exp-LM	0.036	0.036	0.042	0.039	0.048	0.035	0.046	0.056	0.048	0.052
0.6	sup-Wald	0.186	0.104	0.078	0.066	0.060	0.117	0.081	0.075	0.061	0.056
0.6	ave-Wald	0.087	0.060	0.054	0.046	0.054	0.056	0.053	0.054	0.045	0.053
0.6	exp-Wald	0.147	0.071	0.065	0.051	0.053	0.089	0.070	0.072	0.052	0.058
0.6	sup-LM	0.032	0.031	0.051	0.043	0.047	0.040	0.045	0.052	0.049	0.051
0.6	ave-LM	0.041	0.038	0.042	0.045	0.049	0.030	0.036	0.047	0.040	0.048
0.6	exp-LM	0.033	0.039	0.049	0.042	0.048	0.034	0.040	0.049	0.050	0.055
0.9	sup-Wald	0.232	0.129	0.080	0.053	0.053	0.161	0.100	0.077	0.078	0.068
0.9	ave-Wald	0.099	0.063	0.053	0.046	0.058	0.091	0.057	0.064	0.063	0.062
0.9	exp-Wald	0.168	0.098	0.064	0.048	0.059	0.129	0.074	0.066	0.068	0.063
0.9	sup-LM	0.038	0.047	0.042	0.039	0.051	0.041	0.039	0.051	0.064	0.059
0.9	ave-LM	0.051	0.049	0.043	0.041	0.060	0.050	0.046	0.051	0.061	0.057
0.9	exp-LM	0.035	0.046	0.039	0.035	0.054	0.042	0.042	0.052	0.053	0.055

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. We report the rejection frequencies of the wild-bootstrap tests for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 13: Rejection frequencies of the bootstrap tests for the no-threshold-effect hypothesis H_0^* (Midastar, Case #2: weak identification)

		$m = 3$					$m = 12$				
ψ_0	Statistic \ n	125	250	500	1000	2000	125	250	500	1000	2000
0.3	sup-Wald	0.724	0.469	0.296	0.173	0.130	0.718	0.458	0.286	0.201	0.153
0.3	ave-Wald	0.627	0.423	0.282	0.166	0.149	0.618	0.396	0.269	0.203	0.138
0.3	exp-Wald	0.721	0.472	0.303	0.179	0.144	0.705	0.448	0.282	0.202	0.147
0.3	sup-LM	0.515	0.342	0.235	0.150	0.120	0.545	0.366	0.253	0.187	0.140
0.3	ave-LM	0.447	0.340	0.245	0.157	0.145	0.485	0.336	0.249	0.192	0.132
0.3	exp-LM	0.539	0.368	0.252	0.161	0.138	0.571	0.381	0.262	0.192	0.141
0.6	sup-Wald	0.745	0.470	0.265	0.176	0.128	0.751	0.468	0.285	0.207	0.145
0.6	ave-Wald	0.656	0.399	0.258	0.174	0.129	0.651	0.398	0.272	0.199	0.139
0.6	exp-Wald	0.730	0.468	0.270	0.190	0.127	0.745	0.454	0.291	0.217	0.150
0.6	sup-LM	0.503	0.323	0.207	0.149	0.112	0.572	0.357	0.238	0.187	0.135
0.6	ave-LM	0.500	0.332	0.227	0.164	0.127	0.541	0.347	0.252	0.192	0.134
0.6	exp-LM	0.540	0.341	0.225	0.168	0.119	0.613	0.367	0.255	0.200	0.147
0.9	sup-Wald	0.797	0.515	0.350	0.235	0.146	0.760	0.498	0.330	0.212	0.168
0.9	ave-Wald	0.721	0.472	0.341	0.233	0.159	0.692	0.454	0.338	0.242	0.171
0.9	exp-Wald	0.772	0.515	0.342	0.236	0.144	0.745	0.492	0.341	0.231	0.169
0.9	sup-LM	0.545	0.353	0.278	0.202	0.132	0.525	0.371	0.265	0.194	0.158
0.9	ave-LM	0.594	0.408	0.312	0.221	0.155	0.592	0.411	0.310	0.234	0.166
0.9	exp-LM	0.572	0.393	0.302	0.210	0.137	0.565	0.404	0.287	0.214	0.163

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. We report the rejection frequencies of the wild-bootstrap tests for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 14: Rejection frequencies of the bootstrap tests for the no-threshold-effect hypothesis H_0^* (Midastar, Case #3: boundary case of weak identification)

		$m = 3$					$m = 12$				
ψ_0	Statistic \ n	125	250	500	1000	2000	125	250	500	1000	2000
0.3	sup-Wald	0.723	0.623	0.545	0.522	0.446	0.716	0.633	0.575	0.542	0.526
0.3	ave-Wald	0.629	0.548	0.489	0.498	0.444	0.635	0.581	0.537	0.500	0.503
0.3	exp-Wald	0.724	0.617	0.559	0.545	0.477	0.713	0.641	0.587	0.554	0.550
0.3	sup-LM	0.513	0.497	0.472	0.491	0.433	0.544	0.550	0.534	0.515	0.514
0.3	ave-LM	0.456	0.470	0.456	0.481	0.431	0.492	0.514	0.502	0.482	0.494
0.3	exp-LM	0.537	0.533	0.502	0.524	0.457	0.570	0.567	0.552	0.539	0.541
0.6	sup-Wald	0.750	0.637	0.535	0.510	0.521	0.730	0.645	0.562	0.540	0.519
0.6	ave-Wald	0.646	0.577	0.490	0.500	0.500	0.656	0.571	0.508	0.511	0.494
0.6	exp-Wald	0.738	0.651	0.549	0.520	0.534	0.740	0.642	0.568	0.565	0.535
0.6	sup-LM	0.488	0.500	0.465	0.466	0.494	0.551	0.539	0.499	0.512	0.507
0.6	ave-LM	0.486	0.503	0.453	0.483	0.489	0.538	0.505	0.477	0.499	0.492
0.6	exp-LM	0.525	0.544	0.495	0.492	0.520	0.598	0.559	0.524	0.552	0.528
0.9	sup-Wald	0.747	0.652	0.613	0.578	0.514	0.754	0.704	0.611	0.575	0.571
0.9	ave-Wald	0.668	0.634	0.599	0.589	0.555	0.687	0.669	0.610	0.589	0.588
0.9	exp-Wald	0.743	0.655	0.617	0.593	0.550	0.745	0.704	0.630	0.598	0.596
0.9	sup-LM	0.492	0.520	0.526	0.531	0.492	0.515	0.583	0.549	0.537	0.558
0.9	ave-LM	0.531	0.567	0.575	0.574	0.554	0.580	0.619	0.588	0.575	0.579
0.9	exp-LM	0.519	0.543	0.567	0.572	0.535	0.554	0.612	0.587	0.582	0.577

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified (boundary case). Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. We report the rejection frequencies of the wild-bootstrap tests for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 15: Rejection frequencies of the bootstrap tests for the no-threshold-effect hypothesis H_0^* (Midastar, Case #4: semi-strong identification)

		$m = 3$					$m = 12$				
ψ_0	Statistic \ n	125	250	500	1000	2000	125	250	500	1000	2000
0.3	sup-Wald	0.724	0.776	0.896	0.979	0.997	0.695	0.821	0.933	0.987	0.997
0.3	ave-Wald	0.644	0.712	0.868	0.958	0.996	0.610	0.781	0.906	0.981	0.997
0.3	exp-Wald	0.724	0.785	0.908	0.982	0.998	0.684	0.828	0.936	0.992	0.997
0.3	sup-LM	0.489	0.683	0.875	0.975	0.997	0.542	0.769	0.919	0.986	0.997
0.3	ave-LM	0.493	0.636	0.848	0.952	0.995	0.486	0.724	0.885	0.978	0.996
0.3	exp-LM	0.537	0.706	0.890	0.979	0.998	0.565	0.786	0.928	0.991	0.997
0.6	sup-Wald	0.747	0.833	0.910	0.972	1.000	0.720	0.844	0.916	0.980	0.997
0.6	ave-Wald	0.643	0.768	0.868	0.943	0.997	0.625	0.776	0.883	0.971	0.992
0.6	exp-Wald	0.731	0.831	0.914	0.977	1.000	0.717	0.848	0.926	0.979	0.998
0.6	sup-LM	0.491	0.728	0.878	0.968	1.000	0.536	0.777	0.895	0.975	0.997
0.6	ave-LM	0.498	0.701	0.844	0.939	0.997	0.521	0.724	0.867	0.966	0.991
0.6	exp-LM	0.534	0.765	0.892	0.972	1.000	0.567	0.791	0.908	0.979	0.998
0.9	sup-Wald	0.785	0.837	0.931	0.981	0.995	0.765	0.859	0.935	0.982	0.999
0.9	ave-Wald	0.706	0.816	0.916	0.979	0.998	0.704	0.837	0.922	0.975	0.998
0.9	exp-Wald	0.768	0.842	0.930	0.983	0.998	0.772	0.859	0.940	0.983	0.999
0.9	sup-LM	0.518	0.735	0.909	0.977	0.996	0.541	0.768	0.909	0.978	0.999
0.9	ave-LM	0.587	0.767	0.906	0.975	0.998	0.586	0.796	0.911	0.974	0.998
0.9	exp-LM	0.564	0.776	0.908	0.981	0.998	0.582	0.789	0.921	0.982	0.999

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is semi-strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. We report the rejection frequencies of the wild-bootstrap tests for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 16: Rejection frequencies of the bootstrap tests for the no-threshold-effect hypothesis H_0^* (Midastar, Case #5: strong identification)

		$m = 3$					$m = 12$				
ψ_0	Statistic \ n	125	250	500	1000	2000	125	250	500	1000	2000
0.3	sup-Wald	0.718	0.949	1.000	1.000	1.000	0.722	0.956	1.000	1.000	1.000
0.3	ave-Wald	0.617	0.900	0.999	1.000	1.000	0.628	0.922	0.999	1.000	1.000
0.3	exp-Wald	0.711	0.949	1.000	1.000	1.000	0.718	0.960	1.000	1.000	1.000
0.3	sup-LM	0.457	0.912	0.999	1.000	1.000	0.568	0.924	0.999	1.000	1.000
0.3	ave-LM	0.442	0.858	0.998	1.000	1.000	0.494	0.891	0.999	1.000	1.000
0.3	exp-LM	0.501	0.924	1.000	1.000	1.000	0.590	0.937	0.999	1.000	1.000
0.6	sup-Wald	0.739	0.957	1.000	1.000	1.000	0.738	0.944	1.000	1.000	1.000
0.6	ave-Wald	0.622	0.906	0.997	1.000	1.000	0.646	0.902	1.000	1.000	1.000
0.6	exp-Wald	0.730	0.958	1.000	1.000	1.000	0.734	0.947	1.000	1.000	1.000
0.6	sup-LM	0.501	0.903	1.000	1.000	1.000	0.535	0.914	1.000	1.000	1.000
0.6	ave-LM	0.485	0.877	0.997	1.000	1.000	0.526	0.873	0.999	1.000	1.000
0.6	exp-LM	0.530	0.916	1.000	1.000	1.000	0.559	0.926	1.000	1.000	1.000
0.9	sup-Wald	0.778	0.946	1.000	1.000	1.000	0.808	0.976	1.000	1.000	1.000
0.9	ave-Wald	0.711	0.927	1.000	1.000	1.000	0.750	0.961	1.000	1.000	1.000
0.9	exp-Wald	0.761	0.944	1.000	1.000	1.000	0.801	0.973	1.000	1.000	1.000
0.9	sup-LM	0.513	0.905	0.999	1.000	1.000	0.591	0.949	1.000	1.000	1.000
0.9	ave-LM	0.578	0.898	1.000	1.000	1.000	0.642	0.949	0.996	1.000	1.000
0.9	exp-LM	0.558	0.916	1.000	1.000	1.000	0.625	0.956	1.000	1.000	1.000

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #5: $\delta_0 = 0$, hence $\gamma_0 = (d_0, \mu_0)^\top$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x^*_{t-d/m} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x^*_{t-d/m} \geq \mu$. We report the rejection frequencies of the wild-bootstrap tests for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 17: Simulation results on the profiling estimation (aggregated TAR, $m = 3$, Case #1)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	−0.002	0.263	0.001	0.183	0.008	0.125	−0.002	0.090	0.000	0.067
0.3	ϕ_1	−0.019	0.265	−0.007	0.180	−0.004	0.124	−0.005	0.087	−0.002	0.062
0.3	α_2	0.004	0.255	−0.006	0.181	−0.007	0.126	0.003	0.096	−0.001	0.064
0.3	ϕ_2	−0.010	0.240	−0.009	0.170	−0.005	0.121	−0.002	0.086	0.003	0.060
0.6	α_1	0.002	0.276	−0.006	0.182	0.002	0.129	0.000	0.087	−0.004	0.063
0.6	ϕ_1	−0.030	0.259	−0.020	0.173	−0.004	0.128	−0.003	0.083	−0.001	0.060
0.6	α_2	0.006	0.243	−0.000	0.177	0.004	0.131	−0.004	0.092	0.001	0.063
0.6	ϕ_2	−0.007	0.240	−0.014	0.170	−0.004	0.119	−0.006	0.087	0.001	0.063
0.9	α_1	0.007	0.268	−0.012	0.180	0.004	0.124	0.005	0.090	−0.001	0.064
0.9	ϕ_1	−0.051	0.255	−0.023	0.164	−0.007	0.123	−0.004	0.087	−0.005	0.062
0.9	α_2	−0.006	0.245	0.001	0.182	−0.005	0.126	−0.002	0.087	−0.000	0.058
0.9	ϕ_2	−0.033	0.233	−0.022	0.169	−0.020	0.119	−0.003	0.084	−0.001	0.060

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 18: Simulation results on the profiling estimation (aggregated TAR, $m = 3$, Case #2)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.199	0.280	0.097	0.189	0.054	0.129	0.038	0.092	0.023	0.063
0.3	ϕ_1	0.149	0.263	0.099	0.182	0.056	0.122	0.028	0.085	0.015	0.063
0.3	α_2	−0.182	0.266	−0.094	0.181	−0.058	0.131	−0.035	0.087	−0.022	0.066
0.3	ϕ_2	−0.188	0.234	−0.122	0.165	−0.065	0.118	−0.033	0.085	−0.021	0.063
0.6	α_1	0.147	0.276	0.087	0.189	0.052	0.128	0.031	0.091	0.018	0.067
0.6	ϕ_1	0.110	0.258	0.079	0.177	0.043	0.121	0.029	0.088	0.020	0.064
0.6	α_2	−0.141	0.274	−0.085	0.183	−0.049	0.124	−0.033	0.090	−0.018	0.064
0.6	ϕ_2	−0.172	0.217	−0.098	0.165	−0.055	0.122	−0.035	0.087	−0.022	0.063
0.9	α_1	0.047	0.245	0.018	0.177	0.006	0.124	0.002	0.088	0.002	0.061
0.9	ϕ_1	0.016	0.234	0.004	0.154	0.006	0.115	0.003	0.081	0.002	0.059
0.9	α_2	−0.011	0.280	−0.005	0.174	−0.012	0.125	−0.006	0.084	−0.002	0.063
0.9	ϕ_2	−0.095	0.200	−0.045	0.152	−0.024	0.117	−0.010	0.082	−0.007	0.060

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 19: Simulation results on the profiling estimation (aggregated TAR, $m = 3$, Case #3)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.185	0.278	0.133	0.191	0.088	0.124	0.060	0.092	0.040	0.067
0.3	ϕ_1	0.149	0.263	0.109	0.172	0.078	0.126	0.058	0.090	0.040	0.066
0.3	α_2	−0.168	0.262	−0.127	0.179	−0.082	0.127	−0.063	0.088	−0.043	0.062
0.3	ϕ_2	−0.200	0.236	−0.127	0.169	−0.090	0.122	−0.063	0.084	−0.044	0.061
0.6	α_1	0.158	0.286	0.100	0.190	0.078	0.134	0.052	0.089	0.035	0.063
0.6	ϕ_1	0.121	0.260	0.083	0.176	0.063	0.124	0.052	0.087	0.035	0.061
0.6	α_2	−0.138	0.275	−0.093	0.187	−0.070	0.129	−0.052	0.091	−0.032	0.067
0.6	ϕ_2	−0.179	0.225	−0.102	0.163	−0.070	0.118	−0.055	0.086	−0.039	0.060
0.9	α_1	0.048	0.252	0.025	0.175	0.017	0.115	0.010	0.084	0.006	0.060
0.9	ϕ_1	0.012	0.226	0.001	0.162	0.004	0.110	0.002	0.080	0.004	0.055
0.9	α_2	−0.004	0.286	−0.016	0.174	−0.013	0.117	−0.010	0.080	−0.007	0.057
0.9	ϕ_2	−0.108	0.209	−0.057	0.149	−0.030	0.106	−0.015	0.075	−0.010	0.055

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified (boundary case). Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 20: Simulation results on the profiling estimation (aggregated TAR, $m = 3$, Case #4)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.185	0.280	0.150	0.187	0.119	0.130	0.100	0.094	0.084	0.064
0.3	ϕ_1	0.146	0.272	0.133	0.172	0.116	0.128	0.097	0.090	0.084	0.059
0.3	α_2	−0.176	0.265	−0.150	0.181	−0.117	0.130	−0.098	0.090	−0.083	0.067
0.3	ϕ_2	−0.194	0.239	−0.153	0.164	−0.122	0.115	−0.101	0.088	−0.086	0.060
0.6	α_1	0.141	0.282	0.119	0.192	0.104	0.127	0.079	0.095	0.071	0.065
0.6	ϕ_1	0.137	0.261	0.116	0.180	0.093	0.129	0.079	0.092	0.071	0.062
0.6	α_2	−0.130	0.270	−0.114	0.188	−0.106	0.133	−0.077	0.092	−0.071	0.063
0.6	ϕ_2	−0.185	0.224	−0.144	0.160	−0.109	0.120	−0.088	0.082	−0.074	0.061
0.9	α_1	0.045	0.254	0.048	0.165	0.029	0.109	0.022	0.075	0.021	0.048
0.9	ϕ_1	0.008	0.239	0.013	0.154	0.025	0.107	0.020	0.071	0.019	0.050
0.9	α_2	−0.013	0.278	−0.010	0.182	−0.019	0.118	−0.023	0.075	−0.023	0.048
0.9	ϕ_2	−0.107	0.200	−0.053	0.134	−0.038	0.096	−0.031	0.064	−0.025	0.046

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is semi-strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 21: Simulation results on the profiling estimation (aggregated TAR, $m = 3$, Case #5)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.177	0.269	0.176	0.188	0.175	0.128	0.171	0.093	0.169	0.062
0.3	ϕ_1	0.149	0.264	0.158	0.185	0.173	0.127	0.168	0.088	0.169	0.063
0.3	α_2	−0.166	0.265	−0.176	0.187	−0.175	0.137	−0.168	0.088	−0.169	0.069
0.3	ϕ_2	−0.193	0.247	−0.180	0.167	−0.170	0.120	−0.175	0.085	−0.168	0.059
0.6	α_1	0.155	0.279	0.131	0.184	0.144	0.125	0.142	0.087	0.134	0.055
0.6	ϕ_1	0.130	0.258	0.135	0.176	0.135	0.123	0.132	0.084	0.135	0.061
0.6	α_2	−0.130	0.264	−0.136	0.185	−0.129	0.143	−0.134	0.094	−0.127	0.063
0.6	ϕ_2	−0.196	0.237	−0.157	0.168	−0.145	0.119	−0.133	0.081	−0.128	0.054
0.9	α_1	0.043	0.248	0.041	0.155	0.050	0.092	0.064	0.065	0.073	0.046
0.9	ϕ_1	0.004	0.223	0.024	0.150	0.045	0.097	0.059	0.069	0.074	0.047
0.9	α_2	−0.004	0.260	−0.030	0.165	−0.034	0.108	−0.047	0.069	−0.058	0.048
0.9	ϕ_2	−0.103	0.202	−0.065	0.128	−0.054	0.081	−0.059	0.052	−0.059	0.038

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 3$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 22: Simulation results on the profiling estimation (aggregated TAR, $m = 12$, Case #1)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.005	0.277	0.018	0.182	−0.009	0.132	0.001	0.089	−0.000	0.061
0.3	ϕ_1	−0.019	0.266	−0.011	0.183	−0.001	0.125	−0.000	0.088	0.001	0.063
0.3	α_2	−0.008	0.246	−0.013	0.180	0.002	0.128	−0.000	0.094	−0.003	0.063
0.3	ϕ_2	−0.030	0.250	−0.013	0.170	−0.008	0.124	−0.003	0.084	−0.002	0.064
0.6	α_1	0.005	0.279	0.003	0.185	0.000	0.123	0.002	0.091	−0.001	0.064
0.6	ϕ_1	−0.001	0.264	−0.002	0.183	−0.005	0.124	−0.006	0.087	−0.001	0.061
0.6	α_2	−0.002	0.237	−0.003	0.184	−0.004	0.127	−0.004	0.093	−0.003	0.067
0.6	ϕ_2	−0.010	0.234	−0.020	0.172	−0.003	0.123	−0.002	0.088	−0.001	0.061
0.9	α_1	−0.005	0.274	−0.004	0.191	0.001	0.129	−0.000	0.095	0.001	0.061
0.9	ϕ_1	−0.041	0.275	−0.017	0.168	−0.009	0.122	0.005	0.089	0.001	0.060
0.9	α_2	0.012	0.252	−0.001	0.187	−0.001	0.130	−0.006	0.085	−0.006	0.065
0.9	ϕ_2	−0.025	0.229	−0.010	0.169	−0.009	0.125	−0.007	0.084	−0.003	0.064

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 23: Simulation results on the profiling estimation (aggregated TAR, $m = 12$, Case #2)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.198	0.290	0.106	0.185	0.059	0.129	0.037	0.089	0.022	0.064
0.3	ϕ_1	0.164	0.260	0.100	0.176	0.068	0.122	0.034	0.089	0.022	0.062
0.3	α_2	−0.166	0.268	−0.104	0.179	−0.058	0.127	−0.034	0.088	−0.022	0.064
0.3	ϕ_2	−0.197	0.239	−0.112	0.176	−0.076	0.116	−0.037	0.087	−0.026	0.063
0.6	α_1	0.194	0.289	0.110	0.185	0.057	0.128	0.040	0.088	0.019	0.062
0.6	ϕ_1	0.152	0.277	0.097	0.176	0.052	0.125	0.034	0.092	0.021	0.065
0.6	α_2	−0.153	0.268	−0.107	0.179	−0.061	0.132	−0.041	0.088	−0.021	0.066
0.6	ϕ_2	−0.203	0.240	−0.109	0.164	−0.065	0.119	−0.039	0.085	−0.022	0.062
0.9	α_1	0.170	0.280	0.095	0.190	0.050	0.128	0.035	0.089	0.019	0.065
0.9	ϕ_1	0.138	0.262	0.070	0.177	0.044	0.124	0.024	0.088	0.017	0.061
0.9	α_2	−0.144	0.272	−0.080	0.183	−0.051	0.127	−0.028	0.092	−0.016	0.061
0.9	ϕ_2	−0.171	0.232	−0.104	0.173	−0.060	0.122	−0.037	0.084	−0.022	0.058

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 24: Simulation results on the profiling estimation (aggregated TAR, $m = 12$, Case #3)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.182	0.281	0.115	0.189	0.082	0.130	0.061	0.093	0.044	0.064
0.3	ϕ_1	0.155	0.258	0.115	0.189	0.082	0.124	0.059	0.089	0.043	0.061
0.3	α_2	-0.183	0.258	-0.127	0.183	-0.083	0.129	-0.058	0.088	-0.044	0.062
0.3	ϕ_2	-0.199	0.236	-0.130	0.167	-0.093	0.117	-0.067	0.085	-0.045	0.063
0.6	α_1	0.160	0.281	0.134	0.188	0.089	0.132	0.064	0.090	0.046	0.063
0.6	ϕ_1	0.168	0.259	0.125	0.175	0.085	0.119	0.057	0.087	0.045	0.061
0.6	α_2	-0.175	0.254	-0.133	0.188	-0.088	0.129	-0.062	0.094	-0.046	0.066
0.6	ϕ_2	-0.204	0.231	-0.144	0.163	-0.084	0.119	-0.064	0.085	-0.044	0.064
0.9	α_1	0.164	0.287	0.122	0.182	0.069	0.133	0.054	0.088	0.037	0.065
0.9	ϕ_1	0.145	0.259	0.090	0.177	0.070	0.126	0.045	0.085	0.034	0.060
0.9	α_2	-0.152	0.265	-0.117	0.180	-0.065	0.122	-0.051	0.093	-0.038	0.064
0.9	ϕ_2	-0.177	0.221	-0.115	0.169	-0.084	0.116	-0.060	0.086	-0.037	0.061

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is weakly identified (boundary case). Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 25: Simulation results on the profiling estimation (aggregated TAR, $m = 12$, Case #4)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.195	0.268	0.158	0.179	0.125	0.131	0.107	0.089	0.087	0.065
0.3	ϕ_1	0.158	0.261	0.144	0.181	0.121	0.125	0.100	0.089	0.089	0.063
0.3	α_2	−0.190	0.265	−0.150	0.185	−0.124	0.126	−0.107	0.092	−0.085	0.067
0.3	ϕ_2	−0.187	0.230	−0.156	0.158	−0.131	0.123	−0.105	0.088	−0.093	0.058
0.6	α_1	0.187	0.270	0.150	0.189	0.126	0.130	0.103	0.091	0.089	0.063
0.6	ϕ_1	0.156	0.265	0.137	0.172	0.121	0.120	0.100	0.086	0.086	0.060
0.6	α_2	−0.179	0.264	−0.146	0.183	−0.121	0.128	−0.098	0.092	−0.087	0.065
0.6	ϕ_2	−0.190	0.235	−0.151	0.172	−0.121	0.121	−0.101	0.090	−0.088	0.063
0.9	α_1	0.148	0.275	0.132	0.186	0.108	0.132	0.080	0.087	0.074	0.066
0.9	ϕ_1	0.136	0.264	0.105	0.176	0.108	0.122	0.091	0.084	0.074	0.064
0.9	α_2	−0.124	0.269	−0.123	0.181	−0.094	0.132	−0.078	0.093	−0.071	0.064
0.9	ϕ_2	−0.191	0.234	−0.138	0.168	−0.116	0.118	−0.091	0.085	−0.076	0.058

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is semi-strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 26: Simulation results on the profiling estimation (aggregated TAR, $m = 12$, Case #5)

		$n = 125$		$n = 250$		$n = 500$		$n = 1000$		$n = 2000$	
ψ_0	$\boldsymbol{\beta}$	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev	Bias	Stdev
0.3	α_1	0.167	0.277	0.182	0.185	0.177	0.132	0.173	0.096	0.177	0.064
0.3	ϕ_1	0.155	0.266	0.174	0.179	0.175	0.127	0.175	0.088	0.176	0.064
0.3	α_2	−0.175	0.263	−0.171	0.180	−0.175	0.133	−0.172	0.094	−0.174	0.064
0.3	ϕ_2	−0.201	0.230	−0.182	0.172	−0.178	0.116	−0.179	0.088	−0.174	0.063
0.6	α_1	0.176	0.274	0.175	0.186	0.175	0.133	0.173	0.091	0.174	0.066
0.6	ϕ_1	0.160	0.258	0.160	0.181	0.178	0.122	0.176	0.087	0.171	0.061
0.6	α_2	−0.167	0.270	−0.173	0.190	−0.170	0.135	−0.172	0.096	−0.175	0.067
0.6	ϕ_2	−0.197	0.238	−0.190	0.170	−0.184	0.126	−0.177	0.083	−0.174	0.059
0.9	α_1	0.145	0.283	0.161	0.181	0.157	0.129	0.147	0.091	0.145	0.064
0.9	ϕ_1	0.128	0.254	0.143	0.178	0.141	0.128	0.148	0.090	0.143	0.061
0.9	α_2	−0.141	0.272	−0.153	0.188	−0.140	0.140	−0.139	0.097	−0.141	0.066
0.9	ϕ_2	−0.185	0.225	−0.156	0.164	−0.150	0.119	−0.144	0.083	−0.135	0.057

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m = 12$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). The profiling estimator for $(\alpha_1, \phi_1, \alpha_2, \phi_2, d, \mu)$ is computed. For each parameter, bias and standard deviation across $J = 1000$ Monte Carlo samples are reported.

Table 27: Rejection frequencies of the bootstrap sup-LM test for the no-threshold-effect hypothesis H_0^* based on the aggregated TAR model

		$m = 3$					$m = 12$				
ψ_0	n	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
0.3	125	0.024	0.022	0.021	0.025	0.030	0.023	0.029	0.030	0.023	0.033
0.3	250	0.041	0.033	0.038	0.029	0.042	0.046	0.044	0.035	0.038	0.039
0.3	500	0.039	0.045	0.052	0.049	0.046	0.049	0.041	0.041	0.047	0.041
0.3	1000	0.047	0.047	0.052	0.049	0.054	0.047	0.055	0.044	0.048	0.040
0.3	2000	0.052	0.045	0.049	0.063	0.083	0.041	0.063	0.055	0.051	0.062
0.6	125	0.022	0.026	0.049	0.034	0.033	0.029	0.029	0.026	0.029	0.029
0.6	250	0.029	0.043	0.050	0.046	0.048	0.039	0.033	0.045	0.033	0.047
0.6	500	0.034	0.049	0.043	0.062	0.104	0.038	0.040	0.043	0.048	0.049
0.6	1000	0.053	0.052	0.047	0.085	0.181	0.042	0.065	0.043	0.049	0.042
0.6	2000	0.046	0.062	0.055	0.106	0.419	0.040	0.051	0.047	0.059	0.047
0.9	125	0.016	0.146	0.126	0.130	0.141	0.027	0.027	0.029	0.030	0.020
0.9	250	0.035	0.113	0.173	0.250	0.391	0.044	0.046	0.040	0.049	0.043
0.9	500	0.030	0.115	0.194	0.401	0.780	0.046	0.039	0.056	0.060	0.060
0.9	1000	0.044	0.088	0.221	0.585	0.990	0.050	0.075	0.050	0.060	0.127
0.9	2000	0.042	0.079	0.207	0.759	1.000	0.048	0.048	0.059	0.095	0.243

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0$ is weakly identified. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0$ is weakly identified (boundary case). Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0$ is semi-strongly identified. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^m x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). We report the rejection frequencies of the bootstrap sup-LM test for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 28: Rejection frequencies of the bootstrap ave-LM test for the no-threshold-effect hypothesis H_0^* based on the aggregated TAR model

		$m = 3$					$m = 12$				
ψ_0	n	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
0.3	125	0.043	0.039	0.034	0.040	0.033	0.030	0.037	0.041	0.042	0.029
0.3	250	0.046	0.043	0.048	0.043	0.053	0.048	0.044	0.047	0.038	0.041
0.3	500	0.046	0.047	0.045	0.045	0.045	0.053	0.045	0.041	0.055	0.048
0.3	1000	0.034	0.061	0.045	0.033	0.064	0.043	0.047	0.050	0.041	0.046
0.3	2000	0.041	0.045	0.048	0.041	0.087	0.040	0.055	0.052	0.055	0.056
0.6	125	0.045	0.057	0.064	0.069	0.069	0.030	0.043	0.032	0.036	0.047
0.6	250	0.047	0.062	0.074	0.077	0.081	0.052	0.037	0.048	0.035	0.043
0.6	500	0.060	0.060	0.071	0.089	0.117	0.050	0.048	0.039	0.043	0.041
0.6	1000	0.056	0.065	0.050	0.097	0.220	0.032	0.057	0.050	0.051	0.049
0.6	2000	0.062	0.064	0.070	0.143	0.463	0.047	0.050	0.053	0.055	0.043
0.9	125	0.068	0.315	0.302	0.313	0.315	0.051	0.056	0.059	0.049	0.048
0.9	250	0.104	0.240	0.360	0.451	0.605	0.077	0.067	0.068	0.076	0.087
0.9	500	0.090	0.229	0.364	0.575	0.882	0.055	0.062	0.076	0.090	0.101
0.9	1000	0.100	0.196	0.385	0.739	0.993	0.058	0.074	0.063	0.091	0.156
0.9	2000	0.088	0.151	0.375	0.868	1.000	0.068	0.074	0.084	0.110	0.296

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0$ is weakly identified. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0$ is weakly identified (boundary case). Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0$ is semi-strongly identified. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). We report the rejection frequencies of the bootstrap ave-LM test for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 29: Rejection frequencies of the bootstrap exp-LM test for the no-threshold-effect hypothesis H_0^* based on the aggregated TAR model

		$m = 3$					$m = 12$				
ψ_0	n	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
0.3	125	0.026	0.023	0.022	0.030	0.026	0.020	0.026	0.024	0.020	0.033
0.3	250	0.037	0.031	0.042	0.036	0.043	0.047	0.042	0.032	0.037	0.036
0.3	500	0.039	0.049	0.047	0.050	0.043	0.048	0.042	0.045	0.047	0.037
0.3	1000	0.033	0.043	0.051	0.051	0.055	0.040	0.049	0.045	0.045	0.035
0.3	2000	0.047	0.041	0.042	0.060	0.083	0.040	0.063	0.049	0.047	0.055
0.6	125	0.023	0.033	0.046	0.038	0.037	0.030	0.034	0.026	0.032	0.026
0.6	250	0.031	0.044	0.054	0.052	0.058	0.050	0.031	0.043	0.029	0.049
0.6	500	0.039	0.049	0.050	0.076	0.111	0.038	0.042	0.040	0.044	0.042
0.6	1000	0.053	0.052	0.047	0.086	0.202	0.041	0.058	0.043	0.047	0.044
0.6	2000	0.050	0.061	0.058	0.117	0.447	0.038	0.048	0.043	0.058	0.044
0.9	125	0.018	0.177	0.183	0.172	0.181	0.032	0.031	0.033	0.030	0.023
0.9	250	0.044	0.140	0.215	0.301	0.464	0.044	0.052	0.037	0.054	0.053
0.9	500	0.037	0.131	0.244	0.458	0.831	0.047	0.040	0.061	0.064	0.062
0.9	1000	0.052	0.111	0.264	0.641	0.993	0.053	0.064	0.050	0.068	0.137
0.9	2000	0.048	0.089	0.255	0.808	1.000	0.048	0.051	0.064	0.095	0.262

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 \in \{0.3, 0.6, 0.9\}$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $n \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0$ is weakly identified. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0$ is weakly identified (boundary case). Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0$ is semi-strongly identified. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0$ is strongly identified. Model: $y_t = \alpha_1 + \phi_1 y_{t-1} + u_t$ if $x_{t-d} < \mu$ and $y_t = \alpha_2 + \phi_2 y_{t-1} + u_t$ if $x_{t-d} \geq \mu$, where $x_t = m^{-1} \sum_{k=1}^{m} x^*_{t-1+k/m}$ for all $t \in \mathbb{L}$ (i.e., averaging). We report the rejection frequencies of the bootstrap exp-LM test for the no-threshold-effect hypothesis $H_0^* : (\alpha_1, \phi_1) = (\alpha_2, \phi_2)$, where the nominal size is $a = 0.05$; the number of bootstrap samples is $B = 500$; the number of Monte Carlo samples is $J = 1000$.

Table 30: Rejection frequencies of the Diebold-Mariano test for the equal predictive accuracy hypothesis H_0^{eq} ($\psi_0 = 0.3$)

		$m = 3$					$m = 12$				
H_1	N	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
H_1^{eq}	125	0.044	0.090	0.090	0.067	0.071	0.056	0.091	0.086	0.083	0.077
H_1^{eq}	250	0.042	0.054	0.080	0.105	0.139	0.048	0.057	0.086	0.098	0.142
H_1^{eq}	500	0.043	0.050	0.072	0.117	0.276	0.039	0.052	0.068	0.129	0.267
H_1^{eq}	1000	0.037	0.049	0.048	0.158	0.505	0.041	0.033	0.071	0.156	0.520
H_1^{eq}	2000	0.033	0.041	0.056	0.207	0.828	0.048	0.037	0.066	0.216	0.851
H_1^{midas}	125	0.054	0.137	0.131	0.102	0.119	0.057	0.141	0.140	0.123	0.136
H_1^{midas}	250	0.041	0.085	0.113	0.171	0.211	0.053	0.088	0.135	0.146	0.234
H_1^{midas}	500	0.048	0.068	0.101	0.210	0.395	0.051	0.070	0.125	0.204	0.404
H_1^{midas}	1000	0.050	0.067	0.089	0.249	0.647	0.051	0.064	0.107	0.242	0.644
H_1^{midas}	2000	0.043	0.051	0.089	0.312	0.907	0.052	0.059	0.103	0.314	0.922
H_1^{tar}	125	0.050	0.024	0.025	0.027	0.019	0.054	0.017	0.016	0.014	0.029
H_1^{tar}	250	0.051	0.025	0.022	0.017	0.008	0.039	0.019	0.015	0.020	0.012
H_1^{tar}	500	0.044	0.032	0.025	0.005	0.000	0.031	0.029	0.016	0.009	0.000
H_1^{tar}	1000	0.033	0.029	0.018	0.005	0.000	0.027	0.023	0.017	0.005	0.000
H_1^{tar}	2000	0.034	0.025	0.017	0.003	0.000	0.042	0.024	0.022	0.000	0.000

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 = 0.3$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $N \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0$ is weakly identified. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0$ is weakly identified (boundary case). Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0$ is semi-strongly identified. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0$ is strongly identified. The rolling window out-of-sample prediction is performed, where the window size is fixed at $n = 0.8N$. The one-step ahead forecasts based on the Midastar and aggregated TAR models are compared by the asymptotic Diebold-Mariano test. H_0^{eq}: The two forecasts are equally accurate. H_1^{eq}: The two forecasts have different accuracies. H_1^{midas}: The Midastar forecast is more accurate than the TAR forecast. H_1^{tar}: The TAR forecast is more accurate than the Midastar forecast. Rejection frequencies across $J = 1000$ Monte Carlo samples are reported, where the nominal size is $a = 0.05$.

Table 31: Rejection frequencies of the Diebold-Mariano test for the equal predictive accuracy hypothesis H_0^{eq} ($\psi_0 = 0.6$)

		$m = 3$					$m = 12$				
H_1	N	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
H_1^{eq}	125	0.044	0.076	0.078	0.058	0.071	0.056	0.083	0.084	0.091	0.084
H_1^{eq}	250	0.052	0.056	0.065	0.093	0.137	0.049	0.065	0.061	0.091	0.136
H_1^{eq}	500	0.034	0.051	0.071	0.115	0.275	0.059	0.051	0.053	0.112	0.263
H_1^{eq}	1000	0.043	0.040	0.066	0.130	0.506	0.035	0.049	0.054	0.159	0.508
H_1^{eq}	2000	0.032	0.032	0.057	0.213	0.834	0.041	0.056	0.060	0.234	0.848
H_1^{midas}	125	0.049	0.118	0.125	0.110	0.122	0.051	0.127	0.140	0.138	0.145
H_1^{midas}	250	0.054	0.083	0.113	0.157	0.212	0.050	0.091	0.112	0.153	0.235
H_1^{midas}	500	0.039	0.084	0.096	0.187	0.416	0.058	0.080	0.096	0.192	0.382
H_1^{midas}	1000	0.049	0.056	0.099	0.211	0.638	0.049	0.055	0.100	0.263	0.629
H_1^{midas}	2000	0.037	0.042	0.101	0.309	0.903	0.067	0.071	0.099	0.352	0.907
H_1^{tar}	125	0.041	0.027	0.026	0.018	0.014	0.067	0.023	0.025	0.025	0.014
H_1^{tar}	250	0.050	0.033	0.024	0.011	0.008	0.042	0.032	0.017	0.016	0.005
H_1^{tar}	500	0.034	0.031	0.034	0.005	0.000	0.040	0.034	0.017	0.009	0.001
H_1^{tar}	1000	0.045	0.033	0.013	0.007	0.001	0.032	0.038	0.016	0.002	0.000
H_1^{tar}	2000	0.027	0.033	0.014	0.002	0.000	0.030	0.036	0.018	0.000	0.000

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 = 0.6$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $N \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0$ is weakly identified. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0$ is weakly identified (boundary case). Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0$ is semi-strongly identified. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0$ is strongly identified. The rolling window out-of-sample prediction is performed, where the window size is fixed at $n = 0.8N$. The one-step ahead forecasts based on the Midastar and aggregated TAR models are compared by the asymptotic Diebold-Mariano test. H_0^{eq}: The two forecasts are equally accurate. H_1^{eq}: The two forecasts have different accuracies. H_1^{midas}: The Midastar forecast is more accurate than the TAR forecast. H_1^{tar}: The TAR forecast is more accurate than the Midastar forecast. Rejection frequencies across $J = 1000$ Monte Carlo samples are reported, where the nominal size is $a = 0.05$.

Table 32: Rejection frequencies of the Diebold-Mariano test for the equal predictive accuracy hypothesis H_0^{eq} ($\psi_0 = 0.9$)

		$m = 3$					$m = 12$				
H_1	N	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
H_1^{eq}	125	0.046	0.063	0.057	0.050	0.067	0.061	0.083	0.086	0.063	0.083
H_1^{eq}	250	0.054	0.053	0.054	0.065	0.093	0.047	0.054	0.076	0.090	0.109
H_1^{eq}	500	0.048	0.051	0.065	0.088	0.177	0.041	0.045	0.052	0.102	0.230
H_1^{eq}	1000	0.040	0.044	0.048	0.113	0.368	0.037	0.042	0.058	0.150	0.523
H_1^{eq}	2000	0.036	0.051	0.058	0.164	0.701	0.043	0.037	0.049	0.227	0.833
H_1^{midas}	125	0.059	0.097	0.090	0.090	0.096	0.053	0.128	0.139	0.113	0.121
H_1^{midas}	250	0.063	0.079	0.080	0.122	0.158	0.054	0.076	0.111	0.137	0.172
H_1^{midas}	500	0.046	0.060	0.098	0.136	0.286	0.048	0.082	0.091	0.179	0.366
H_1^{midas}	1000	0.048	0.067	0.080	0.197	0.501	0.044	0.059	0.094	0.236	0.644
H_1^{midas}	2000	0.049	0.052	0.092	0.238	0.791	0.045	0.051	0.092	0.320	0.890
H_1^{tar}	125	0.052	0.022	0.023	0.024	0.022	0.044	0.014	0.023	0.025	0.022
H_1^{tar}	250	0.053	0.034	0.031	0.015	0.011	0.031	0.033	0.016	0.020	0.009
H_1^{tar}	500	0.044	0.040	0.021	0.021	0.002	0.038	0.024	0.019	0.009	0.002
H_1^{tar}	1000	0.047	0.029	0.030	0.011	0.001	0.028	0.022	0.018	0.004	0.000
H_1^{tar}	2000	0.025	0.046	0.022	0.003	0.000	0.038	0.029	0.018	0.002	0.000

DGP for the target variable: $y_t = \alpha_{10} + \phi_{10} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} < \mu_0$ and $y_t = \alpha_{20n} + \phi_{20n} y_{t-1} + \epsilon_t$ if $x^*_{t-d_0/m} \geq \mu_0$, where $\alpha_{10} = 0$, $\phi_{10} = 0.2$, $\alpha_{20n} = \alpha_{10} + \lambda_0 n^{-\delta_0}$, $\phi_{20n} = \phi_{10} + \lambda_0 n^{-\delta_0}$, $d_0 = 1$, $\mu_0 = 0$, $\epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and $m \in \{3, 12\}$. DGP for the threshold variable: $x^*_t = \psi_0 x^*_{t-1/m} + \nu^*_t$, where $\psi_0 = 0.9$ and $\nu^*_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. Sample size: $N \in \{125, 250, 500, 1000, 2000\}$. Case #1: $\delta_0 \to \infty$, hence $\boldsymbol{\gamma}_0 = (d_0, \mu_0)^\top$ is unidentified. Case #2: $\delta_0 = 0.75$, hence $\boldsymbol{\gamma}_0$ is weakly identified. Case #3: $\delta_0 = 0.5$, hence $\boldsymbol{\gamma}_0$ is weakly identified (boundary case). Case #4: $\delta_0 = 0.25$, hence $\boldsymbol{\gamma}_0$ is semi-strongly identified. Case #5: $\delta_0 = 0$, hence $\boldsymbol{\gamma}_0$ is strongly identified. The rolling window out-of-sample prediction is performed, where the window size is fixed at $n = 0.8N$. The one-step ahead forecasts based on the Midastar and aggregated TAR models are compared by the asymptotic Diebold-Mariano test. H_0^{eq}: The two forecasts are equally accurate. H_1^{eq}: The two forecasts have different accuracies. H_1^{midas}: The Midastar forecast is more accurate than the TAR forecast. H_1^{tar}: The TAR forecast is more accurate than the Midastar forecast. Rejection frequencies across $J = 1000$ Monte Carlo samples are reported, where the nominal size is $a = 0.05$.