Supplemental Appendix to “Forecasting with Dynamic Panel Data Models”
Laura Liu, Hyungsik Roger Moon, and Frank Schorfheide

A Theoretical Derivations and Proofs

A.1 Proofs for Section 3.2

A.1.1 Preliminaries

Throughout the proofs, we use the notation $\epsilon$ for a small positive constant such that

$$0 < \epsilon < \epsilon_0.$$ 

In addition, we will make use of the following two lemmas.

**Lemma A.1** If $A_N(\pi) = o_u(\pi N^+)$ and $B_N(\pi) = o_u(\pi N^+)$, then $C_N(\pi) = A_N(\pi) + B_N(\pi) = o_u(\pi N^+)$.

**Proof of Lemma A.1.** Take an arbitrary $\epsilon > 0$. We need to show that there exists a sequence $\eta^c_N(\epsilon)$ such that

$$N^{-\epsilon}C_N(\pi) \leq \eta^c_N(\epsilon).$$

Write

$$N^{-\epsilon}C_N(\pi) = N^{-\epsilon}(A_N(\pi) + B_N(\pi)).$$

Because $A_N$ and $B_N$ are subpolynomial, there exist sequences $\eta^a_N(\epsilon)$ and $\eta^b_N(\epsilon)$ such that

$$N^{-\epsilon}(A_N(\pi) + B_N(\pi)) \leq \eta^a_N(\epsilon) + \eta^b_N(\epsilon).$$

Thus, we can choose $\eta^c_N(\epsilon) = \eta^a_N(\epsilon) + \eta^b_N(\epsilon) \rightarrow 0$ to establish the claim. ■

**Lemma A.2** If $A_N(\pi) = o_u(\pi N^+)$ and $B_N(\pi) = o_u(\pi N^+)$, then $C_N(\pi) = A_N(\pi)B_N(\pi) = o_u(\pi N^+)$. 
Proof of Lemma A.2. Take an arbitrary $\epsilon > 0$. We need to show that there exists a sequence $\eta_N^c(\epsilon)$ such that

$$N^{-\epsilon}C_N(\pi) \leq \eta_N^c(\epsilon).$$

Write

$$N^{-\epsilon}C_N(\pi) = \left(N^{-\epsilon/2}A_N(\pi)\right)\left(N^{-\epsilon/2}B_N(\pi)\right).$$

Because $A_N$ and $B_N$ are subpolynomial, there exist sequences $\eta_N^a(\epsilon/2)$ and $\eta_N^b(\epsilon/2)$ such that

$$\left(N^{-\epsilon/2}A_N(\pi)\right)\left(N^{-\epsilon/2}B_N(\pi)\right) \leq \eta_N^a(\epsilon/2)\eta_N^b(\epsilon/2).$$

Thus, we can choose $\eta_N^c(\epsilon) = \eta_N^a(\epsilon/2)\eta_N^b(\epsilon/2) \longrightarrow 0$ to establish the claim. ■

A.1.2 Main Theorem

Proof of Theorem 3.7. The goal is to prove that for any given $\epsilon_0 > 0$

$$\limsup_{N \to \infty} \sup_{\pi \in \Pi} \frac{R_N(\hat{Y}_{T+1}, \pi) - R_{N}^{opt}(\pi)}{N \mathbb{E}^{Y,\lambda_i}_{\theta,\pi}[\lambda_i]} \leq N^{\epsilon_0}, \quad (A.1)$$

where

$$R_N(\hat{Y}_{T+1}, \pi) = N \mathbb{E}^{Y,\lambda_i}_{\theta,\pi} \left[\lambda_i + \rho Y_i - \hat{Y}_{i+1}\right]^2 + N\sigma^2,$$

$$R_{N}^{opt}(\pi) = N \mathbb{E}^{Y,\lambda_i}_{\theta,\pi} \left[\lambda_i - \mathbb{E}^{\lambda_i}_{\theta,\pi,Y_i}[\lambda_i]\right]^2 + N\sigma^2.$$

Here we used the fact that there is cross-sectional independence and symmetry in terms of $i$. The statement is equivalent to

$$\limsup_{N \to \infty} \sup_{\pi \in \Pi} \frac{N \mathbb{E}^{Y,\lambda_i}_{\theta,\pi} \left[\lambda_i + \rho Y_i - \hat{Y}_{i+1}\right]^2}{N \mathbb{E}^{Y,\lambda_i}_{\theta,\pi}[\lambda_i]} \leq 1. \quad (A.2)$$

Here we made the dependence on $\pi$ of the risks and the posterior moments explicit. In the calculations below, we often drop the $\pi$ argument to simplify the notation.

In the main text we asserted that

$$p_*(\hat{\lambda}_i, y_i) = \mathbb{E}^{\gamma^{(-i)}}_{\theta,\pi,Y_i}[\hat{p}^{(-i)}(\hat{\lambda}_i, y_{i0})]. \quad (A.3)$$
This assertion can be verified as follows. Taking expectations with respect to \((\hat{\lambda}_j, y_{j,0})\) for \(j \neq i\) yields

\[
\mathbb{E}_{\theta, \pi, \mathcal{Y}_{i,\pi}}[\hat{p}^{(-i)}(\hat{\lambda}_i, y_{i,0})] = \sum_{j \neq i} \int \int \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_i - \hat{\lambda}_j}{B_N} \right) \frac{1}{B_N} \phi \left( \frac{y_{0} - y_{j,0}}{B_N} \right) p(\hat{\lambda}_j, y_{j,0}) d\hat{\lambda}_j dy_{j,0}
\]

\[
= \int \int \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_i - \hat{\lambda}_j}{B_N} \right) \frac{1}{B_N} \phi \left( \frac{y_{0} - y_{j,0}}{B_N} \right) p(\hat{\lambda}_j, y_{j,0}) d\hat{\lambda}_j dy_{j,0}
\]

The second equality follows from the symmetry with respect to \(j\) and the fact that we integrate out \((\hat{\lambda}_j, y_{j,0})\). We now substitute in

\[
p(\hat{\lambda}_j, y_{j,0}) = \int p(\hat{\lambda}_j | \lambda_j) \pi(\lambda_j, y_{j,0}) d\lambda_j,
\]

where

\[
p(\hat{\lambda}_j | \lambda_j) = \frac{1}{\sigma/T} \phi \left( \frac{\hat{\lambda}_j - \lambda_j}{\sigma/T} \right),
\]

and change the order of integration. This leads to:

\[
\mathbb{E}_{\theta, \pi, \mathcal{Y}_{i,\pi}}[\hat{p}^{(-i)}(\hat{\lambda}_i, y_{i,0})]
\]

\[
= \int \int \left[ \int \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_i - \hat{\lambda}_j}{B_N} \right) p(\hat{\lambda}_j | \lambda_j) d\hat{\lambda}_j \right] \frac{1}{B_N} \phi \left( \frac{y_{0} - y_{j,0}}{B_N} \right) \pi(\lambda_j, y_{j,0}) d\lambda_j dy_{j,0}
\]

\[
= \int \frac{1}{\sqrt{\sigma^2/T + B_N^2}} \phi \left( \frac{\hat{\lambda}_i - \lambda_j}{\sqrt{\sigma^2/T + B_N^2}} \right) \frac{1}{B_N} \phi \left( \frac{y_{0} - y_{j,0}}{B_N} \right) \pi(\lambda_j, y_{j,0}) d\lambda_j dy_{j,0}
\]

\[
= \int \frac{1}{\sqrt{\sigma^2/T + B_N^2}} \phi \left( \frac{\hat{\lambda}_i - \lambda_j}{\sqrt{\sigma^2/T + B_N^2}} \right) \left[ \int \frac{1}{B_N} \phi \left( \frac{y_{0} - y_{j,0}}{B_N} \right) \pi(y_{j,0} | \lambda_j) dy_{j,0} \right] \pi(\lambda_j) d\lambda_j.
\]

Now re-label \(\lambda_j\) and \(\lambda_i\) and \(y_{j,0}\) as \(\tilde{y}_{i,0}\) to obtain:

\[
p_i(\tilde{\lambda}_i, \tilde{y}_{i,0})
\]

\[
= \int \frac{1}{\sqrt{\sigma^2/T + B_N^2}} \phi \left( \frac{\tilde{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T + B_N^2}} \right) \left[ \int \frac{1}{B_N} \phi \left( \frac{\tilde{y}_{i,0} - \tilde{y}_{i,0}}{B_N} \right) \pi(\tilde{y}_{i,0} | \lambda_i) d\tilde{y}_{i,0} \right] \pi(\lambda_i) d\lambda_i.
\]

**Risk Decomposition.** We begin by decomposing the forecast error. Let

\[
\mu(\lambda, \omega^2, p(\lambda, y_0)) = \lambda + \omega^2 \frac{\partial \ln p(\lambda, y_0)}{\partial \lambda}.
\]  

(A.4)
Using the previously developed notation, we expand the prediction error due to parameter estimation as follows:

\[
\hat{Y}_{iT+1} - \lambda_i - \rho Y_{iT} = \left[ \mu(\hat{\lambda}_i(\hat{\rho}), \hat{\sigma}^2/T + B_N^2, \hat{p}(-i)(\hat{\lambda}_i(\hat{\rho}), Y_{i0})) \right]_{CN} - m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) + m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i + (\hat{\rho} - \rho) Y_{iT}
\]

Now write

\[
N \mathbb{E}_{\theta, \pi}^N \left[ \left( \lambda_i + \rho Y_{iT} - \hat{Y}_{iT+1} \right)^2 \right] = N \mathbb{E}_{\theta, \pi}^N [(A_{1i} + A_{2i} + A_{3i})^2].
\]

We deduce from the \( C_r \) inequality that the statement of the theorem follows if we can show that

(i) \( N \mathbb{E}_{\theta, \pi}^N [A_{1i}^2] = o_{u, \pi}(N^{\epsilon_0}) \),

(ii) \( \limsup_{N \to \infty} \sup_{\pi \in \Pi} \frac{N \mathbb{E}_{\theta, \pi}^N, \lambda_i^\pi [A_{2i}^2]}{N \mathbb{E}_{\theta, \pi}^N, \lambda_i^\pi [(\lambda_i - \mathbb{E}_{\theta, \pi, y_0}^\lambda_i)^2] + N^{\epsilon_0}} \leq 1 \),

(iii) \( N \mathbb{E}_{\theta, \pi}^N [A_{3i}^2] = o_{u, \pi}(N^+) \).

The required bounds are provided in Lemmas A.3 (term \( A_{1i} \)), A.4 (term \( A_{2i} \)), A.5 (term \( A_{3i} \)).

### A.1.3 Three Important Lemmas

**Truncations.** The remainder of the proof involves a number of truncations that we will apply when analyzing the risk terms. We take the sequence \( C_N \) as given from Assumption 3.3. Recall that

\[
\frac{2}{M_2} \ln N \leq C_N < \frac{1}{B_N}.
\]

We introduce a new sequence diverging sequence \( L_N \) with the properties

\[
\liminf_N L_N B_N > 1 \quad \text{and} \quad L_N = o(N^+). \quad (A.5)
\]

Even though we do not indicate this explicitly through our notation, we also restrict the domain of \((\lambda, y_0)\) arguments that appear in numerous expressions throughout the proof.
to the support of the distribution of the random variables \((\lambda_i, Y_{i0})\), which is defined as \(\text{Supp}_{\lambda,Y_{0}} = \{ (\lambda, y_0) \in \mathbb{R}^2 | \pi(\lambda, y_0) > 0 \}\).

1. Define the truncated region \(T_\lambda = \{|\lambda| \leq C_N\}\). From Assumption 3.2 we obtain for \(C_N \geq M_3\) that

\[
N^{1-\epsilon} \mathbb{P}(T^c_\lambda) \leq M_1 \exp \left( (1 - \epsilon) \ln N - M_2(C_N - M_3) \right) \tag{A.6}
= \tilde{M}_1 \exp \left( -M_2 \left[ C_N - \frac{1 - \epsilon}{M_2} \ln N \right] \right)
= o(1),
\]

for all \(0 < \epsilon\) because, according to Assumption 3.3, \(C_N > 2(\ln N)/M_2\). Thus, we can deduce

\[
N \mathbb{P}(T^c_\lambda) = o_{u,\pi}(N^+).
\]

2. Define the truncated region \(T_{Y_0} = \{ \max_{1 \leq i \leq N} |Y_{i0}| \leq L_N \}\). Then,

\[
N^{1-\epsilon} \mathbb{P}(T^c_{Y_0}) = N^{1-\epsilon} \mathbb{P}\{ \max_{1 \leq i \leq N} |Y_{i0}| \geq L_N \} \leq N^{1-\epsilon} \sum_{i=1}^{N} \mathbb{P}\{ |Y_{i0}| \geq L_N \}
= N^{2-\epsilon} \int_{|y_0| \geq L_N} \pi(y_0) dy_0
\leq \tilde{M}_1 \exp \left( -M_2 \left[ L_N - \frac{2 - \epsilon}{M_2} \ln N \right] \right)
= o(1),
\]

for all \(\epsilon > 0\) because according to (A.5) \(L_N > (2/M_2) \ln N\). Thus, we deduce that

\[
N \mathbb{P}(T^c_{Y_0}) = o_{u,\pi}(N^+).
\]

3. Define the truncated region \(T_\rho = \{|\hat{\rho} - \rho| \leq 1/L_N^2\}\). By Chebyshev’s inequality, Assumption 3.6, and (A.5), we can bound

\[
N \mathbb{P}(T^c_\rho) = N \mathbb{P}\{ |\hat{\rho} - \rho| > 1/L_N^2 \} \leq L_N^4 \mathbb{E} \left[ N(\hat{\rho} - \rho)^2 \right] = o_{u,\pi}(N^+). \tag{A.8}
\]

4. Define the truncated region \(T_{\sigma^2} = \{|\hat{\sigma}^2 - \sigma^2| \leq 1/L_N\}\). By Chebyshev’s inequality,
Assumption 3.6, and (A.5), we can bound

\[
N \mathbb{P}(\cT_{\Delta}^c) = N \mathbb{P}\{ |\hat{\sigma}^2 - \sigma^2| > 1/L_N \} \leq L_N^2 \mathbb{E}[N(\hat{\sigma}^2 - \sigma^2)^2] = o_{u.p}(N^+). \tag{A.9}
\]

5. Let \( \bar{U}_{i-1}(\rho) = \frac{1}{T} \sum_{t=2}^{T} U_{it-1}(\rho) \) and \( U_{it}(\rho) = U_{it} + \rho U_{it-1} + \cdots + \rho^{i-1} U_{i1} \). Define the truncated region \( \cT_{\bar{U}} = \{ \max_{1 \leq i \leq N} |\bar{U}_{i-1}(\rho)| \leq L_N \} \). Notice that \( \bar{U}_{i-1}(\rho) \sim iid N(0, \sigma^2_{\bar{U}}) \) with \( 0 < \sigma^2_{\bar{U}} < \infty \). Thus, we have

\[
N \mathbb{P}(\cT_{\bar{U}}^c) = N \mathbb{P}\{ \max_{1 \leq i \leq N} |\bar{U}_{i-1}(\rho)| \geq L_N \} \leq N \sum_{i=1}^{N} \mathbb{P}\{ |\bar{U}_{i-1}(\rho)| \geq L_N \} = N^2 \mathbb{P}\{ |\bar{U}_{i-1}(\rho)| \geq L_N \} \\
\leq 2N^2 \exp\left(-\frac{L_N^2}{2\sigma_{\bar{U}}^2}\right) = 2 \exp\left(-\frac{L_N^2}{2\sigma_{\bar{U}}^2} + 2 \ln N\right) \\
\leq 2 \exp\left(-2 \left( \frac{\ln N}{M_2^2\sigma_{\bar{U}}^2} - 1 \right) \ln N\right) = o_{u.p}(N^+),
\]

where the last inequality holds by (A.5).

6. Let \( \bar{Y}_{i-1} = C_1(\rho) Y_{i0} + C_2(\rho) \lambda_i + \bar{U}_{i-1}(\rho) \), where \( C_1(\rho) = \frac{1}{T} \sum_{t=1}^{T} \rho^{i-1} \), \( C_2(\rho) = \frac{1}{T} \sum_{t=2}^{T} (1 + \cdots + \rho^{i-2}) \). Because \( T \) is finite and \( |\rho| \) is bounded, there exists a finite constant, say \( M \) such that \( |C_1(\rho)| \leq M \) and \( |C_2(\rho)| \leq M \). Then, in the region \( \cT_{\lambda} \cap \cT_{\lambda_0} \cap \cT_{\bar{U}} \):

\[
\max_{1 \leq i \leq N} |\bar{Y}_{i-1}| \leq |C_1(\rho)| \max_{1 \leq i \leq N} |\lambda_i| + |C_2(\rho)| \max_{1 \leq i \leq N} |Y_{i0}| + \max_{1 \leq i \leq N} |\bar{U}_{i-1}(\rho)| \\
\leq M(C_N + L_N + L_N),
\]

which leads to

\[
\max_{1 \leq i \leq N} |\bar{Y}_{i-1} - Y_{i-1}| \leq 2 \max_{1 \leq i \leq N} |\bar{Y}_{i-1}| \leq 2 M (C_N + 2L_N) = o_{u.p}(N^+). \tag{A.11}
\]

7. For the region \( \cT_{\lambda} \cap \cT_{\lambda_0} \cap \cT_{\bar{U}} \cap \cT_{\hat{\rho}} \) and with some finite constant \( M \), we obtain the bound

\[
\max_{1 \leq i,j \leq N} |(\hat{\rho} - \rho)(\bar{Y}_{j-1} - Y_{i-1})| \leq \frac{M(C_N + L_N)}{L_N^2} = o_{u.p}(N^+). \tag{A.12}
\]
8. Define the regions $\mathcal{T}_m = \{ |m(\hat{\lambda}_i, Y_{i0})| \leq C_N \}$ and $\mathcal{T}_{m*} = \{ |m_*(\hat{\lambda}_i, Y_{i0})| \leq C_N \}$. By Chebyshev’s inequality and Assumption 3.5, we deduce

$$N \mathbb{P}(\mathcal{T}^c_m) \leq \frac{1}{C_N^2} N \mathbb{E}(m(\hat{\lambda}_i, Y_{i0})^2 \mathcal{T}^c_m) \leq o_{u,\pi}(N^+) \quad (A.13)$$

$$N \mathbb{P}(\mathcal{T}^c_{m*}) \leq \frac{1}{C_N^2} N \mathbb{E}(m_*(\hat{\lambda}_i, Y_{i0})^2 \mathcal{T}^c_{m*}) \leq o_{u,\pi}(N^+).$$

We will subsequently use indicator function notation, abbreviating, say, $\mathbb{I}\{\lambda \in \mathcal{T}_\lambda\}$ by $\mathbb{I}(\mathcal{T}_\lambda)$ and $\mathbb{I}(\mathcal{T}_\lambda)\mathbb{I}(\mathcal{T}_{Y_0})$ by $\mathbb{I}(\mathcal{T}_\lambda \mathcal{T}_{Y_0})$.

A.1.3.1 Term $A_{1i}$

**Lemma A.3** Suppose the assumptions in Theorem 3.7 hold. Then,

$$N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}\left[ \left[ \left( \mu(\lambda_i(\hat{\rho}), \hat{\sigma}^2/T + B_N^2, \hat{\rho}^{(-i)}(\lambda_i(\hat{\rho}), Y_{i0})) \right) \right]_{C_N} - m_*(\lambda_i, y_{i0}; \pi, B_N) \right]^2 = o_{u,\pi}(N^{c_{\mathcal{T}}}).$$

**Proof of Lemma A.3.** We begin with the following bound: since $(a + b)^2 \leq 2a^2 + 2b^2$,

$$|A_{1i}|^2 = \left( \left[ \mu(\lambda_i(\hat{\rho}), \hat{\sigma}^2/T + B_N^2, \hat{\rho}^{(-i)}(\lambda_i(\hat{\rho}), Y_{i0})) \right]_{C_N} - m_*(\lambda_i, y_{i0}; \pi, B_N) \right)^2 \leq 2C_N^2 + 2m_N^2(\hat{\lambda}_i, y_{i0}; \pi, B_N)$$

$$= 2C_N^2 + 2m_N^2(\hat{\lambda}_i, y_{i0}; \pi, B_N) \mathbb{I}(\mathcal{T}_{m*}) + 2m_N^2(\hat{\lambda}_i, y_{i0}; \pi, B_N) \mathbb{I}(\mathcal{T}^c_{m*}) \leq 4C_N^2 + 2m_N^2(\hat{\lambda}_i, y_{i0}; \pi, B_N) \mathbb{I}(\mathcal{T}_{m*}). \quad (A.14)$$

Then,

$$N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}[A_{1i}^2] \leq N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}[A_{1i}^2 \mathbb{I}(\mathcal{T}_{\mathcal{T}_\mathcal{U}} \mathcal{T}_{\mathcal{T}_{Y_0}} \mathcal{T}_{m*} \mathcal{T}_\lambda)] + N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}[A_{1i}^2 \mathbb{I}(\mathcal{T}^c_{\mathcal{T}_\mathcal{U}} \mathcal{T}^c_{\mathcal{T}_{Y_0}} \mathcal{T}^c_{m*} \mathcal{T}_\lambda)]$$

$$\leq N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}[A_{1i}^2 \mathbb{I}(\mathcal{T}_{\mathcal{T}_\mathcal{U}} \mathcal{T}_{\mathcal{T}_{Y_0}} \mathcal{T}_{m*} \mathcal{T}_\lambda)] + 4C_N^2 N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}[\mathbb{I}(\mathcal{T}_{\mathcal{T}_\mathcal{U}} \mathcal{T}_{\mathcal{T}_{Y_0}} \mathcal{T}_{m*} \mathcal{T}_\lambda)]$$

$$+ 12N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}[A_{1i}^2 \mathbb{I}(\mathcal{T}_{\mathcal{T}_\mathcal{U}} \mathcal{T}_{\mathcal{T}_{Y_0}} \mathcal{T}_{m*} \mathcal{T}_\lambda)]$$

$$\leq N \mathbb{E}_{\theta,\pi}^{Y_\mathcal{T}}[A_{1i}^2 \mathbb{I}(\mathcal{T}_{\mathcal{T}_\mathcal{U}} \mathcal{T}_{\mathcal{T}_{Y_0}} \mathcal{T}_{m*} \mathcal{T}_\lambda)] + o_{u,\pi}(N^+) + o_{u,\pi}(N^+).$$

The first $o_{u,\pi}(N^+)$ follows from the properties of the truncation regions discussed above and the second $o_{u,\pi}(N^+)$ follows from Assumption 3.5. In the remainder of the proof we will
construct the desired bound, $o_{u,\pi}(N^{\epsilon_0})$, for the first term on the right-hand side of (A.15). We proceed in two steps.

**Step 1.** We introduce two additional truncation regions, $\mathcal{T}_{\hat{\lambda}Y_0}$ and $\mathcal{T}_{p(\cdot)}$, which are defined as follows:

\[
\mathcal{T}_{\hat{\lambda}Y_0} = \left\{ (\hat{\lambda}_i, Y_{i0}) \mid -C'_N \leq \hat{\lambda}_i \leq C'_N, -C'_N \leq Y_{i0} \leq C'_N \right\},
\]

\[
\mathcal{T}_{p(\cdot)} = \left\{ (\hat{\lambda}_i, Y_{i0}) \mid p(\hat{\lambda}_i, Y_{i0}) \geq \frac{N^\epsilon}{N} \right\},
\]

where it is assumed that $0 < \epsilon < \epsilon_0$.

Notice that since $C_N = o(N^+)$ and $\sqrt{\ln N} = o(N^+)$, $C'_N = o(N^+)$ (A.16).

In the first truncation region both $\hat{\lambda}_i$ and $Y_{i0}$ are bounded by $C'_N$. In the second truncation region the density $p(\hat{\lambda}_i, Y_{i0})$ is not too low. We will show that

\[
N\mathbb{E}^Y_{\theta,\pi} \left[ A^2_{1i}(\mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}) \right] \leq o_{u,\pi}(N^{\epsilon_0}) \quad (A.17)
\]

\[
N\mathbb{E}^Y_{\theta,\pi} \left[ A^2_{1i}(\mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}) \right] \leq o_{u,\pi}(N^+). \quad (A.18)
\]

**Step 1.1.** First, we consider the case where $(\hat{\lambda}_i, Y_{i0})$ are bounded and the density $p(\hat{\lambda}_i, y_{i0})$ is “low” in (A.17). Using the bound for $|A_{1i}|$ in (A.14) we obtain:

\[
N\mathbb{E}^Y_{\theta,\pi} \left[ A^2_{1i}(\mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}) \right] \\
\leq 4NC^2_N \mathbb{P}(\mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}) + 2N\mathbb{E}^Y_{\theta,\pi} \left[ m^2_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) (\mathcal{T}_{m(\cdot)}) \right] \\
= 4NC^2_N \int_{\hat{\lambda}_i = -C'_N}^{C'_N} \int_{y_{i0} = -C'_N}^{C'_N} \mathbb{I} \left\{ p(\hat{\lambda}_i, y_{i0}) < \frac{N^\epsilon}{N} \right\} p(\hat{\lambda}_i, y_{i0}) d(\hat{\lambda}_i, y_{i0}) + o_{u,\pi}(N^+) \\
\leq 4NC^2_N \int_{\hat{\lambda}_i = -C'_N}^{C'_N} \int_{y_{i0} = -C'_N}^{C'_N} \left( \frac{N^\epsilon}{N} \right) dy_{i0} d\hat{\lambda}_i + o_{u,\pi}(N^+) \\
= 4C^2_N (2C'_N)^2 N^\epsilon + o_{u,\pi}(N^+) \\
\leq o_{u,\pi}(N^{\epsilon_0}).
\]

The $o_{u,\pi}(N^+)$ term in the first equality follows from Assumption 3.5. The last equality holds because $C_N, C'_N = o_{u,\pi}(N^+)$ (Assumption 3.3 and (A.16)) and $0 < \epsilon < \epsilon_0$. This establishes (A.17).
Step 1.2. Next, we consider the case where \((\hat{\lambda}_i, y_{i0})\) exceed the \(C_N'\) bound and the density \(p(\hat{\lambda}_i, y_{i0})\) is “high.” We will immediately replace the contribution of \(2Nm^2(\hat{\lambda}_i, y_{i0}; \pi, B_N)\Pi(\mathcal{T}_m^c)\) to the expected value of \(A^2_{i1}\) by \(o_u\pi(N^+)\).

\[
N\mathbb{E}^N_{\pi} \left[ A^2_{i1} \Pi(\mathcal{T}_\lambda^c \mathcal{T}_{Y_{i0}}^c) \right] \\
\leq 4NC_N^2 \mathbb{P}(\mathcal{T}_\lambda^c \cap \mathcal{T}_{Y_{i0}}^c) + o_u\pi(N^+) \\
= 4NC_N^2 \int_{Y_{i0}} \left[ \int_{\lambda_i \geq C_N} \frac{1}{\sigma/\sqrt{T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sigma/\sqrt{T}} \right) \pi(\lambda_i, y_{i0}) d\lambda_i \right] d(\hat{\lambda}_i, y_{i0}) + o_u\pi(N^+) \\
\leq 4NC_N^2 \int_{\lambda_i \geq C_N} \int_{|y_{i0}| > C_N} \left[ \int \frac{1}{\sigma/\sqrt{T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sigma/\sqrt{T}} \right) d\hat{\lambda}_i \right] \pi(\lambda_i) d\lambda_i + 4NC_N^2 \int_{|y_{i0}| > C_N} \left[ \int \pi(\lambda_i | y_{i0}) d\lambda_i \right] \pi(\lambda_i) dy_{i0} + o_u\pi(N^+) \\
\leq 4NC_N^2 \int_{\lambda_i \leq C_N} \left[ \int \frac{1}{\sigma/\sqrt{T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sigma/\sqrt{T}} \right) d\hat{\lambda}_i \right] \pi(\lambda_i) d\lambda_i \\
+ 4NC_N^2 \int_{|y_{i0}| > C_N} \left[ \int \pi(\lambda_i | y_{i0}) d\lambda_i \right] \pi(\lambda_i) dy_{i0} + o_u\pi(N^+) \\
= B_1 + o_u\pi(N^+) + o_u\pi(N^+), \text{ say.}
\]

The second equality is obtained by integrating out \(\hat{\lambda}_i\), recognizing that the integrand is a properly scaled probability density function that integrates to one. The last line follows from the calculations in (A.6), Lemma A.2, and \(C_N' > C_N\).
We will first analyze term $B_1$. Note by the change of variable that

$$
\int_{|\lambda_i| > C_N'} \frac{1}{\sigma/\sqrt{T}} \phi \left( \frac{\lambda_i - \mu_i}{\sigma/\sqrt{T}} \right) d\lambda_i
= \int_{-\infty}^{-\sqrt{T}(C_N' + |\lambda_i|)/\sigma} \phi(\lambda_i) d\lambda_i + \int_{\sqrt{T}(C_N' - |\lambda_i|)/\sigma}^{\infty} \phi(\lambda_i) d\lambda_i
\leq \int_{-\infty}^{-\sqrt{T}(C_N' - |\lambda_i|)/\sigma} \phi(\lambda_i) d\lambda_i + \int_{\sqrt{T}(C_N' - |\lambda_i|)/\sigma}^{\infty} \phi(\lambda_i) d\lambda_i
\leq 2 \int_{-\sqrt{T}(C_N' - |\lambda_i|)/\sigma}^{\infty} \phi(\lambda_i) d\lambda_i
\leq 2 \frac{\phi(\sqrt{T}(C_N' - |\lambda_i|)/\sigma)}{\sqrt{T}(C_N' - |\lambda_i|)/\sigma} \pi(\lambda_i),
$$

where we used the inequality $\int_x^\infty \phi(x) d\lambda \leq \phi(x)/x$. Using the definition of $C_N'$ in Assumption 3.3 we obtain the bound (for $\sqrt{2 \ln N} \geq 1$):

$$
B_1 \leq 4NC_N^2 \int_{|\lambda| < C_N} \frac{\phi(\sqrt{T}(C_N' - |\lambda|)/\sigma)}{\sqrt{T}(C_N' - |\lambda|)/\sigma} \pi(\lambda_i) d\lambda_i
\leq 4NC_N^2 \int_{|\lambda| < C_N} \phi(\sqrt{2 \ln N}) \pi(\lambda_i) d\lambda_i
\leq 4NC_N^2 \exp \left( - \ln N \right) \int_{|\lambda| < C_N} \pi(\lambda_i) d\lambda_i
\leq 4C_N^2
= o_u.(N^+).
$$

This leads to the desired bound in (A.18).

Step 2. It remains to be shown that

$$
N^E_{\theta, \pi}[A_1^2 \| \mathcal{T}_{\mathcal{T}} \mathcal{T}_U \mathcal{T}_Y \mathcal{T}_m \mathcal{T}_\lambda \mathcal{T}_{\lambda_{\mathcal{Y}_0}} \mathcal{T}_{\theta(\cdot)}] \leq o_u.(N^+). \quad (A.19)
$$
We introduce the following notation:

\[
\begin{align*}
\tilde{p}_{-i}^{-} &= p_{-i}(\hat{\lambda}(\hat{\rho}), Y_{i0}) \\
\hat{p}_{-i}^{-} &= \frac{1}{\hat{\lambda}_{i}(\hat{\rho})} \partial \hat{p}^{-i}(\hat{\lambda}(\hat{\rho}), Y_{i0}) \\
\hat{p}_{i}^{-} &= p_{i}(\hat{\lambda}_{i}(\rho), Y_{i0}) \\
\hat{p}_{i}^{-} &= \frac{1}{\partial \hat{\lambda}_{i}(\rho)} \partial p^{-i}(\hat{\lambda}(\rho), Y_{i0}) \\
p_{i} &= p(\hat{\lambda}_{i}(\rho), Y_{i0}) \\
p_{*i} &= p_{*}(\hat{\lambda}_{i}(\rho), Y_{i0}) \\
dp_{*i} &= \frac{1}{\partial \hat{\lambda}_{i}(\rho)} \partial p_{*}(\hat{\lambda}(\rho), Y_{i0}).
\end{align*}
\]

Moreover, we introduce another truncation

\[
\mathcal{T}_{\tilde{p}(\cdot)} = \left\{ (\hat{\lambda}_{i}, Y_{i0}) \mid \tilde{p}_{i}^{-} > \frac{p_{*i}}{2} \right\} .
\]

On the set \( \mathcal{T}_{m*} \), we have \( |m_{*}(\hat{\lambda}_{i}, y_{i0}; \pi, B_{N})| \leq C_{N} \), and so

\[
|A_{11}| \leq 2C_{N}.
\]

For the required result of Step 2 in (A.19), we show the following two inequalities; see Steps 2.1 and 2.2 below:

\[
\begin{align*}
\mathbb{E}_{\theta, \pi}^{Y_{i0}} \left[ A_{11}^{2} \mathcal{T}_{\tilde{p}(\cdot)} \right] &\leq o_{u.\pi}(N^{+}) \\
\mathbb{E}_{\theta, \pi}^{Y_{i0}} \left[ A_{11}^{2} \mathcal{T}_{\tilde{p}(\cdot)} \right] &\leq o_{u.\pi}(N^{+}).
\end{align*}
\]
Step 2.1. Using the triangle inequality, we obtain

\[
|A_{1i}| = \left| \mu(\hat{\lambda}_i(\hat{\rho}), Y_{i0}, \hat{\sigma}_i^2/T + B_N^2, \hat{p}_i^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0})) \right|^{C_N}_{m_*} - \left| m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) \right|
\]

\[
\leq \left| \mu(\hat{\lambda}_i(\hat{\rho}), Y_{i0}, \hat{\sigma}_i^2/T + B_N^2, \hat{p}_i^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0})) - m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) \right|
\]

\[
= \hat{\lambda}_i(\hat{\rho}) - \lambda_i(\rho) + \left( \frac{\hat{\sigma}_i^2}{T} - \frac{\sigma_i^2}{T} \right) \frac{dp_{si}}{p_{si}} + \left( \frac{\hat{\sigma}_i^2}{T} + B_N^2 \right) \left( \frac{dp_{si}^{(-i)}}{p_{si}^{(-i)}} - \frac{dp_{si}}{p_{si}} \right)
\]

\[
\leq |\hat{\rho} - \rho| |\hat{Y}_{i-1}| + \left| \frac{\hat{\sigma}_i^2}{T} - \frac{\sigma_i^2}{T} \right| \frac{dp_{si}}{p_{si}} + \left( \frac{\hat{\sigma}_i^2}{T} + B_N^2 \right) \left( \frac{dp_{si}^{(-i)}}{p_{si}^{(-i)}} - \frac{dp_{si}}{p_{si}} \right),
\]

\[
= A_{11i} + A_{12i} + A_{13i}, \quad \text{say.}
\]

Recall that \( \hat{Y}_{i-1} = \frac{1}{T} \sum_{t=1}^{T} Y_{it-1} \). Using the Cauchy-Schwarz inequality, it suffices to show that

\[
N\mathbb{E}_{Y}^{Y_\pi} \left[ A_{11i}^2 \mathcal{T}_2^2 \mathcal{T}_U \mathcal{T}_0 \mathcal{T}_{m^*} \mathcal{T}_Y \mathcal{T}_{\lambda Y_0} \mathcal{T}_{\hat{\rho}(i)} \mathcal{T}_{\hat{p}(i)} \right] \leq o_{u,\pi}(N^+), \quad j = 1, 2, 3.
\]

For term A_{11i}. First, using a slightly more general argument than the one used in the proof of Lemma A.5 below, we can show that

\[
N\mathbb{E}_{Y}^{Y_\pi} \left[ A_{11i}^2 \right] = \mathbb{E}_{Y}^{Y_\pi} \left[ N(\hat{\rho} - \rho)^2 \hat{Y}_{i-1}^2 \right] = o_{u,\pi}(N^+).
\]

For term A_{12i}. Second, in the region \( \mathcal{T}_{\lambda Y_0} \cap \mathcal{T}_{m^*} \) we can bound the Tweedie correction term under \( p_{si} \) by

\[
\left( \frac{\sigma_i^2}{T} + B_N^2 \right) \frac{dp_{si}}{p_{si}} = \left| m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \hat{\lambda}_i(\rho) \right| \leq C_N + C_N'. \tag{A.25}
\]

Using Assumption 3.3, Assumption 3.6, and (A.16), we obtain the bound

\[
N\mathbb{E}_{Y}^{Y_\pi} \left[ A_{12i}^2 \mathcal{T}_0^2 \right] \leq \frac{1}{(\sigma_i^2/T + B_N^2)^2} \mathbb{E}_{Y}^{Y_\pi} \left[ N(\hat{\sigma}_i^2 - \sigma_i^2)^2 \right] (C_N' + C_N)^2 = o_{u,\pi}(N^+).
\]

For term A_{13i}. Finally, note that

\[
A_{13i}^2 \mathcal{T}_2^2 \leq \left( \frac{\sigma_i^2}{T} + B_N^2 + \frac{1}{T} \frac{1}{L_N} \right)^2 \left( \frac{dp_{si}^{(-i)}}{p_{si}^{(-i)}} - \frac{dp_{si}}{p_{si}} \right)^2.
\]
Thus, the desired result follows if we show

\[
N E_{\theta, \pi}^{2N} \left[ \left( \frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - \frac{dp_{si}}{p_{si}} \right)^2 \| (T_\theta T_U T_\gamma T_{\lambda_Y} T_{\lambda_Y} T_{p(\gamma)} T_{\bar{p}(\gamma)} ) \right] = o_{u, \pi} (N^+) \quad (A.26)
\]

Decompose

\[
\frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - \frac{dp_{si}}{p_{si}} = \frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - \frac{dp_{si}}{p_{si}} - \frac{dp_{si}}{p_{si}} \left( \frac{\bar{p}_i^{(-i)}}{p_{si}} - p_{si} + p_{si} \right).
\]

Using the \( C_r \) inequality, we obtain

\[
N E_{\theta, \pi}^{2N} \left[ \left( \frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - \frac{dp_{si}}{p_{si}} \right)^2 \| (T_\theta T_U T_\gamma T_{\lambda_Y} T_{\lambda_Y} T_{p(\gamma)} T_{\bar{p}(\gamma)} ) \right] 
\leq 2N E_{\theta, \pi}^{2N} \left[ \left( \frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - \frac{dp_{si}}{p_{si}} \right)^2 \| (T_\theta T_U T_\gamma T_{\lambda_Y} T_{\lambda_Y} T_{p(\gamma)} T_{\bar{p}(\gamma)} ) \right]
+ 2N E_{\theta, \pi}^{2N} \left[ \left( \frac{dp_{si}}{p_{si}} \right)^2 \left( \frac{\bar{p}_i^{(-i)}}{p_{si}} - p_{si} + p_{si} \right)^2 \| (T_\theta T_U T_\gamma T_{\lambda_Y} T_{\lambda_Y} T_{p(\gamma)} T_{\bar{p}(\gamma)} ) \right]
\]

\[
= 2E_{\theta, \pi}^{2N} \left[ \left( \sqrt{N} \left( \frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - \frac{dp_{si}}{p_{si}} \right) \right)^2 \| (T_\theta T_U T_\gamma T_{\lambda_Y} T_{\lambda_Y} T_{p(\gamma)} T_{\bar{p}(\gamma)} ) \right]
+ 2o_{u, \pi} (N^+) E_{\theta, \pi}^{2N} \left[ \left( \sqrt{N} \left( \frac{\bar{p}_i^{(-i)}}{p_{si}} - p_{si} \right) \right)^2 \| (T_\theta T_U T_\gamma T_{\lambda_Y} T_{\lambda_Y} T_{p(\gamma)} T_{\bar{p}(\gamma)} ) \right]
\]

\[
= 2B_{1i} + 2o_{u, \pi} (N^+) B_{2i},
\]

say. The \( o_{u, \pi} (N^+) \) bound follows from (A.25). Using the mean-value theorem, we can express

\[
\sqrt{N} \left( \frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - dp_{si} \right) = \sqrt{N} \left( \frac{dp_i^{(-i)}}{\bar{p}_i^{(-i)}} - dp_{si} \right) + \sqrt{N} (\hat{\rho} - \rho) R_{1i}(\bar{\rho})
\]

\[
\sqrt{N} \left( \frac{\bar{p}_i^{(-i)}}{p_{si}} - p_{si} \right) = \sqrt{N} \left( \frac{\bar{p}_i^{(-i)}}{p_{si}} - p_{si} \right) + \sqrt{N} (\hat{\rho} - \rho) R_{2i}(\bar{\rho}),
\]

respectively.
where

\[ R_{1i}(\rho) = -\frac{1}{N-1} \sum_{j \neq i}^{N} \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right) \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right)^2 \left( \hat{Y}_{j,-1} - \hat{Y}_{i,-1} \right) \frac{1}{B_N} \phi \left( \frac{Y_{j0} - Y_{i0}}{B_N} \right), \]

\[ R_{2i}(\rho) = \frac{1}{N-1} \sum_{j \neq i}^{N} \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right) \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right) \left( \hat{Y}_{j,-1} - \hat{Y}_{i,-1} \right) \frac{1}{B_N} \phi \left( \frac{Y_{j0} - Y_{i0}}{B_N} \right), \]

and \( \hat{\rho} \) is located between \( \hat{\rho} \) and \( \rho \).

We proceed with the analysis of \( B_{2i} \). Over the region \( \mathcal{T}_{\hat{p}(\cdot)} \), \( \tilde{p}_{i}^{(-i)} - p_{s_i} + p_{s_i} > p_{s_i}/2 \). Using this, the \( C_r \) inequality, and the law of iterated expectations, we obtain

\[ B_{2i} \leq 4 \mathbb{E}_{\theta, \pi}^{\gamma_i} \left[ \frac{1}{p_{s_i}} \mathbb{E}_{\theta, \pi, Y_i}^{\gamma_i} \left[ N (\tilde{p}_{i}^{(-i)} - p_{s_i})^2 \mathbb{I}(T_{p} T_{\mathcal{T}_y} T_{\mathcal{T}_\pi} T_{\mathcal{T}_{\hat{p}(\cdot)}}) \right] \right] + 4 \mathbb{E}_{\theta, \pi}^{\gamma_i} \left[ \frac{1}{p_{s_i}} \mathbb{E}_{\theta, \pi, Y_i}^{\gamma_i} \left[ (\hat{\rho} - \rho)^2 R_{2i}(\hat{\rho}) \mathbb{I}(T_{p} T_{\mathcal{T}_y} T_{\mathcal{T}_\pi} T_{\mathcal{T}_{\hat{p}(\cdot)}}) \right] \right] \]

\[ = 4 \mathbb{E}_{\theta, \pi}^{\gamma_i} [B_{21i} + B_{22i}], \]

say.

According to Lemma A.8(c) (see Section A.1.4),

\[ \mathbb{E}_{\theta, \pi, Y_i}^{\gamma_i} \left[ N (\tilde{p}_{i}^{(-i)} - p_{s_i})^2 \mathbb{I}(T_{p} T_{\mathcal{T}_y} T_{\mathcal{T}_\pi} T_{\mathcal{T}_{\hat{p}(\cdot)}}) \right] \leq \frac{M}{B_N^2} p_{s_i} \mathbb{I}(T_{\mathcal{T}_{\hat{p}(\cdot)}}). \]

This leads to

\[ \mathbb{E}_{\theta, \pi}^{\gamma_i} [B_{21i}] \leq \frac{M}{B_N^2} \mathbb{E}_{\theta, \pi}^{\gamma_i} \left[ \frac{p_{s_i}}{p_{s_i}} \mathbb{I}(T_{\mathcal{T}_{\hat{p}(\cdot)}}) \right] = \frac{M}{B_N^2} \int_{T_{\mathcal{T}_{\hat{p}(\cdot)}}} \frac{p_{s_i}^2}{p_{s_i}} d\hat{\lambda}_i dy_{i0}. \]

According to Lemma A.8(c) (see Section A.1.4),

\[ \int_{T_{\mathcal{T}_{\hat{p}(\cdot)}}} \frac{p_{s_i}^2}{p_{s_i}} d\hat{\lambda}_i dy_{i0} = o_{u, \pi}(N^+). \]

Because \( 1/B_N^2 = o(N^+) \) according to Assumption 3.3, we can deduce that

\[ \mathbb{E}_{\theta, \pi}^{\gamma_i} [B_{21i}] \leq o_{u, \pi}(N^+). \]
Using the Cauchy-Schwarz Inequality, we obtain

\[ B_{2i} \leq \frac{1}{p^{*}_{i}} \sqrt{E_{\theta,\pi,Y}^{(\cdot)}} \left[ N^2(\hat{\rho} - \rho)^4 \right] \sqrt{E_{\theta,\pi,Y}^{(\cdot)}} \left[ R_{2i}^{4}(\hat{\rho})(T_{\hat{\rho}T_{U}T_{Y_{0}}T_{m}T_{\lambda_{0}}T_{p}}(T_{p})) \right]. \]

Using the inequality once more leads to

\[ E_{\theta,\pi}[B_{2i}] \leq o_{u,\pi}(N^+) \sqrt{E_{\theta,\pi}^{(\cdot)}} \left[ R_{2i}^{4}(\hat{\rho})(T_{\hat{\rho}T_{U}T_{Y_{0}}T_{m}T_{\lambda_{0}}T_{p}}(T_{p})) \right]. \]

The second inequality follows from Assumption 3.6.

According to Lemma A.8(a) (see Section A.1.4),

\[ E_{\theta,\pi,Y}^{(\cdot)} \left[ R_{2i}^{4}(\hat{\rho})(T_{\hat{\rho}T_{U}T_{Y_{0}}T_{m}T_{\lambda_{0}}T_{p}}(T_{p})) \right] \leq ML_{N}^{4}p_{i}^{4}\|T_{\lambda_{0}}T_{p}\|, \]

where \( L_{N} = o(N^+) \) was defined in (A.5). This leads to the bound

\[ E_{\theta,\pi}[B_{2i}] \leq o_{u,\pi}(N^+)ML_{N}^{2} \sqrt{E_{\theta,\pi}^{(\cdot)}} \left[ \left( \frac{p_{i}}{p_{*i}} \right)^{4} \|T_{\lambda_{0}}T_{p}\| \right] \]

\[ = o_{u,\pi}(N^+)ML_{N}^{2} \sqrt{\int_{T_{\lambda_{0}} \cap T_{p}} \left( \frac{p_{i}}{p_{*i}} \right)^{4} p_{i} d\lambda_{i} dy_{0}} \]

\[ \leq o_{u,\pi}(N^+)M_{*}L_{N}^{2} \sqrt{\int_{T_{\lambda_{0}} \cap T_{p}} \left( \frac{p_{i}}{p_{*i}} \right)^{4} d\lambda_{i} dy_{0}} \]

\[ \leq o_{u,\pi}(N^+). \]

The second inequality holds because the density \( p_{i} \) is bounded from above and \( M_{*} \) is a constant. The last inequality is proved in Lemma A.8(e) (see Section A.1.4).

We deduce that \( B_{2i} = o_{u,\pi}(N^+) \). A similar argument can be used to establish that \( B_{1i} = o_{u,\pi}(N^+) \).

**Step 2.2.** Recall from (A.22) that over \( T_{m^*} \),

\[ |A_{1i}| \leq 2C_{N} = o_{u,\pi}(N^+). \]
Then,

\[
\mathbb{E}_{\theta,\pi}^{\gamma}( A_{ij}^2 \mathbb{I}(T_{\tilde{\rho}} T_{\tilde{\rho}} T_{\tilde{U}} T_{Y_0} T_{\tilde{\lambda}} T_{\tilde{\lambda}_Y0} T_{p(\cdot)} T_{\tilde{p}(\cdot)}) T_{p(\cdot)}^c) \geq o_{\pi,\pi}(N^+) \mathbb{E}_{\theta,\pi}^{\gamma}(T_{\tilde{\rho}} T_{\tilde{U}} T_{Y_0} T_{\tilde{\lambda}} T_{\tilde{\lambda}_Y0} T_{p(\cdot)} T_{\tilde{p}(\cdot)}).
\]

Notice that

\[
\mathcal{T}_{\tilde{p}(\cdot)}^c = \left\{ \tilde{p}_i^{(-i)} - p_{si} + (\hat{\rho} - \rho) R_{2i}(\hat{\rho}) < -\frac{p_{si}}{2} \right\}
\subset \left\{ \tilde{p}_i^{(-i)} - p_{si} - |\hat{\rho} - \rho||R_{2i}(\hat{\rho})| < -\frac{p_{si}}{2} \right\}
\subset \left\{ \tilde{p}_i^{(-i)} - p_{si} < -\frac{p_{si}}{4} \right\} \cup \left\{ |\hat{\rho} - \rho||R_{2i}(\hat{\rho})| > \frac{p_{si}}{4} \right\}.
\]

Then,

\[
\mathbb{E}_{\theta,\pi,\gamma}^{\gamma(-i)}(T_{\tilde{\rho}} T_{\tilde{U}} T_{Y_0} T_{\tilde{\lambda}} T_{\tilde{\lambda}_Y0} T_{p(\cdot)} T_{\tilde{p}(\cdot)})
\leq \mathbb{E}_{\theta,\pi,\gamma}^{\gamma(-i)} \left\{ \tilde{p}_i^{(-i)} - p_{si} < -\frac{p_{si}}{4} \right\}
+ \mathbb{E}_{\theta,\pi,\gamma}^{\gamma(-i)} \left[ \left\{ |\hat{\rho} - \rho||R_{2i}(\hat{\rho})| > \frac{p_{si}}{4} \right\} \mathbb{I}(T_{\tilde{\rho}} T_{\tilde{U}} T_{Y_0} T_{\tilde{\lambda}} T_{\tilde{\lambda}_Y0} T_{p(\cdot)}) \right]
\leq \mathbb{E}_{\theta,\pi,\gamma}^{\gamma(-i)} \left\{ \tilde{p}_i^{(-i)} - p_{si} < -\frac{p_{si}}{4} \right\} + \frac{16L_4^2}{p_{si}^2} \mathbb{E}_{\theta,\pi,\gamma}^{\gamma(-i)} \left[R_{2i}(\hat{\rho})^2 \mathbb{I}(T_{\tilde{\rho}} T_{\tilde{U}} T_{Y_0} T_{\tilde{\lambda}} T_{\tilde{\lambda}_Y0} T_{p(\cdot)}) \right]
\leq \mathbb{E}_{\theta,\pi,\gamma}^{\gamma(-i)} \left\{ \tilde{p}_i^{(-i)} - p_{si} < -\frac{p_{si}}{4} \right\} + \frac{ML_6^6}{p_{si}^2} \mathbb{I}(T_{\tilde{\lambda}_Y0} T_{p(\cdot)}).
\]

The first inequality is based on the superset of \(\mathcal{T}_{\tilde{p}(\cdot)}^c\) from above. The second inequality is based on Chebychev’s inequality and truncation \(T_{\tilde{\rho}}\). The third inequality uses a version of the result in Lemma A.8(a) in which the remainder is raised to the power of two instead of to the power of four. Assumption 3.4 implies that \(p_i\) is bounded from above:

\[
p_i = \int p(\lambda|\lambda)\pi(Y_{i0}|\lambda)\pi(\lambda)d\lambda \leq M \int \pi(\lambda)d\lambda = \tilde{M} \leq M < \infty, \quad (A.27)
\]

because \(p(\lambda|\lambda)\) is the density of a \(N(\lambda, \sigma^2/T)\) and \(\pi(Y_{i0}|\lambda)\) is bounded for every \(\pi \in \Pi\) according to Assumption 3.4. Thus, in the previous calculation we can absorb one of the \(p_i\) terms in the constant \(M\).

In Lemma A.8(f) (see Section A.1.4) we apply Bernstein’s inequality to bound the prob-
ability \( E^{\mathbf{Y}(-i)}_{\hat{\theta},\pi,\mathbf{y}} \left\{ \hat{p}^{(-i)}_i - p^{\ast} i < -\frac{p^{\ast} i}{4} \right\} \) uniformly over \((\hat{\lambda}_i, Y_0)\) in the region \( T_{\hat{\lambda}_0} \), showing that

\[
N E^{\mathbf{Y}}_{\hat{\theta},\pi} \left[ E^{\mathbf{Y}(-i)}_{\hat{\theta},\pi,\mathbf{y}} \left\{ \hat{p}^{(-i)}_i - p^{\ast} i < -\frac{p^{\ast} i}{4} \right\} \mathbb{I}(T_{\hat{\lambda}_0} T_{p(\cdot)}) \right] = o_u.\pi(N^+),
\]

as desired. Moreover, according to Lemma A.8(e) (see Section A.1.4)

\[
E^{\mathbf{Y}}_{\hat{\theta},\pi} \left[ \frac{p_i^2}{p^{\ast} i} \mathbb{I}(T_{\hat{\lambda}_0} T_{p(\cdot)}) \right] = \int_{T_{\hat{\lambda}_0} \cap T_{p(\cdot)}} \left( \frac{p_i}{p^{\ast} i} \right)^2 d\hat{\lambda}_i dy_0 = o_u.\pi(N^+),
\]

which gives us the required result for Step 2.2. Combining the results from Steps 2.1 and 2.2 yields (A.19).

The bound in (A.15) now follows from (A.17), (A.18), and (A.19), which completes the proof of the lemma. ■

### A.1.3.2 Term \( A_{2i} \)

**Lemma A.4** Suppose the assumptions in Theorem 3.7 hold. Then, for every \( \epsilon_0 > 0 \),

\[
\lim_{N \to \infty} \sup_{\pi \in \Pi} \frac{N E^{\mathbf{Y},\lambda}_i \left[ \left( m^* (\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right]}{N E^{\mathbf{Y},\lambda}_i \left[ (\lambda_i - E^{\mathbf{Y},\lambda}_i (\lambda_i))^2 \right] + N \epsilon_0} \leq 1.
\]

**Proof of Lemma A.4.** Recall that \( m^* (\hat{\lambda}_i, y_{i0}; \pi, B_N) \) can be interpreted as the posterior mean of \( \lambda_i \) under the \( p^* (\hat{\lambda}_i, y_{i0}; \pi) \) defined in (17). We will use \( \mathbb{E}^{\mathbf{Y},\lambda}_i [\cdot] \) to denote the integral

\[
\mathbb{E}^{\mathbf{Y},\lambda}_i [\cdot] = \int \cdot \pi_i (\hat{\lambda}_i | \lambda) \pi_\pi (\lambda) d(\hat{\lambda}_i, \lambda, y_0),
\]

where

\[
p^* (\hat{\lambda} | \lambda) = \frac{1}{\sqrt{\sigma^2/T + B_N^2}} \phi \left( \frac{\hat{\lambda} - \lambda}{\sqrt{\sigma^2/T + B_N^2}} \right)
\]

and

\[
\pi^* (y_0 | \lambda) = \int \frac{1}{B_N} \phi \left( \frac{y_0 - \tilde{y}_0}{B_N} \right) \pi(y_0 | \lambda) d\tilde{y}_0.
\]
The desired result follows if we can show that

\[
\begin{align*}
(i) \quad & \limsup_{N \to \infty} \limsup_{\pi \in \Pi} \frac{N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] + N^{\epsilon_0}}{N \mathbb{E}_{\theta, \pi}^Y \left[ \left( \lambda_i - m(\hat{\lambda}_i, y_{i0}; \pi) \right)^2 \right] + N^{\epsilon_0}} \leq 1 \\
(ii) \quad & \limsup_{N \to \infty} \limsup_{\pi \in \Pi} \frac{N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] + N^{\epsilon_0}}{N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] + N^{\epsilon_0}} \leq 1.
\end{align*}
\]

**Part (i):** Notice that the denominator is bounded below by \( N^{\epsilon_0} \). We will proceed by constructing an upper bound for the numerator. Using the fact that the posterior mean \( \hat{\lambda}_i \) minimizes the integrated risk, we obtain

\[
\begin{align*}
& N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] \\
& \leq N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( m(\hat{\lambda}_i, y_{i0}; \pi) - \lambda_i \right)^2 \right] + N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( \lambda_i - m(\hat{\lambda}_i, y_{i0}; \pi) \right)^2 \right] \\
& \leq N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( m(\hat{\lambda}_i, y_{i0}; \pi) - \lambda_i \right)^2 \right] \mathbb{E}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{\beta} \mathcal{T}_m \mathcal{T}_\lambda) \\
& + N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( m(\hat{\lambda}_i, y_{i0}; \pi) - \lambda_i \right)^2 \right] (\mathbb{E}(\mathcal{T}_{\lambda Y_0}^c) + \mathbb{E}(\mathcal{T}_{\beta}^c \mathcal{T}_{\lambda Y_0}^c) + \mathbb{E}(\mathcal{T}_m^c) + \mathbb{E}(\mathcal{T}_\lambda^c)) \\
& = B_{1i} + B_{2i},
\end{align*}
\]

say.

A bound for \( B_{1i} \) can be obtained as follows:

\[
\begin{align*}
B_{1i} &= N \int \int \left( m(\hat{\lambda}_i, y_{i0}; \pi) - \lambda_i \right)^2 \mathbb{E}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{\beta} \mathcal{T}_m \mathcal{T}_\lambda \lambda_0) \pi_s(\hat{\lambda}_i | \lambda_i) \pi_s(y_{i0} | \lambda_i) \pi(\lambda_i) d(\hat{\lambda}_i, \lambda_i, y_{i0}) \\
& \leq (1 + o(1)) N \int \int \left( m(\hat{\lambda}_i, y_{i0}; \pi) - \lambda_i \right)^2 \mathbb{E}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{\beta} \mathcal{T}_m \mathcal{T}_\lambda \lambda_0) \pi(\hat{\lambda}_i | \lambda_i) \pi(y_{i0} | \lambda_i) \pi(\lambda_i) d(\hat{\lambda}_i, \lambda_i, y_{i0}) \\
& \leq (1 + o(1)) N \mathbb{E}_{\pi, \lambda, \theta}^Y \left[ \left( \lambda_i - m(\hat{\lambda}_i, y_{i0}; \pi) \right)^2 \right].
\end{align*}
\]

The first inequality is based on Assumption 3.4 and an argument similar to the one used in the analysis of term \( I \) in the proof of Lemma A.7. The \( o(1) \) term does not depend on \( \pi \in \Pi \).
To derive a bound for $B_{2i}$ first consider the inequalities

\[
\left( m(\hat{\lambda}_i, y_{i0}; \pi) - \lambda_i \right)^2 \leq 2m^2(\hat{\lambda}_i, y_{i0}; \pi) \left( \|T_m\| + \|T_c^c\| \right) + 2\lambda_i^2 \left( \|T_\lambda\| + \|T_\lambda^c\| \right)
\]

\[
\leq 4C_N^2 + 2m(\hat{\lambda}_i, y_{i0}; \pi)^2 \|T_m^c\| + 2\lambda_i^2 \|T_c^c\|.
\]

Thus,

\[
B_{2i} \leq 4C_N^2 \cdot N \left( \mathbb{P}(T_{\lambda\theta} T_{c(\cdot)}^c) + \mathbb{P}(T_{\lambda\theta}^c) + \mathbb{P}(T_m^c) + \mathbb{P}(T_\lambda^c) \right) + 8\mathbb{E}_{\theta, \pi}^{\lambda_i, \lambda_i} \left[ m(\hat{\lambda}_i, y_{i0}; \pi)^2 \|T_m^c\| \right] + 8\mathbb{E}_{\theta, \pi}^{\lambda_i, \lambda_i} \left[ \lambda_i^2 \|T_\lambda^c\| \right].
\]

Notice that $C_N^2 = o_u(N)$, $N \mathbb{P}(T_{\lambda\theta} T_{c(\cdot)}^c) = o_u(N^\epsilon_0)$ (see Step 1.1), $N \mathbb{P}(T_{\lambda\theta}^c) = o_u(N^+) \) (see Step 1.2), and $N \mathbb{P}(T_m^c) = o_u(N^+)$ (see Truncation 1). Also, notice that

\[
\mathbb{E}_{\theta, \pi}^{\lambda_i, \lambda_i} \left[ \lambda_i^2 \|T_\lambda^c\| \right] = N \int_{|\lambda| > C_N} \lambda^2 \pi(\lambda) d\lambda
\]

\[
\leq \sqrt{\int_{|\lambda| > C_N} \lambda^4 \pi(\lambda) d\lambda} \sqrt{N^2 \int_{|\lambda| > C_N} \pi(\lambda) d\lambda}
\]

\[
\leq M \sqrt{N^2 \int_{|\lambda| > C_N} \pi(\lambda) d\lambda}
\]

\[
= o_u(N^+).
\]

The first inequality is the Cauchy-Schwartz inequality, the second inequality holds by Assumption 3.2, and the last line follows from calculations similar to the ones in (A.6). Therefore,

\[
B_{2i} \leq o_u(N^\epsilon_0).
\]
Combining the bounds for $B_{1i}$ and $B_{2i}$, we have

\[
\limsup_{N \to \infty} \frac{\limsup_{\pi \in \Pi} N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right]}{N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( \lambda_i - m(\hat{\lambda}_i, y_{i0}; \pi) \right)^2 \right] + N^{\epsilon_0}}
\leq \limsup_{N \to \infty} \frac{(1 + o(1)) N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right]}{N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( \lambda_i - m(\hat{\lambda}_i, y_{i0}; \pi) \right)^2 \right] + N^{\epsilon_0}}
\leq \limsup_{N \to \infty} \frac{(1 + o(1)) \left[ N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] + N^{\epsilon_0} \right]}{N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( \lambda_i - m(\hat{\lambda}_i, y_{i0}; \pi) \right)^2 \right] + N^{\epsilon_0}}
= 1,
\]

where the term $o(1)$ holds uniformly in $\pi \in \Pi$. We have the required result for Part (i).

**Part (ii):** The proof of Part (ii) is similar to that of Part (i). We construct an upper bound for the numerator as follows

\[
N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] \\
\leq N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] + N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( \lambda_i - m(\hat{\lambda}_i, y_{i0}; \pi) \right)^2 \right] + N^{\epsilon_0}
\leq \left[ N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right] + N^{\epsilon_0} \right]
\]

\[
= B_{1i} + B_{2i},
\]
say. Now consider the term $B_{1i}$:

\[
B_{1i} = N \int \int \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 p_*(\hat{\lambda}_i|\lambda_i) p_*(y_{i0}|\lambda_i) \frac{p(\lambda_i|\lambda_i) p(y_{i0}|\lambda_i)}{p_*(\lambda_i|\lambda_i) p_*(y_{i0}|\lambda_i)} \pi(\lambda_i)
\times \left( \mathcal{T}_{\lambda Y_0} \mathcal{T}_{\rho(\cdot)} \mathcal{T}_{m*} \mathcal{T}_{\lambda} \right) d(\hat{\lambda}_i, \lambda_i, y_{i0})
\leq (1 + o(1)) N \int \int \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 p_*(\hat{\lambda}_i|\lambda_i) p_*(y_{i0}|\lambda_i) \pi(\lambda_i)
\times \left( \mathcal{T}_{\lambda Y_0} \mathcal{T}_{\rho(\cdot)} \mathcal{T}_{m*} \mathcal{T}_{\lambda} \right) d(\hat{\lambda}_i, d\lambda_i, d y_{i0})
\leq (1 + o(1)) N \mathbb{E}_{\theta, \pi}^{Y_i, \lambda_i} \left[ \left( m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \right)^2 \right],
\]
where the $o(1)$ term is uniform in $\pi \in \Pi$. Using the a similar argument as in Part (i) we can establish that $B_{2i} = o_{u.\pi}(N^0)$, which leads to the desired result.

### A.1.3.3 Term $A_{3i}$

**Lemma A.5** Suppose the assumptions in Theorem 3.7 hold. Then,

$$N\mathbb{E}_{\theta,\pi}^{Y^N_n}[(\hat{\rho} - \rho)^2Y_{iT}^2] = o_{u.\pi}(N^+).$$

**Proof of Lemma A.5.** Using the Cauchy-Schwarz inequality, we can bound

$$\mathbb{E}_{\theta,\pi}^{Y^N_n}[(\sqrt{N}(\hat{\rho} - \rho))^2Y_{iT}^2] \leq \sqrt{\mathbb{E}_{\theta,\pi}^{Y^N_n}[(\sqrt{N}(\hat{\rho} - \rho))^4]} \mathbb{E}_{\theta,\pi}^{Y^N_n}[Y_{iT}^4].$$

By Assumption 3.6, we have

$$\mathbb{E}_{\theta,\pi}^{Y^N_n}[(\sqrt{N}(\hat{\rho} - \rho))^4] \leq o_{u.\pi}(N^+).$$

For the second term, write

$$Y_{iT} = \rho^TY_{i0} + \sum_{\tau=0}^{T-1} \rho^\tau(\lambda_i + U_{iT-\tau}).$$

Using the $C_r$ inequality and noting that $T$ is finite and $U_{it} \sim iidN(0,\sigma^2)$, we deduce that there is a finite constant $M$ that does not depend on $\pi \in \Pi$ such that

$$\mathbb{E}_{\theta,\pi}^{Y^N_n}[Y_{iT}^4] \leq M \left( \mathbb{E}_{\theta}^{Y^N_n}[Y_{i0}^4] + \mathbb{E}_{\theta,\pi}^{Y^N_n}[\Lambda_1^4] + \mathbb{E}_{\theta,\pi}^{Y^N_n}[U_{i1}^4] \right) = o_{u.\pi}(N^+).$$

The last line holds according to Assumption 3.2 and because $U_{i1}$ is normally distributed and therefore all its moments are finite.
A.1.4 Further Details

We now provide more detailed derivations for some of the bounds used in Section A.1.3. Recall that

\[
R_1(i) = \frac{1}{N-1} \sum_{j \neq i}^{N} \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right) \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right)^2 \left( \bar{Y}_{j,-1} - \bar{Y}_{i,-1} \right) \frac{1}{B_N} \phi \left( \frac{Y_{j,0} - Y_{i,0}}{B_N} \right)
\]

\[
R_2(i) = \frac{1}{N-1} \sum_{j \neq i}^{N} \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right) \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} \right) \left( \bar{Y}_{j,-1} - \bar{Y}_{i,-1} \right) \frac{1}{B_N} \phi \left( \frac{Y_{j,0} - Y_{i,0}}{B_N} \right)
\]

For expositional purposes, our analysis focuses on the slightly simpler term \(R_2(i)\). The extension to \(R_1(i)\) is fairly straightforward. By definition,

\[
\hat{\lambda}_j(\bar{\rho}) - \hat{\lambda}_i(\bar{\rho}) = \hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho) - (\bar{\rho} - \rho)(\bar{Y}_{j,-1} - \bar{Y}_{i,-1}).
\]

Therefore,

\[
R_2(\bar{\rho}) = \frac{1}{N-1} \sum_{j \neq i}^{N} \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} - (\bar{\rho} - \rho) \left( \bar{Y}_{j,-1} - \bar{Y}_{i,-1} \right) \right)
\]

\[
\times \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} - (\bar{\rho} - \rho) \left( \bar{Y}_{j,-1} - \bar{Y}_{i,-1} \right) \right)
\]

\[
\times \left( \bar{Y}_{j,-1} - \bar{Y}_{i,-1} \right) \frac{1}{B_N} \phi \left( \frac{Y_{j,0} - Y_{i,0}}{B_N} \right).
\]

Consider the region \(T_{\bar{\rho}} \cap T_U \cap T_{\bar{Y}_0}\). First, using (A.12) we can bound

\[
\max_{1 \leq i, j \leq N} \left| (\bar{\rho} - \rho)(\bar{Y}_{j,-1} - \bar{Y}_{i,-1}) \right| \leq \frac{M}{L_N}.
\]
Thus,

\[
\phi\left(\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} - (\bar{\rho} - \rho) \left(\frac{\bar{Y}_{j,-1} - \bar{Y}_{i,-1}}{B_N}\right)\right) I(T \bar{U} \cap T_{Y_0}) \\
\leq \phi\left(\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} + \left(\frac{M}{L_N B_N}\right)\right) I\left\{\left|\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N}\right| \leq -\frac{M}{L_N B_N}\right\} \\
+ \phi(0) I\left\{\left|\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N}\right| \leq \frac{M}{L_N B_N}\right\} \\
+ \phi\left(\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} - \left(\frac{M}{L_N B_N}\right)\right) I\left\{\left|\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N}\right| \geq \frac{M}{L_N B_N}\right\} \\
= \tilde{\phi}\left(\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N}\right),
\]

say. The function \(\tilde{\phi}(x)\) is flat for \(|x| < \frac{M}{L_N B_N}\) and is proportional to a Gaussian density outside of this region.

Second, we can use the bound

\[
\left|\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N} - (\bar{\rho} - \rho) \left(\frac{Y_{j,-1} - Y_{i,-1}}{B_N}\right)\right| \leq \left|\frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{B_N}\right| + \frac{M}{L_N B_N}.
\]

Third, for the region \(T_{\bar{U}} \cap T_{Y_0}\) we can deduce from (A.11) that

\[
\max_{1 \leq i,j \leq N} |\bar{Y}_{j,-1} - \bar{Y}_{i,-1}| \leq M L_N.
\]

Therefore,

\[
\left|\frac{Y_{j,0} - Y_{i,0}}{B_N}\right| \leq \frac{M L_N}{B_N} \phi\left(\frac{Y_{j,0} - Y_{i,0}}{B_N}\right).
\]

Now, define the function

\[
\tilde{\phi}_*(x) = \tilde{\phi}(x) \left(\left|x\right| + \frac{M}{L_N B_N}\right).
\]

Because for random variables with bounded densities and Gaussian tails all moments exist and because \(L_N B_N > 1\) by definition of \(L_N\) in (A.5), the function \(\tilde{\phi}_*(x)\) has the property that for any finite positive integer \(m\) there is a finite constant \(M\) such that

\[
\int \tilde{\phi}_*(x)^m dx \leq M.
\]
Combining the previous results we obtain the following bound for \( R_{2i}(\bar{\rho}) \):

\[
| R_{2i}(\bar{\rho}) \| (\mathcal{T}_\rho \mathcal{T}_\theta \mathcal{T}_{Y_0}) \| \leq \frac{MLN}{N - 1} \sum_{j \neq i}^N \frac{1}{BN} \tilde{\phi}_* \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{BN} \right) \frac{1}{BN} \phi \left( \frac{Y_{j0} - Y_{i0}}{BN} \right). \quad (A.28)
\]

For the subsequent analysis it is convenient define the function

\[
f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) = \frac{1}{BN} \tilde{\phi}_* \left( \frac{\hat{\lambda}_j(\rho) - \hat{\lambda}_i(\rho)}{BN} \right) \phi \left( \frac{Y_{j0} - Y_{i0}}{BN} \right). \quad (A.29)
\]

In the remainder of this section we will state and prove three technical lemmas that establish moment bounds for \( R_{1i}(\bar{\rho}) \) and \( R_{2i}(\bar{\rho}) \). The bounds are used in Section A.1.3. We will abbreviate \( \mathbb{E}_{\theta,\pi,Y}^{\lambda(-i)}[\cdot] = \mathbb{E}_i[\cdot] \) and simply use \( \mathbb{E}[\cdot] \) to denote \( \mathbb{E}_{\theta,\pi}^{\lambda} [\cdot] \).

**Lemma A.6** Suppose the assumptions in Theorem 3.7 hold. Then, for any finite positive integer \( m \geq 1 \), over the regions \( \mathcal{T}_{\hat{\lambda}Y_0} \) and \( \mathcal{T}_{\rho(\cdot)} \), there exists a finite constant \( M \) that does not depend on \( \pi \) such that

\[
\mathbb{E}_i \left[ f^m(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) \right] \leq \frac{M}{B_{N}^{2(m-1)p_i}}.
\]

**Proof of Lemma A.6.** We have

\[
\mathbb{E}_i \left[ f^m(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) \right] = \frac{1}{B_{N}^{2(m-1)}} \int \frac{1}{BN} \tilde{\phi}_* \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{BN} \right)^m \frac{1}{BN} \phi \left( \frac{y_0 - Y_{i0}}{BN} \right)^m p(\hat{\lambda}, y_0) d(\hat{\lambda}, y_0) \\
= \frac{1}{B_{N}^{2(m-1)}} \int \left[ \int \frac{1}{BN} \tilde{\phi}_* \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{BN} \right)^m \frac{1}{BN} \phi \left( \frac{y_0 - Y_{i0}}{BN} \right)^m p(\hat{\lambda}, y_0|\lambda_i) d(\hat{\lambda}, y_0) \right] \pi(\lambda_i) d\lambda_i \\
= \frac{1}{B_{N}^{2(m-1)}} \int \left[ \int \frac{1}{BN} \tilde{\phi}_* \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{BN} \right)^m \frac{1}{BN} \phi \left( \frac{y_0 - Y_{i0}}{BN} \right)^m p(\hat{\lambda}, y_0|\lambda_i) d(\hat{\lambda}, y_0) \right] \pi(\lambda_i) d\lambda_i \\
+ \frac{1}{B_{N}^{2(m-1)}} \int \mathcal{T}_\lambda \left[ \int \frac{1}{BN} \tilde{\phi}_* \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{BN} \right)^m \frac{1}{BN} \phi \left( \frac{y_0 - Y_{i0}}{BN} \right)^m p(\hat{\lambda}, y_0|\lambda_i) d(\hat{\lambda}, y_0) \right] \pi(\lambda_i) d\lambda_i.
\]
The required result of the lemma follows if we show

\[
I = \frac{1}{p(\lambda_i, Y_{i0})} \int_{\mathcal{T}_\lambda} \left[ \int \frac{1}{B_N} \tilde{\phi_\star} \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{B_N} \phi \left( \frac{y_0 - Y_{i0}}{B_N} \right)^m p(\hat{\lambda}, y_0|\lambda_i) d(\hat{\lambda}, y_0) \right] \pi(\lambda_i) d\lambda_i \leq M
\]  

(A.30)

\[
II = \frac{1}{p(\lambda_i, Y_{i0})} \int_{\mathcal{T}_\lambda} \left[ \int \frac{1}{B_N} \tilde{\phi_\star} \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{B_N} \phi \left( \frac{y_0 - Y_{i0}}{B_N} \right)^m p(\hat{\lambda}, y_0|\lambda_i) d(\hat{\lambda}, y_0) \right] \pi(\lambda_i) d\lambda_i \leq M
\]  

(A.31)

over the regions \(\mathcal{T}_{\lambda_0}\) and \(\mathcal{T}_{p(\cdot)}\) and uniformly in \(\pi\).

For (A.30), notice that the inner integral of term \(I\) is

\[
\int \frac{1}{B_N} \tilde{\phi_\star} \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{B_N} \phi \left( \frac{y_0 - Y_{i0}}{B_N} \right)^m p(\hat{\lambda}, y_0|\lambda_i) d(\hat{\lambda}, y_0)
= \int \frac{1}{B_N} \tilde{\phi_\star} \left( \frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{\sigma/\sqrt{T}} \exp \left( -\frac{1}{2} \left( \frac{\hat{\lambda} - \lambda_i}{\sigma/\sqrt{T}} \right)^2 \right) d\hat{\lambda}
\times \int \frac{1}{B_N} \phi \left( \frac{y_0 - Y_{i0}}{B_N} \right)^m \pi(y_0|\lambda) dy_0
= I_1 \times I_2,
\]

say.
Notice that

\[ I_1 = \int \frac{1}{B_N} \phi_s \left( \frac{\lambda - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{\sigma/\sqrt{T}} \exp \left( -\frac{1}{2} \left( \frac{\lambda - \lambda_i}{\sigma/\sqrt{T}} \right)^2 \right) d\lambda \]

\[ = \int \phi_s (\lambda^*)^m \frac{1}{\sigma/\sqrt{T}} \exp \left( -\frac{1}{2} \left( \frac{\lambda_i - \lambda + B_N \lambda^*}{\sigma/\sqrt{T}} \right)^2 \right) d\lambda^* \]

\[ = \int \phi_s (\lambda^*)^m \exp \left( -\left( (\hat{\lambda}_i - \lambda_i) B_N \lambda^* / \sigma^2 / T \right) \right) \exp \left( -\frac{1}{2} \left( \frac{B_N \lambda^*}{\sigma/\sqrt{T}} \right)^2 \right) d\lambda^* \]

\[ \times \left[ \frac{1}{\sigma/\sqrt{T}} \exp \left( -\frac{1}{2} \left( \frac{\lambda_i - \lambda_i}{\sigma/\sqrt{T}} \right)^2 \right) \right] \]

\[ \leq \int \tilde{\phi}_s (\lambda^*)^m \exp \left( -\left( (\hat{\lambda}_i - \lambda_i) B_N \lambda^* / \sigma^2 / T \right) \lambda^* \right) d\lambda^* \]

\[ \times \left[ \frac{1}{\sigma/\sqrt{T}} \exp \left( -\frac{1}{2} \left( \frac{\lambda_i - \lambda_i}{\sigma/\sqrt{T}} \right)^2 \right) \right] \]

\[ \leq M \left( \int_0^\infty \tilde{\phi}_s (\lambda^*)^m \exp (v_N \lambda^*) d\lambda^* \right) \left[ \frac{1}{\sigma/\sqrt{T}} \exp \left( -\frac{1}{2} \left( \frac{\lambda_i - \lambda_i}{\sigma/\sqrt{T}} \right)^2 \right) \right] \]

\[ \leq M \left[ \frac{1}{\sigma/\sqrt{T}} \exp \left( -\frac{1}{2} \left( \frac{\lambda_i - \lambda_i}{\sigma/\sqrt{T}} \right)^2 \right) \right] \]

\[ = M p(\hat{\lambda}_i | \lambda_i), \]

where \( v_N = \frac{T}{\sigma^2} (C'_N + C_N) B_N. \)

Here, for the second equality, we used the change-of-variable \( \lambda_* = (\hat{\lambda} - \hat{\lambda}_i) / B_N \) to replace \( \hat{\lambda} \). The first inequality holds because the exponential function \( \exp \left( -\frac{1}{2} \left( \frac{B_N \lambda^*}{\sigma/\sqrt{T}} \right)^2 \right) \) is bounded by one. Moreover, under truncations \( T_{\lambda Y} \) and \( T_{\lambda}, |\hat{\lambda}_i| \leq C'_N \) and \( |\lambda_i| \leq C_N. \) According to Assumption 3.3, \( v_N = \frac{T}{\sigma^2} (C'_N + C_N) B_N = o(1). \) Thus, the last inequality holds because \( \int_0^\infty \tilde{\phi}_s (\lambda^*)^m \exp (v_N \lambda^*) d\lambda^* \) is finite.

We now proceed with a bound for the second integral, \( I_2. \) Using the fact that the Gaussian
pdf $\phi(x)$ is bounded and by Assumption 3.4, we can write

$$I_2 = \int \frac{1}{B_N} \phi \left( \frac{y_0 - Y_{i0}}{B_N} \right) \pi(y_0 | \lambda) dy_0$$

$$\leq M \int \frac{1}{B_N} \phi \left( \frac{y_0 - Y_{i0}}{B_N} \right) \pi(y_0 | \lambda) dy_0$$

$$= M(1 + o(1)) \pi(Y_{i0} | \lambda),$$

uniformly in $|y_0| \leq C_N'$ and $|\lambda| \leq C_N$ and in $\pi \in \Pi$.

Combining the bounds for $I_1$ and $I_2$ and integrating over $\lambda$, we obtain

$$I \leq M \frac{1}{p(\hat{\lambda}_i, Y_{i0})} \int_{T_\lambda} \left[ p(\hat{\lambda}_i | \lambda_i) \pi(Y_{i0} | \lambda_i) \right] \pi(\lambda_i) d\lambda_i$$

$$\leq M \frac{1}{p(\hat{\lambda}_i, Y_{i0})} \int_{T_\lambda} p(\hat{\lambda}_i | \lambda_i) \pi(Y_{i0} | \lambda_i) \pi(\lambda_i) d\lambda_i = M,$$

as required for (A.30).

Next, for (A.31), since $\tilde{\phi}_*(x), \phi(x), p(\hat{\lambda}, y_0 | \lambda_i)$ are bounded uniformly in $\pi$ and $p(\hat{\lambda}_i, Y_{i0}) > N^{-1}$ over $T_{p(*)}$, we have

$$II \leq M \frac{1}{p(\hat{\lambda}_i, Y_{i0}) B_N^2} \int_{T_\lambda} \pi(\lambda_i) d\lambda_i$$

$$\leq MN^{-\epsilon} \left( \frac{1}{B_N^2} \right) \left( N \int_{T_\lambda} \pi(\lambda_i) d\lambda_i \right)$$

$$\leq MN^{-\epsilon} o_u.\pi(N^+) o_u.\pi(N^+)$$

$$\leq M,$$

where the second-to-last line holds because according to Assumption 3.3 $1/B_N^2 = o_u.\pi(N^+)$ and because of the tail bound in (A.6). This yields the required result for (A.31). ■

**Lemma A.7** Suppose the assumptions required for Theorem 3.7 are satisfied. Then,

(a) $\sup_{\pi \in \Pi} \sup_{(\hat{\lambda}_i, Y_{i0}) \in T_{X_0} \cap T_{p(*)}} \left| \frac{p_{\pi i}}{p_i} - 1 \right| = o(1)$,

(b) $\sup_{\pi \in \Pi} \sup_{(\hat{\lambda}_i, Y_{i0}) \in T_{X_0} \cap T_{p(*)}} \left| \frac{p_i}{p_{\pi i}} - 1 \right| = o(1)$. 
Proof of Lemma A.7.

Part (a). Denote

\[ p(\hat{\lambda}_i, y_{i0}|\lambda_i) = \frac{1}{\sqrt{\sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}} \right) \pi(y_{i0}|\lambda_i) \]

\[ p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) = \frac{1}{\sqrt{B_N^2 + \sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{B_N^2 + \sigma^2/T}} \right) \left[ \int \frac{1}{B_N} \phi \left( \frac{y_{i0} - \tilde{y}_{i0}}{B_N} \right) \pi(\tilde{y}_{i0}|\lambda_i) d\tilde{y}_{i0} \right], \]

so that

\[ p_i = \int p(\hat{\lambda}_i, y_{i0}|\lambda_i) \pi(\lambda_i) d\lambda_i, \quad p_{si} = \int p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) \pi(\lambda_i) d\lambda_i. \]

Notice that

\[ \left| \frac{p_{si} - p_i}{p_i} \right| = \left| \frac{p_{si} - p_i}{p_i} \right| \leq \frac{1}{p_i} \int \left| p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i) \right| \pi(\lambda_i) d\lambda_i 
\]

\[ = \frac{1}{p_i} \int_{\mathcal{T}_\lambda} \left| p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i) \right| \pi(\lambda_i) d\lambda_i 
\]

\[ + \frac{1}{p_i} \int_{\mathcal{T}_\lambda} \left| p(\hat{\lambda}_i, y_{i0}|\lambda_i) - p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) \right| \pi(\lambda_i) d\lambda_i 
\]

\[ = I + II, \text{ say.} \]

For term I, since \(|\lambda_i| \leq C_N\) in the region \(\mathcal{T}_\lambda\) and \(|\hat{\lambda}_i| \leq C'_N\) in the region \(\mathcal{T}_{\hat{\lambda}_0}\), we can choose a constant \(M\) that does not depend on \(\pi\) such that for \(N\) sufficiently large

\[ \frac{1}{\sqrt{B_N^2 + \sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{B_N^2 + \sigma^2/T}} \right) = \frac{1}{\sqrt{\sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}} \right) \]

\[ \times \frac{\sqrt{\sigma^2/T}}{\sqrt{B_N^2 + \sigma^2/T}} \exp \left\{ \frac{1}{2} \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{B_N^2 + \sigma^2/T}} \right)^2 \frac{B_N^2}{\sigma^2/T} \right\} \]

\[ \leq \frac{1}{\sqrt{\sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}} \right) \exp(M(C'_N + C_N)^2 B_N^2) \]

\[ = \frac{1}{\sqrt{\sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}} \right) (1 + o(1)), \]

where the inequality holds by Assumption 3.3 which implies that \((C'_N + C_N)B_N = o(1)\), and the \(o(1)\) term in the last line is uniform in \((\hat{\lambda}_i, Y_{i0}) \in \mathcal{T}_{\hat{\lambda}_0} \cap \mathcal{T}_p\) and in \(\pi \in \Pi\).
According to Assumption 3.4,

\[
\int \frac{1}{B_N} \phi \left( \frac{y_{i0} - \bar{y}_{i0}}{B_N} \right) \pi(\bar{y}_{i0}|\lambda_i) d\bar{y}_{i0} = (1 + o(1)) \pi(y_{i0}|\lambda_i)
\]

uniformly in \( |y_{i0}| \leq C'_N \) and \( |\lambda_i| \leq C_N \) and in \( \pi \in \Pi \).

Then,

\[
I = \frac{1}{p_i} \int_{T_\lambda} \left| p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i) \right| \pi(\lambda_i) d\lambda_i
\]

\[
= \frac{1}{p_i} \int_{T_\lambda} \left| \frac{1}{\sqrt{B_N^2 + \sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{B_N^2 + \sigma^2/T}} \right) \right| \frac{1}{B_N} \phi \left( \frac{y_{i0} - \bar{y}_{i0}}{B_N} \right) \pi(\bar{y}_{i0}|\lambda_i) d\bar{y}_{i0}
\]

\[
- \frac{1}{\sqrt{\sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}} \right) \pi(y_{i0}|\lambda_i) \right| \pi(\lambda_i) d\lambda_i
\]

\[
\leq |(1 + o(1))^2 - 1| \frac{1}{p_i} \int_{T_\lambda} \frac{1}{\sqrt{\sigma^2/T}} \phi \left( \frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}} \right) \pi(y_{i0}|\lambda_i) \pi(\lambda_i) d\lambda_i
\]

\[
\leq |(1 + o(1))^2 - 1| = o(1).
\]

Note that the \( o(1) \) term does not depend on \((\hat{\lambda}_i, Y_{i0}) \in T_{\lambda Y_{i0}} \cap T_{p(i)}\) nor on \( \pi \in \Pi \).

For term \( II \), calculations similar to the one in (A.27) imply that the densities \( p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) \) and \( p(\hat{\lambda}_i, y_{i0}|\lambda_i) \) are bounded, say, by \( M \). Thus, we have

\[
II = \frac{1}{p_i} \int_{T_\lambda} \left| p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i) \right| \pi(\lambda_i) d\lambda_i
\]

\[
\leq \frac{2M}{p_i} \int_{T_\lambda} \pi(\lambda_i) d\lambda_i
\]

\[
\leq 2M \sup_{\pi \in \Pi} N^{1-\epsilon} \int_{T_\lambda} \pi(\lambda_i) d\lambda_i
\]

\[
= o(1),
\]

where the second inequality holds since \( p_i > \frac{N^\epsilon}{N} \) under the truncation \( T_{p(i)} \) and the last line holds according to (A.6).

Combining the upper bounds of \( I \) and \( II \) yields the required result for Part (a).

Part (b). According to Part (a),

\[
p_{si} = p_i(1 + o(1))
\]
uniformly in $(\hat{\lambda_i}, Y_{i0}) \in \mathcal{T}_{\lambda Y_0} \cap \mathcal{T}_p(\cdot)$ and in $\pi \in \Pi$. Then, for some finite constant $M$ that does not depend on $(\hat{\lambda_i}, Y_{i0})$ and $\pi$,

$$\sup_{\pi \in \Pi} \sup_{(\hat{\lambda_i}, Y_{i0}) \in \mathcal{T}_{\lambda Y_0} \cap \mathcal{T}_p(\cdot)} \frac{|p_i - p_{si}|}{p_{si}} = \sup_{\pi \in \Pi} \sup_{(\hat{\lambda_i}, Y_{i0}) \in \mathcal{T}_{\lambda Y_0} \cap \mathcal{T}_p(\cdot)} \frac{|p_i - p_{si}|}{p_{si}} \frac{p_i}{p_{si}}$$

$$= \sup_{\pi \in \Pi} \sup_{(\hat{\lambda_i}, Y_{i0}) \in \mathcal{T}_{\lambda Y_0} \cap \mathcal{T}_p(\cdot)} \frac{|p_i - p_{si}|}{p_{si}} \frac{p_i}{p_{si}} (1 + o(1))$$

$$\leq M \sup_{\pi \in \Pi} \sup_{(\hat{\lambda_i}, Y_{i0}) \in \mathcal{T}_{\lambda Y_0} \cap \mathcal{T}_p(\cdot)} \frac{|p_i - p_{si}|}{p_{si}}$$

$$= o(1),$$

as required for Part (b). $\blacksquare$

**Lemma A.8** Under the assumptions required for Theorem 3.7, we obtain the following bounds:

(a) $\mathbb{E}_i \left[ R_{2i}^i(\tilde{\rho}) \mathbb{I}(\mathcal{T}_p \mathcal{T}_U \mathcal{T}_Y \mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)} \mathcal{T}_{\tilde{p}(\cdot)}) \right] \leq M L_N^{1/4} p_i \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_p(\cdot))$,

(b) $\mathbb{E}_i \left[ R_{2i}^i(\tilde{\rho}) \mathbb{I}(\mathcal{T}_p \mathcal{T}_U \mathcal{T}_Y \mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)} \mathcal{T}_{\tilde{p}(\cdot)}) \right] \leq M \frac{L_N^{1/4}}{B_N} p_i \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_p(\cdot))$,

(c) $\mathbb{E}_i \left[ N(\tilde{p}^{(-i)}_i) - p_{si}) \mathbb{I}(\mathcal{T}_p \mathcal{T}_U \mathcal{T}_Y \mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)} \mathcal{T}_{\tilde{p}(\cdot)}) \right] \leq M \frac{B_N}{p_i} \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_p(\cdot))$,

(d) $\mathbb{E}_i \left[ N(d\tilde{p}^{(-i)}_i - dp_{si}) \mathbb{I}(\mathcal{T}_p \mathcal{T}_U \mathcal{T}_Y \mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)} \mathcal{T}_{\tilde{p}(\cdot)}) \right] \leq M \frac{B_N}{p_i} \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_p(\cdot))$, where $M$ is a finite constant that does not depend on $\pi \in \Pi$.

(e) For any finite $m > 1$, $\int_{\mathcal{T}_{\lambda Y_0} \cap \mathcal{T}_p(\cdot)} \left( \frac{p_i}{p_{si}} \right)^m d\lambda_i d\gamma_{i0} = o_{\pi}(N^+)$.

(f) $\mathbb{N}\mathbb{E} \mathbb{I}\left\{ \tilde{p}^{(-i)}_i - p_{si} < -p_{si}/4 \right\} \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_p(\cdot)) = o_{\pi}(N^+)$.  

**Proof of Lemma A.8.** Part (a). Recall the following definitions

$$\tilde{\phi}(x) = \phi \left( x + \frac{M}{L_N B_N} \right) \mathbb{I} \left\{ x \leq -\frac{M}{L_N B_N} \right\} + \phi(0) \mathbb{I} \left\{ x \geq \frac{M}{L_N B_N} \right\}$$

$$\phi_*(x) = \phi(0) \left( x + \frac{M}{L_N B_N} \right).$$
First, recall that according to (A.28) and (A.29), in the region $T_\tilde{\rho} \cap T_\theta \cap T_{y_0}$

$$|R_{2i}(\tilde{\rho})| \leq \frac{MLN}{N-1} \sum_{j \neq i}^N f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}).$$

Then,

$$|R_{2i}(\tilde{\rho})|^4 \leq \left[ \frac{MLN}{N-1} \sum_{j \neq i}^N f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) \right]^4$$

$$= \left[ \frac{MLN}{N-1} \sum_{j \neq i}^N \left( f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) - \mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})] \right) \right]^4$$

$$\leq MLN^4 \left[ \frac{1}{N-1} \sum_{j \neq i}^N \left( f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) - \mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})] \right) \right]^4$$

$$+ MLN^4 \left[ \mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})] \right]^4$$

$$= MLN^4 (A_1 + A_2),$$

say. The first equality holds since $f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})$ are iid conditional on $(\hat{\lambda}_i, Y_{i0})$. The second inequality holds because $|x + y|^4 \leq 8(|x|^4 + |y|^4)$.

The term $(N - 1)^4 A_1$ takes the form

$$\left( \sum a_j \right)^4 = \left( \sum a_j^2 + 2 \sum_j \sum_{i > j} a_j a_i \right)^2$$

$$= \left( \sum a_j^2 \right)^2 + 4 \left( \sum a_j^2 \right) \left( \sum_j \sum_{i > j} a_j a_i \right) + 4 \left( \sum \sum_{i > j} a_j a_i \right)^2$$

$$= \sum a_j^4 + 6 \sum_j \sum_{i > j} a_j^2 a_i^2$$

$$+ 4 \left( \sum a_j^2 \right) \left( \sum_j \sum_{i > j} a_j a_i \right) + 4 \sum_j \sum_{i > j} \sum_{l \neq j} \sum_{k > l} a_j a_l a_k,$$

where

$$a_j = f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) - \mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})], \quad j \neq i.$$
across $j \neq i$. This implies that

$$
\mathbb{E}_i \left[ \left( \sum a_j \right)^4 \right] = \sum \mathbb{E}_i [a_i^4] + 6 \sum_{j>i} \mathbb{E}_i [a_j^2 a_i^2].
$$

The remaining terms drop out because they involve at least one term $a_j$ that is raised to the power of one and therefore has mean zero.

Using the $C_r$ inequality, Jensen’s inequality, the conditional independence of $a_j^2$ and $a_i^2$ and Lemma A.6, we can bound

$$
\mathbb{E}_i [a_j^4] \leq \frac{M}{B_N^6} p_i, \quad \mathbb{E}_i [a_j^2 a_i^2] \leq \frac{M}{B_N^4} p_i^2.
$$

Thus, in the region $T_{\bar{p}} \cap T_C \cap T_{Y_0} \cap T_{\lambda Y_0} \cap T_{p(\cdot)} \cap T_{\bar{p}(\cdot)},$

$$
\mathbb{E}_i [A_1] \leq \frac{M p_i}{N^3 B_N^6} + \frac{M p_i^2}{N^2 B_N^4} \leq M p_i^4,
$$

The second inequality holds because over $T_{\bar{p}(\cdot)}$, $p_i \geq \frac{N^4}{N} \geq \frac{M}{N B_N^2}$ and for $N$ large, $N^3 B_N^6$ and $N^2 B_N^4$ are larger than one under Assumption 3.3. Here $M$ is uniform in $\pi \in \Pi$.

Using a similar argument, we can also deduce that

$$
\mathbb{E}_i [A_2] \leq M p_i^4,
$$

which proves Part (a) of the lemma.

**Part (b).** Similar to proof of Part (a).

**Part (c).** Can be established using existing results for the variance of a kernel density estimator.

**Part (d).** Similar to proof of Part (c).

**Part (e).** We have the desired result because by Lemma A.7 we can choose a constant $c$ that does not depend on $\pi$ such that

$$
p_i - p_{*i} \leq c p_{*i}
$$
over the region $T_{\hat{\lambda}Y_0} \cap T_{p(\cdot)}$. Thus,

$$
\left( \frac{p_i}{p_{si}} \right)^m = \left( 1 + \frac{p_i - p_{si}}{p_{si}} \right)^m \leq (1 + c)^m.
$$

We deduce that

$$
\int_{T_{\hat{\lambda}Y_0} \cap T_{p(\cdot)}} \left( \frac{p_i}{p_{si}} \right)^m d\hat{\lambda}_i dy_{i0} \leq (1 + c)^m \int_{T_{\hat{\lambda}Y_0} \cap T_{p(\cdot)}} d\hat{\lambda}_i dy_{i0} \leq (1 + c)^m (2C_N')^2 = o_{u.p}(N^+),
$$
as required.

**Part (f).** Define

$$
\psi_i(\hat{\lambda}_j, Y_{j0}) = \phi \left( \frac{\hat{\lambda}_j - \hat{\lambda}_i}{B_N} \right) \phi \left( \frac{Y_{j0} - Y_{i0}}{B_N} \right)
$$

and write

$$
p_i^{(-i)} - p_{si} = \frac{1}{N - 1} \sum_{j \neq i} \left\{ \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_j - \hat{\lambda}_i}{B_N} \right) \frac{1}{B_N} \phi \left( \frac{Y_{j0} - Y_{i0}}{B_N} \right) \right. \\
- \mathbb{E}_i \left[ \frac{1}{B_N} \phi \left( \frac{\hat{\lambda}_j - \hat{\lambda}_i}{B_N} \right) \frac{1}{B_N} \phi \left( \frac{Y_{j0} - Y_{i0}}{B_N} \right) \right] \left\} \\
= \frac{1}{B_N^2(N - 1)} \sum_{j \neq i} \left( \psi_i(\hat{\lambda}_j, Y_{j0}) - \mathbb{E}_i[\psi_i(\hat{\lambda}_j, Y_{j0})] \right).
$$

Notice that conditional on $(\hat{\lambda}_i, Y_{i0})$, $\psi_i(\hat{\lambda}_j, Y_{j0}) \sim iid$ across $j \neq i$ with $|\psi_i(\hat{\lambda}_j, Y_{j0})| \leq M$ for some finite constant $M$. Then, by Bernstein’s inequality \(^{15}\) (e.g., Lemma 19.32 in van der Vaart (1998)),

$$
N\mathbb{P} \left\{ \left( \frac{1}{N - 1} \right)^{N} \sum_{j \neq i} \left( \psi_i(\hat{\lambda}_j, Y_{j0}) - \mathbb{E}_i[\psi_i(\hat{\lambda}_j, Y_{j0})] \right) < -\frac{p_{si}}{4} \right\} \leq 2N \exp \left( -\frac{1}{4E[\psi_i(\hat{\lambda}_j, Y_{j0})] + MB_N^2p_{si}/4} \right) \mathbb{P} \left( (T_{\hat{\lambda}Y_0} \cap T_{p(\cdot)}) \right).
$$

\(^{15}\)For a bounded function $f$ and a sequence of iid random variables $X_i$,

$$
\mathbb{P} \left\{ \left| \frac{1}{\sqrt{N}} \sum_{i=1}^N (f(X_i) - \mathbb{E}[f(X_i)]) \right| > x \right\} \leq 2 \exp \left( -\frac{1}{4 \mathbb{E}[f(X_i)]^2 + \frac{x^2}{\sqrt{N}} \sup_x |f(x)|} \right).
$$
Using an argument similar to the proof of Lemma A.6 one can show that

\[ \mathbb{E}[\psi_i(\lambda_j, Y_{j0})^2 / B_N^4] \leq Mp_i / B_N^2. \]

In turn

\[ N \mathbb{P}_i \left\{ \frac{\hat{p}_i^{(-i)}}{p_i} - p_{si} < -\frac{p_{si}}{4} \right\} \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)}) \leq 2 \exp \left( -MB_N^2 \frac{p_{si}^2}{p_i + p_{si}} + \ln N \right) \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)}). \]

From Lemma A.7 we can find a constant \( c \) such that \( p_i \leq (1 + c)p_{si} \) and \( p_{si} \leq (1 + c)p_i \). This leads to

\[ \frac{p_{si}^2}{p_i + p_{si}} \geq \frac{p_i}{(2 + c)(1 + c)^2}. \]

Then, on the region \( T_{p(\cdot)} \)

\[
N \mathbb{E} \left[ \mathbb{P}_i \left\{ \frac{\hat{p}_i^{(-i)}}{p_i} - p_{si} < -\frac{p_{si}}{4} \right\} \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)}) \right] \\
\leq 2 \mathbb{E} \left[ \exp \left( -MB_N^2 \frac{p_{si}^2}{p_i + p_{si}} + \ln N \right) \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)}) \right] \\
\leq 2 \mathbb{E} \left[ \exp \left( -MB_N^2 p_i + \ln N \right) \mathbb{I}(\mathcal{T}_{\lambda Y_0} \mathcal{T}_{p(\cdot)}) \right] \\
\leq 2 \exp \left( -MB_N^2 N^c + \ln N \right) \\
= o(1),
\]

where the last line holds by Assumption 3.3 and the \( o(1) \) bound in the last line is uniform in \( \pi \in \Pi \). Then, we have the required result for Part (f). ■

A.2 Proofs for Section 3.3

Proof of Theorem 3.8.

Part (i): we verify that our assumptions hold uniformly for the multivariate normal distributions \( \pi \in \Pi \).

Assumption 3.2. Because \( \lambda \) is normally distributed, the uncentered fourth moment is finite for each \( \pi(\lambda) \in \Pi_\lambda \) and can be bounded uniformly. Note that

\[
\mathbb{P}(|\lambda| \geq C) \leq \mathbb{P}(|\lambda - \mu_\lambda| \geq C - |\mu_\lambda|) \leq 2 \exp \left( -\frac{C - |\mu_\lambda|}{2\sigma^2_\lambda} \right) \quad (A.32)
\]
for $C > |\mu_\lambda| + 1$. By plugging the bounds from (22) into (A.32) one can obtain constants $M_1$, $M_2$, and $M_3$ such that Assumption 3.2(i) is satisfied. The second part can be verified by noting that $Y_0 \sim N(\mu_y, \sigma_y^2)$, where $\mu_y = \alpha_0 + \alpha_1 \mu_\lambda$ and $\sigma_y^2 = \sigma_{y\mid \lambda}^2 + \alpha_1^2 \sigma_\lambda^2$.

**Assumption 3.4** The boundedness of the conditional density follows from $0 < \delta_{\sigma_{y\mid \lambda}} \leq \sigma_{y\mid \lambda}$ in (22). To verify Part (ii) define $\mu_y(\lambda) = \alpha_0 + \alpha_1 \lambda$ and notice that $\tilde{y} \mid \lambda \sim N(\mu_y(\lambda), \sigma_{y\mid \lambda}^2 + B_N^2)$. Thus, we can write

$$
\sup_{\pi \in \Pi} \sup_{|y| \leq C'_N / \lambda < C_N} \left| \frac{\int \phi \left( \frac{\tilde{y} - y}{B_N} \right) \pi(\tilde{y} \mid \lambda) d\tilde{y}}{\pi(y \mid \lambda)} - 1 \right| = \sup_{\pi \in \Pi} \sup_{|y| \leq C'_N / \lambda < C_N} |R_{1,N} \cdot R_{2,N} - 1|,
$$

where

$$
R_{1,N} = \sqrt{\frac{\sigma_{y\mid \lambda}^2}{\sigma_{y\mid \lambda}^2 + B_N^2}} \leq 1, \quad R_{2,N} = \exp \left\{ -\frac{1}{2} (y - \mu_y(\lambda))^2 \left( \frac{1}{\sigma_{y\mid \lambda}} + \frac{1}{\sigma_{y\mid \lambda}^2 + B_N^2} \right) \right\} \geq 1.
$$

$R_{1,N}$ can be bounded from below by replacing $\sigma_{y\mid \lambda}$ with $\delta_{\sigma_{y\mid \lambda}}$. Because $B_N \rightarrow 0$ as $N \rightarrow \infty$, $R_{1,N} \rightarrow 1$ uniformly. For the term $R_{2,N}$ notice that

$$
(y - \mu_y(\lambda))^2 \left( \frac{1}{\sigma_{y\mid \lambda}^2} - \frac{1}{\sigma_{y\mid \lambda}^2 + B_N^2} \right) = (y - \alpha_0 - \alpha_1 \lambda)^2 \frac{B_N^2}{\sigma_{y\mid \lambda}^2 \left( \sigma_{y\mid \lambda}^2 + B_N^2 \right)} \\
\leq 3 \left( (C'_N)^2 + M_{\alpha_0}^2 + M_{\alpha_1}^2 C_N^2 \right) \frac{B_N^2}{(\delta_{\sigma_{y\mid \lambda}})^2} \rightarrow 0
$$

as $N \rightarrow \infty$ because $B_N C_N = o(1)$ and $B_N C'_N = o(1)$ according to Assumption 3.3. Thus, $R_{2,N} \rightarrow 1$ uniformly which delivers the desired result.

**Assumption 3.5.** The first step is to derive the conditional prior distribution $\pi(\lambda \mid y)$ which is of the form $\lambda \mid y \sim N(\mu_{\lambda \mid y}, \sigma_{y\mid \lambda}^2)$. The prior mean function is of the form $\mu_{\lambda \mid y} = \gamma_0 + \gamma_1 y$. If the prior for $\lambda$ is a point mass, i.e., $\sigma_\lambda^2 = 0$, then the distribution of $(\lambda \mid y)$ is also a point mass with $\mu_{\lambda \mid y} = \mu_{\lambda}$ and $\sigma_{\lambda \mid y}^2 = 0$. It can be verified that the coefficients $\gamma_0$, $\gamma_1$, and the variance $\sigma_{y\mid \lambda}^2$ are bounded from above in absolute value.

The prior is combined with the Gaussian likelihood function $\hat{\lambda} \mid \lambda \sim N(\lambda, \sigma^2 / T)$, which leads to a posterior mean function that is linear in $\hat{\lambda}$ and $y$:

$$
m(\hat{\lambda}, y; \pi) = \left( \frac{1}{\sigma_{y\mid \lambda}} + \frac{1}{\sigma^2 / T} \right)^{-1} \left( \frac{1}{\sigma_{y\mid \lambda}^2} (\gamma_0 + \gamma_1 y_0) + \frac{1}{\sigma^2 / T} \hat{\lambda} \right) = \bar{\gamma}_0 + \bar{\gamma}_1 y_0 + \bar{\gamma}_2 \hat{\lambda}. \quad (A.33)
$$
The \( \gamma \) coefficients are also bounded in absolute value for \( \pi \in \Pi \).

The sampling distribution of \((\hat{\lambda}, y_0)\) is jointly normal with mean and covariance matrix

\[
\begin{bmatrix}
\mu_{\hat{\lambda}, y} \\
\Sigma_{\hat{\lambda}, y}
\end{bmatrix} =
\begin{bmatrix}
\mu_{\lambda} \\
\alpha_0 + \alpha_1 \mu_{\lambda}
\end{bmatrix},
\begin{bmatrix}
\sigma^2_{\lambda} + \sigma^2/T \\
\gamma_1\left(\frac{\sigma^2_{y|\lambda} + \alpha_1^2 \sigma^2_{\lambda}}{\sigma^2_{y|\lambda} + \alpha_1^2 \sigma^2_{\lambda}}\right)
\end{bmatrix}
\]

(A.34)

It can be verified that the covariance matrix is always positive definite. The variances of \( \hat{\lambda} \) and \( y_0 \) are strictly greater than some \( \delta > 0 \) and the two random variables are never perfectly correlated because \( \hat{\lambda} = \lambda + \frac{\sum_{t=1}^T u_t}{T} \). Moreover, the covariance matrix can be bounded from above. By combining (A.33) and (A.34) we can deduce that the posterior mean has a Gaussian sampling distribution and we can use standard moment and tail bounds to establish the validity of the assumption. Calculations under the \( p_* (\cdot) \) distributions are very similar.

**Part (ii):** we verify that our assumptions hold uniformly for the finite mixtures of multivariate normal distributions \( \pi_{\text{mix}} \in \Pi_{\text{mix}}^{(K)} \).

**Assumption 3.2.** Consider the marginal density of \( \lambda \) given by \( \pi_{\text{mix}}(\lambda) = \sum_{k=1}^K \omega_k \pi_k(\lambda) \).

Thus, for any non-negative integrable function \( h(\cdot) \) we can use the crude bound

\[
\int h(\lambda) \pi_{\text{mix}}(\lambda) d\lambda \leq \sum_{k=1}^K \int h(\lambda) \pi_k(\lambda) d\lambda.
\]

In turn, uniform tail probability and moment bounds for the mixture components translate into uniform bounds for \( \pi_{\text{mix}}(\cdot) \).

**Assumption 3.4** The key insight is that we can express

\[
\pi_{\text{mix}}(y|\lambda) = \frac{\sum_{k=1}^K \omega_k \pi_k(\lambda, y)}{\sum_{k=1}^K \omega_k \pi_k(\lambda)} = \sum_{k=1}^K \left( \frac{\omega_k \pi_k(\lambda)}{\sum_{k=1}^K \omega_k \pi_k(\lambda)} \right) \frac{\pi_k(\lambda, y)}{\pi_k(\lambda)} \leq \sum_{k=1}^K \pi_k(y|\lambda).
\]

This allows us to directly translate bounds for the mixture components \( \pi_k(y|\lambda) \in \Pi_{y|\lambda} \) into results for \( \pi_{\text{mix}}(y|\lambda) \). Using a similar argument, we can also deduce that

\[
\sup_{\pi_{\text{mix}} \in \Pi_{\text{mix}}^{(K)}} \sup_{\frac{B_N}{N} |y| \leq C_N, |\lambda| < C_N} \left| \frac{\int \frac{1}{B_N} \phi \left( \frac{\bar{y} - y}{B_N} \right) [\pi_{\text{mix}}(\bar{y}|\lambda) - \pi_{\text{mix}}(y|\lambda)] d\bar{y}}{\pi_{\text{mix}}(y|\lambda)} \right| \leq \sum_{k=1}^K \sup_{\pi_k \in \Pi_k} \sup_{\frac{B_N}{N} |y| \leq C_N, |\lambda| < C_N} \left| \frac{\int \frac{1}{B_N} \phi \left( \frac{\bar{y} - y}{B_N} \right) [\pi_k(\bar{y}|\lambda) - \pi_k(y|\lambda)] d\bar{y}}{\pi_k(y|\lambda)} \right| = o(1).
\]
Assumption 3.5. The prior distribution of $\lambda$ given $y$ is a mixture of normals with weights that are a function of $y$:

$$
\pi_{mix}(\lambda|y) = \sum_{k=1}^{K} \left( \frac{\omega_k \pi_k(y)}{\sum_{k=1}^{K} \omega_k \pi_k(y)} \right) \frac{\pi_k(\lambda, y)}{\pi_k(y)} = \sum_{k=1}^{K} \omega_k(y) \pi_k(\lambda|y). \tag{A.35}
$$

Because $\hat{\lambda}|\lambda \sim N(\lambda, \sigma^2/T)$, the posterior mean function is given by

$$
m(\hat{\lambda}, y; \pi_{mix}) = \sum_{k=1}^{K} \left( \frac{\omega_k(y) \int \pi_k(\lambda|y) \phi_N(\hat{\lambda}; \lambda, \sigma^2/T) d\lambda}{\sum_{k=1}^{K} \omega_k(y) \int \pi_k(\lambda|y) \phi_N(\hat{\lambda}; \lambda, \sigma^2/T) d\lambda} \right) \frac{\int \lambda \pi_k(\lambda|y) \phi_N(\hat{\lambda}; \lambda, \sigma^2/T) d\lambda}{\int \pi_k(\lambda|y) \phi_N(\hat{\lambda}; \lambda, \sigma^2/T) d\lambda} = \sum_{k=1}^{K} \hat{\omega}_k(\hat{\lambda}, y)m(\hat{\lambda}, y; \pi_k). \tag{A.36}
$$

Thus, the posterior mean is a weighted average of the posterior means derived from the $K$ mixture components. The $\hat{\omega}(\hat{\lambda}, y)$ can be interpreted as posterior probabilities of the mixture components. We can bound the posterior mean as follows:

$$
\left| m(\hat{\lambda}, y; \pi_{mix}) \right| \leq \sum_{k=1}^{K} \left| m(\hat{\lambda}, y; \pi_k) \right| = \sum_{k=1}^{K} \left| \tilde{\gamma}_{0,k} + \tilde{\gamma}_{1,k}y + \tilde{\gamma}_{2,k}\hat{\lambda} \right| \leq M_0 + M_y|y| + M_\lambda|\hat{\lambda}|, \tag{A.37}
$$

where the $\tilde{\gamma}$ coefficients were defined in (A.33) and are bounded for $\pi_k \in \Pi$. Thus, that the overall bound for $\left| m(\hat{\lambda}, y; \pi_{mix}) \right|$ is piecewise linear in $y$ and $\hat{\lambda}$. The joint sampling distribution of $(\hat{\lambda}, y)$ is given by the following mixture of normals:

$$
p(\hat{\lambda}, y; \pi_{mix}) = \int p(\hat{\lambda}|\lambda) \sum_{k=1}^{K} \omega_k \pi_k(\lambda, y) d\lambda = \sum_{k=1}^{K} \omega_k p(\hat{\lambda}, y; \pi_k). \tag{A.38}
$$

Based on (A.37) and (A.38) one can establish the uniform tailbounds in the assumption. Calculations under the $p_*(\cdot)$ distributions are very similar.

A.3 Derivations for Section 4

A.3.1 Consistency of QMLE in Experiments 2 and 3

We show for the basic dynamic panel data model that even if the Gaussian correlated random effects distribution is misspecified, the pseudo-true value of the QMLE estimator of $\theta$
corresponds to the “true” $\theta_0$. We do so, by calculating

$$(\theta_\ast, \xi_\ast) = \arg\max_{\theta, \xi} \mathbb{E}_{\theta_0}^Y [\ln p(Y, X_2|H, \theta, \xi)],$$

(A.39)

and verifying that $\theta_\ast = \theta_0$. Because the observations are conditionally independent across $i$ and the likelihood function is symmetric with respect to $i$, we can drop the $i$ subscripts.

We make some adjustment to the notation. The covariance matrix $\Sigma$ only depends on $\gamma$, but not on $(\rho, \alpha)$. Moreover, we will split $\xi$ into the parameters that characterize the conditional mean of $\lambda$, denoted by $\Phi$, and $\omega$, which are the non-redundant elements of the prior covariance matrix $\Omega$. Finally, we define

$$\tilde{Y}(\theta_1) = Y - X\rho - Z\alpha$$

with the understanding that $\theta_1 = (\rho, \alpha)$ and excludes $\gamma$. Moreover, let $\phi = \text{vec}(\Phi')$ and $\tilde{h}' = I \otimes h'$, such that we can write $\Phi h = \tilde{h}' \phi$. Using this notation, we can write

$$\ln p(y, x_2|h, \theta_1, \gamma, \phi, \omega) = C - \frac{1}{2} \ln |\Sigma(\gamma)| - \frac{1}{2} ((\tilde{y}(\theta_1) - w\hat{\lambda}(\theta))'\Sigma^{-1}(\gamma)(\tilde{y}(\theta_1) - w\hat{\lambda}(\theta))$$

$$- \frac{1}{2} \ln |\Omega| + \frac{1}{2} \ln |\tilde{\Omega}(\gamma, \omega)|$$

$$- \frac{1}{2} \left( \hat{\lambda}(\theta)'w'\Sigma^{-1}(\gamma)w\hat{\lambda}(\theta) + \phi'\tilde{h}'\tilde{\Omega}^{-1}\tilde{h}'\phi - \tilde{\lambda}'(\theta, \xi)\tilde{\Omega}^{-1}(\gamma, \omega)\tilde{\lambda}(\theta, \xi) \right),$$

where

$$\hat{\lambda}(\theta) = (w'\Sigma^{-1}(\gamma)w)^{-1}w'\Sigma^{-1}(\gamma)\tilde{y}(\theta_1)$$

$$\tilde{\Omega}^{-1}(\gamma, \omega) = \Omega^{-1} + w'\Sigma^{-1}(\gamma)w, \quad \tilde{\lambda}(\theta, \xi) = \tilde{\Omega}(\gamma, \omega)(\tilde{\Omega}^{-1}\tilde{h}'\phi + w'\Sigma^{-1}(\gamma)w\hat{\lambda}(\theta)).$$

In the basic dynamic panel data model $\lambda$ is scalar, $w = \iota$, $\Sigma(\gamma) = \gamma^2 I$, $x_2 = \emptyset$, $z = \emptyset$, $h = [1, y_0]'$, $\Omega = \omega^2$. Thus, splitting the $(T - 1)(\ln \gamma^2)/2$, we can write

$$\ln p(y|h, \rho, \gamma, \phi, \omega) = - \frac{1}{2} \ln |\gamma^2| - \frac{1}{2\gamma^2} ((\tilde{y}(\rho) - i\hat{\lambda}(\rho))'((\tilde{y}(\rho) - i\hat{\lambda}(\rho))$$

$$- \frac{1}{2} \ln |\omega^2| - \frac{1}{2} \ln |\gamma^2/T| + \frac{1}{2} \ln (1/T) + \frac{1}{2} \ln |\tilde{\Omega}(\gamma, \omega)|$$

$$- \frac{1}{2} \left( \frac{T}{\gamma^2} \tilde{\lambda}^2(\rho) + \frac{1}{\omega^2} \phi'\tilde{h}'\phi - \frac{1}{\tilde{\Omega}(\gamma, \omega)}\tilde{\lambda}^2(\theta, \xi) \right),$$
To evaluate $\phi$
We deduce that

$\underline{\Omega}^{-1}(\gamma, \omega) = \frac{1}{\omega^2 + \frac{1}{\gamma^2/T}}$, $\bar{\lambda}(\theta, \xi) = \bar{\Omega}(\gamma, \omega) \left( \frac{1}{\omega^2} \tilde{h}' \phi + \frac{T}{\gamma^2} \tilde{\lambda}(\rho) \right)$.

Note that

$$-\frac{1}{2} \ln |\omega^2| + \frac{1}{2} \ln |T/\gamma^2| + \frac{1}{2} \ln |\bar{\Omega}(\gamma, \omega)| = \frac{1}{2} \ln \left| \frac{\frac{T}{\gamma^2}}{\frac{1}{\omega^2} + \frac{1}{\gamma^2/T}} \right| = -\frac{1}{2} \ln |\omega^2 + \gamma^2/T|.$$

In turn, we can write

$$\ln p(y|h, \rho, \gamma, \phi, \omega) = C - T - \frac{1}{2} \ln |\gamma^2| - \frac{1}{2} \tilde{y}(\rho)'(I - \omega'/T)\tilde{y}(\rho) - \frac{1}{2} \ln |\omega^2 + \gamma^2/T|$$

$$-\frac{1}{2} \left( \frac{T}{\gamma^2} \hat{\lambda}^2(\rho) + \frac{1}{\omega^2} \phi' \tilde{h} \tilde{h}' \phi - \frac{\omega^2 \gamma^2 / T}{\omega^2 + \gamma^2 / T} \left( \frac{1}{\omega^2} \tilde{h}' \phi + \frac{T}{\gamma^2} \hat{\lambda}(\rho) \right)^2 \right)$$

$$= C - T - \frac{1}{2} \ln |\gamma^2| - \frac{1}{2} \tilde{y}(\rho)'(I - \omega'/T)\tilde{y}(\rho) - \frac{1}{2} \ln |\omega^2 + \gamma^2/T|$$

$$-\frac{1}{2(\omega^2 + \gamma^2 / T)} \left( \phi' \tilde{h} \tilde{h}' \phi - 2 \hat{\lambda}(\rho) \tilde{h}' \phi + \hat{\lambda}^2(\rho) \right).$$

Taking expectations (we omit the subscripts from the expectation operator), we can write

$$\mathbb{E} \left[ \ln p(Y|H, \rho, \gamma, \phi, \omega) \right]$$

$$= C - T - \frac{1}{2} \ln |\gamma^2| - \frac{1}{2} \mathbb{E}[\tilde{y}(\rho)'(I - \omega'/T)\tilde{y}(\rho)] - \frac{1}{2} \ln |\omega^2 + \gamma^2 / T|$$

$$-\frac{1}{2(\omega^2 + \gamma^2 / T)} \left( \phi - \mathbb{E}[\tilde{H} \tilde{H}']^{-1} \mathbb{E}[\tilde{H} \hat{\lambda}(\rho)] \right)^2 \mathbb{E}[\tilde{H} \tilde{H}'] \left( \phi - \mathbb{E}[\tilde{H} \tilde{H}']^{-1} \mathbb{E}[\tilde{H} \hat{\lambda}(\rho)] \right)$$

$$-\mathbb{E}[\hat{\lambda}(\rho) \tilde{H}^2(\mathbb{E}[\tilde{H} \tilde{H}']^{-1} \mathbb{E}[\tilde{H} \hat{\lambda}(\rho)] + \mathbb{E}[\hat{\lambda}^2(\rho)])].$$

We deduce that

$$\phi_*(\rho) = \mathbb{E}[\tilde{H} \tilde{H}']^{-1} \mathbb{E}[\tilde{H} \hat{\lambda}(\rho)].$$

To evaluate $\phi_*(\rho_0)$, note that $\hat{\lambda}(\rho_0) = \lambda + \ell' u / T$. Using that fact that the initial observation $Y_{i0}$ is uncorrelated with the shocks $U_{it}$, $t \geq 1$, we deduce that $\mathbb{E}[\tilde{H} \hat{\lambda}(\rho_0)] = \mathbb{E}[\tilde{H} \lambda]$. Thus,

$$\phi_*(\rho_0) = \mathbb{E}[\tilde{H} \tilde{H}']^{-1} \mathbb{E}[\tilde{H} \lambda].$$
The pseudo-true value is obtained through a population regression of $\lambda$ on $H$.

Plugging the pseudo-true value for $\phi$ into (A.41) yields the concentrated objective function

$$
\mathbb{E}\left[ \ln p(Y|H, \rho, \gamma, \phi_*(\rho), \omega) \right] = C - \frac{T-1}{2} \ln |\mathbb{E}[\tilde{Y}(\rho)'(I - u'/T)\tilde{Y}(\rho)]|
$$

$$
- \frac{1}{2} \ln \left| \mathbb{E}[\tilde{\lambda}^2(\rho)] - \mathbb{E}[\tilde{\lambda}(\rho)\tilde{H}'](\mathbb{E}[\tilde{H}\tilde{H}'])^{-1}\mathbb{E}[\tilde{H}\tilde{\lambda}(\rho)] \right|.
$$

Using well-known results for the maximum likelihood estimator of a variance parameter in a Gaussian regression model, we can immediately deduce that

$$
\gamma_*(\rho) = \frac{1}{T-1} \mathbb{E}[\tilde{Y}(\rho)'(I - u'/T)\tilde{Y}(\rho)]
$$

$$
\omega_*(\rho) + \gamma_*(\rho)/T = \left( \mathbb{E}[\tilde{\lambda}^2(\rho)] - \mathbb{E}[\tilde{\lambda}(\rho)\tilde{H}'](\mathbb{E}[\tilde{H}\tilde{H}'])^{-1}\mathbb{E}[\tilde{H}\tilde{\lambda}(\rho)] \right).
$$

At $\rho = \rho_0$ we obtain $\tilde{Y}(\rho_0) = \mu + u$. Thus, $\mathbb{E}[\tilde{\lambda}^2(\rho_0)] = \gamma_0^2/T + \mathbb{E}[\lambda^2]$ and $\mathbb{E}[\tilde{H}\tilde{\lambda}(\rho_0)] = \mathbb{E}[\tilde{H}\lambda]$. In turn,

$$
\gamma_*(\rho_0) = \gamma_0^2, \quad \omega_*(\rho_0) = \mathbb{E}[\lambda^2] - \mathbb{E}[\lambda\tilde{H}'](\mathbb{E}[\tilde{H}\tilde{H}'])^{-1}\mathbb{E}[\tilde{H}\lambda].
$$

Given $\rho = \rho_0$ the pseudo-true value for $\gamma^2$ is the “true” $\gamma_0^2$ and the pseudo-true variance of the correlated random-effects distribution is given by the expected value of the squared residual from a projection of $\lambda$ onto $H$.

Using (A.45), we can now concentrate out $\gamma^2$ and $\omega^2$ from the objective function (A.44):

$$
\mathbb{E}\left[ \ln p(Y|H, \rho, \gamma_*(\rho), \phi_*(\rho), \omega_*(\rho)) \right] = C - \frac{T-1}{2} \ln \left| \mathbb{E}[\tilde{Y}(\rho)'(I - u'/T)\tilde{Y}(\rho)] \right|
$$

$$
- \frac{1}{2} \ln \left| \mathbb{E}[\tilde{Y}'(\rho)u't\tilde{Y}(\rho)] - \mathbb{E}[\tilde{Y}'(\rho)\tilde{H}'](\mathbb{E}[\tilde{H}\tilde{H}'])^{-1}\mathbb{E}[\tilde{H}u't\tilde{Y}(\rho)] \right|.
$$

To find the maximum of $\mathbb{E}\left[ \ln p(Y|H, \rho, \gamma_*(\rho), \phi_*(\rho), \omega_*(\rho)) \right]$ with respect to $\rho$ we will calculate the first-order condition. Differentiating (A.47) with respect to $\rho$ yields

$$
\text{F.O.C.}(\rho) = (T-1) \frac{\mathbb{E}[X'(I - u'/T)\tilde{Y}(\rho)]}{\mathbb{E}[\tilde{Y}(\rho)'(I - u'/T)\tilde{Y}(\rho)]}
$$

$$
+ \frac{\mathbb{E}[X'ut\tilde{Y}(\rho)] - \mathbb{E}[X'\tilde{H}'](\mathbb{E}[\tilde{H}\tilde{H}'])^{-1}\mathbb{E}[\tilde{H}u't\tilde{Y}(\rho)]}{\mathbb{E}[\tilde{Y}'(\rho)u't\tilde{Y}(\rho)] - \mathbb{E}[\tilde{Y}'(\rho)\tilde{H}'](\mathbb{E}[\tilde{H}\tilde{H}'])^{-1}\mathbb{E}[\tilde{H}u't\tilde{Y}(\rho)]}.
$$
We will now verify that F.O.C.($\rho_0$) = 0. Because both denominators are strictly positive, we can rewrite the condition as

$$\text{F.O.C.}(\rho_0) = (T - 1)\mathbb{E}[X'(I - \mu' / T)\tilde{Y}(\rho_0)]$$

\[ \times \left( \mathbb{E}[\tilde{Y}'(\rho_0)\mu' \tilde{Y}(\rho_0)] - \mathbb{E}[\tilde{Y}'(\rho_0)\mu' \tilde{H}'](\mathbb{E}[\tilde{H} \tilde{H}'])^{-1}\mathbb{E}[\tilde{H}' \tilde{Y}(\rho_0)] \right) 
+ \mathbb{E}[\tilde{Y}(\rho_0)'(I - \mu' / T)\tilde{Y}(\rho_0)] 
\times \left( \mathbb{E}[X'\mu' \tilde{Y}(\rho_0)] - \mathbb{E}[X'\mu' \tilde{H}'](\mathbb{E}[\tilde{H} \tilde{H}'])^{-1}\mathbb{E}[\tilde{H}' \tilde{Y}(\rho_0)] \right). \]

Using again the fact that $\tilde{Y}(\rho_0) = \iota \lambda + U$, we can rewrite the terms appearing in the first-order condition as follows:

$$\mathbb{E}[X'(I - \mu' / T)\tilde{Y}(\rho_0)] = \mathbb{E}[X'(I - \mu' / T)u] = \mathbb{E}[X'u] - \mathbb{E}[X'\mu' u]/T = -\mathbb{E}[X'\mu' u]/T$$

$$\mathbb{E}[\tilde{Y}'(\rho_0)\mu' \tilde{Y}(\rho_0)] = \mathbb{E}[(\lambda u' + \mu')\mu' (\iota \lambda + u)] = T^2\mathbb{E}[\lambda^2] + \mathbb{E}[\mu' \mu' u] = T^2\mathbb{E}[\lambda^2] + T\gamma_0^2$$

$$\mathbb{E}[\tilde{H}' \tilde{Y}(\rho_0)] = \mathbb{E}[\tilde{H}'(\iota \lambda + u)] = T\mathbb{E}[\tilde{H} \lambda]$$

$$\mathbb{E}[\tilde{Y}(\rho_0)'(I - \mu' / T)\tilde{Y}(\rho_0)] = \mathbb{E}[\mu'(I - \mu' / T)u] = (T - 1)\gamma_0^2$$

$$\mathbb{E}[X'\mu' \tilde{Y}(\rho_0)] = \mathbb{E}[X'\mu'(\iota \lambda + u)] = T\mathbb{E}[X' \iota \lambda] + \mathbb{E}[X'\mu' u].$$

For the first equality we used the fact that $X_{it} = Y_{it-1}$ is uncorrelated with $U_{it}$. We can now re-state the first-order condition (A.48) as follows:

$$\text{F.O.C.}(\rho_0)$$

\[ = -(T - 1)(\mathbb{E}[X'\mu' u]) \left( \gamma_0^2 + T(\mathbb{E}[\lambda^2] - \mathbb{E}[\lambda \tilde{H}'](\mathbb{E}[\tilde{H} \tilde{H}'])^{-1}\mathbb{E}[\tilde{H} \lambda]) \right) 
+ \left( \mathbb{E}[X'\mu' u] + T(\mathbb{E}[X' \iota \lambda] - \mathbb{E}[X'\mu' \tilde{H}'](\mathbb{E}[\tilde{H} \tilde{H}'])^{-1}\mathbb{E}[\tilde{H} \lambda]) \right)(T - 1)\gamma_0^2 
\times \left( \mathbb{E}[X'\mu' u] - \mathbb{E}[X'\mu' \tilde{H}'](\mathbb{E}[\tilde{H} \tilde{H}'])^{-1}\mathbb{E}[\tilde{H} \lambda] \right) - \mathbb{E}[X'\mu' u] \left( \mathbb{E}[\lambda^2] \right) \right] \right). \]

We now have to analyze the terms involving $X'\mu$. Note that we can express

$$Y_t = \rho_0^t Y_0 + \sum_{\tau=0}^{t-1} \rho_0^\tau (\lambda + U_{t-\tau}).$$
Define $a_t = \sum_{\tau=0}^{t-1} \rho^\tau_0$ and $b = \sum_{t=1}^{T-1} a_t$. Thus, we can write

$$Y_t = \rho^t_0 Y_0 + \lambda a_t + \sum_{\tau=0}^{t-1} \rho^\tau_0 U_{t-\tau}, \quad t > 0.$$  

Consequently,

$$X_t' = \sum_{t=0}^{T-1} Y_t = Y_0 \left( \sum_{t=0}^{T-1} \rho^t_0 \right) + \lambda \left( \sum_{t=1}^{T-1} a_t \right) + \sum_{t=1}^{T-1} \sum_{\tau=0}^{t-1} \rho^\tau_0 U_{t-\tau} = a_T y_0 + b \lambda + \sum_{t=1}^{T-1} a_t U_{T-t}.$$  

Thus, we obtain

$$\mathbb{E}[X'_t u'] = \mathbb{E} \left[ \left( a_T Y_0 + b \lambda + \sum_{t=1}^{T-1} a_t U_{T-t} \right) \left( \sum_{t=1}^{T} U_t \right) \right] = b \gamma^2_0$$

$$\mathbb{E}[X'_t \lambda] = \mathbb{E} \left[ \left( a_T Y_0 + b \lambda + \sum_{t=1}^{T-1} a_t U_{T-t} \right) \lambda \right] = a_T \mathbb{E}[Y_0 \lambda] + b \mathbb{E}[[\lambda^2]$$

$$\mathbb{E}[X'_t \tilde{H}'] = \mathbb{E} \left[ \left( a_T Y_0 + b \lambda + \sum_{t=1}^{T-1} a_t U_{T-t} \right) \tilde{H}' \right] = a_T \mathbb{E}[Y_0 \tilde{H}'] + b \mathbb{E}[[\lambda \tilde{H}']$$

Using these expressions, most terms that appear in (A.49) cancel out and the condition simplifies to

$$\text{F.O.C.}(\rho_0) = T(T-1) \gamma_0 a_T \left( \mathbb{E}[Y_0 \lambda] - \mathbb{E}[Y_0 \tilde{H}'] (\mathbb{E}[\tilde{H} \tilde{H}'])^{-1} \mathbb{E}[\tilde{H} \lambda] \right). \quad (A.50)$$

Now consider

$$\mathbb{E}[Y_0 \tilde{H}'] (\mathbb{E}[\tilde{H} \tilde{H}'])^{-1} \mathbb{E}[\tilde{H} \lambda]$$

$$= \frac{1}{\mathbb{E}[Y_0^2] - (\mathbb{E}[Y_0])} \left[ \begin{array}{cc} \mathbb{E}[Y_0] & \mathbb{E}[Y_0^2] \\ -\mathbb{E}[Y_0] & 1 \end{array} \right] \left[ \begin{array}{c} \mathbb{E}[Y_0] \\ -\mathbb{E}[Y_0] \end{array} \right]$$

$$= \mathbb{E}[Y_0 \lambda].$$

Thus, we obtain the desired result that $\text{F.O.C.}(\rho_0) = 0$. To summarize, the pseudo-true values are given by

$$\rho_* = \rho_0, \quad \gamma_*^2 = \gamma_0, \quad \phi_* = (\mathbb{E}[\tilde{H} \tilde{H}'])^{-1} \mathbb{E}[\tilde{H} \lambda], \quad (A.51)$$

$$\omega_*^2 = \mathbb{E}[\lambda^2] - \mathbb{E}[\lambda \tilde{H}'] (\mathbb{E}[\tilde{H} \tilde{H}'])^{-1} \mathbb{E}[\tilde{H} \lambda].$$

■
A.3.2 Computation of the Oracle Predictor in Experiment 3

We are using a Gibbs sampler to compute the oracle predictor under the mixture distributions for both $\lambda_i$ and $U_{it}$.

Here we combine the scale mixture and the location mixture in a unified framework. Let $a_{it} = 1$ if $U_{it}$ is generated from the mixture component with mean $\mu_+$ and variance $\gamma^2_+$, and $a_{it} = 0$ if $U_{it}$ is generated from the mixture component with mean $\mu_-$ variance $\gamma^2_-$. Then, $\mu_+ = \mu_- = 0$ for the scale mixture, and $\gamma^2_+ = \gamma^2_- = \gamma^2$ for the location mixture. Also, let $b_i$ be an indicator of the components in the correlated random effects distribution, such that

$$
\phi(Y_{i0}, b_i) = \begin{cases} 
\phi_+(Y_{i0}), & \text{if } b_i = 1, \\
\phi_-(Y_{i0}), & \text{if } b_i = 0.
\end{cases}
$$

Omitting $i$ subscripts from now on, define

$$
\tilde{Y}_t = Y_t - \rho Y_{t-1} - (a_t \mu_+ + (1 - a_t) \mu_-), \quad \gamma^2(a_t) = a_t \gamma^2_+ + (1 - a_t) \gamma^2_-, 
$$

so we have

$$
\tilde{Y}_t(a_t) \mid (\lambda, a_t) \sim N(\lambda, \gamma^2(a_t)).
$$

Conditional on $b$, the prior distribution is

$$
\lambda \mid (Y_0, b) \sim N(\phi(Y_0, b), \Omega),
$$

and we obtain a posterior distribution of the form

$$
\lambda \mid (Y_{0:T}, a_{1:T}, b) \sim N(\tilde{\lambda}(a_{1:T}, b), \tilde{\Omega}(a_{1:T})), \quad (A.52)
$$

where

$$
\tilde{\Omega}(a_{1:T}) = (\Omega^{-1} + \sum_{t=1}^{T} (\gamma^2(a_t))^{-1})^{-1}
$$

and

$$
\tilde{\lambda}(a_{1:T}, b) = \tilde{\Omega}(a_{1:T})(\Omega^{-1} \phi(Y_0, b) + \sum_{t=1}^{T} (\gamma^2(a_t))^{-1} \tilde{Y}_t(a_t)).
$$
The posterior probability of $a_t = 1$ conditional on $(\lambda, Y_{0:T})$ is given by
\[
\mathbb{P}(a_t = 1 | \lambda, Y_{0:T}) \quad (A.53)
\]
\[
= \frac{p_u(\gamma_+)^{-1} \exp \left\{ \frac{-1}{2\gamma_+} (\bar{Y}_t(1) - \lambda)^2 \right\}}{p_u(\gamma_+)^{-1} \exp \left\{ \frac{-1}{2\gamma_+} (\bar{Y}_t(1) - \lambda)^2 \right\} + (1 - p_u)(\gamma_-)^{-1} \exp \left\{ \frac{-1}{2\gamma_-} (\bar{Y}_t(0) - \lambda)^2 \right\}},
\]
And the posterior probability of $b = 1$ conditional on $(\lambda, Y_{0:T}, a_{1:T})$ is given by
\[
\mathbb{P}(b = 1 | \lambda, Y_{0:T}, a_{1:T}) \quad (A.54)
\]
\[
= \frac{p_\lambda \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T} \frac{(\bar{Y}_t(a_t) - \phi_+(Y_0))^2}{\Omega + \gamma^2(a_t)} \right\}}{p_\lambda \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T} \frac{(\bar{Y}_t(a_t) - \phi_+(Y_0))^2}{\Omega + \gamma^2(a_t)} \right\} + (1 - p_\lambda) \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T} \frac{(\bar{Y}_t(a_t) - \phi_-(Y_0))^2}{\Omega + \gamma^2(a_t)} \right\}}.
\]

The posterior mean $\mathbb{E}[\lambda | Y_{0:T}]$ can be approximated with the following Gibbs sampler. Generate a sequence of draws $\{\lambda^s, a^s_{1:T}, b^s\}_{s=1}^{N_{sim}}$ by iterating over the conditional distributions given in (A.52), (A.53), and (A.54). Denote $\bar{p}(a_{1:T}, b) = \mathbb{P}(b | \lambda, Y_{0:T}, a_{1:T})$, then,
\[
\hat{\mathbb{E}}[\lambda | Y_{0:T}] = \frac{1}{N_{sim}} \sum_{s=1}^{N_{sim}} \sum_{b=0}^{1} \bar{p}(a^s_{1:T}, b) \lambda(a^s_{1:T}, b), \quad (A.55)
\]
\[
\hat{\mathbb{V}}[\lambda | Y_{0:T}] = \frac{1}{N_{sim}} \sum_{s=1}^{N_{sim}} \left( \Omega(a^s_{1:T}) + \sum_{b=0}^{1} \bar{p}(a^s_{1:T}, b) \lambda^2(a^s_{1:T}, b) \right) - \left( \hat{\mathbb{E}}[\lambda | Y_{0:T}] \right)^2.
\]
A.4 Proofs for Section 5.3

Lemma A.9 Suppose that $T \geq k_w + 1 \geq 2$. Suppose that $W$ is a $T \times k_w$ matrix with rank$(W) = k_w$. Let $\Sigma$ be a $T \times T$ matrix of rank $T$. Let $S = \Sigma W$. Then, rank$(M_{S \otimes S}) = T$, where $M_{S \otimes S}$ and $B$ are defined in the proof of Theorem 5.3.

Proof of Lemma A.9. Notice that the matrix $B$ is a $T^2 \times T$ selection matrix that has one at positions $(1, 1), (T + 2, 2), (2T + 3, 3), \ldots, (T^2, T)$ and zeros at the other positions. Notice that since $\Sigma$ is full rank, rank$(S) = \text{rank}(\Sigma W) = \text{rank}(W) = k_w$. If rank$(S) = k_w$, then rank$(S \otimes S) = k_w^2$. Since the rank of the projection matrix is the same as its trace, we have rank$(M_{S \otimes S}) = tr(M_{S \otimes S}) = T^2 - k_w^2$.

By the spectral decomposition, we can decompose $M_{S \otimes S} = F \Lambda F'$, where $F$ is a $T^2 \times T$ orthogonal matrix and $\Lambda$ is a $T^2 \times T^2$ diagonal matrix whose first $T^2 - k_w^2$ elements are one and the rest are zero. Since $F$ is full rank, rank$(M_{S \otimes S} B) = \text{rank}(F \Lambda F' B) = \text{rank}(\Lambda F' B)$. Notice that $F' B$ is a $T^2 \times T$ matrix that collects the columns of $F'$ in the positions of $1, T + 2, 2T + 3, \ldots, T^2$. Since the columns of $F'$ are linearly independent, rank$(F' B) = T$. Notice that $\Lambda F' B$ is a submatrix of $F' B$ that selects the first $T^2 - k_w^2$ rows. Since $T - 1 \geq k_w$ and $T \geq 2$ implies that $T^2 - k_w^2 \geq 2T - 1 > T$, the $(T^2 - k_w^2) \times T$ submatrix of $F' B$, $\Lambda F' B$, has rank $T$. □

The matrix $E[(W_{it}', X_{it}', Z_{it}')'(W_{it}', X_{it}', Z_{it})]$ has full rank for $t = 1, \ldots, T$. The matrices $\sum_{s=t+1}^{T} W_{is-1} W_{is-1}'$ are invertible with probability one for all $t = 1, \ldots, T - k_w$ and $i = 1, \ldots, N$.

Proof of Theorem 5.3. (i) The parameters $\alpha$ and $\rho$ are identifiable by Assumption 5.2.

(ii) Let $Y_i, W_i, X_i, Z_i$ and $U_i$ denote the matrices vectors that stack $Y_{it}, W_{it-1}', X_{it-1}', Z_{it-1}', U_{it}$, respectively, for $t = 1, \ldots, T$. Define

$$
\Sigma_i^{1/2}(\gamma) = \text{diag}(\sigma_1(h_i, \gamma_1), \ldots, \sigma_T(h_i, \gamma_T)),
$$

$$
S_i(\gamma) = \Sigma_i^{-1/2}(\gamma) W_i, \quad M_i(\gamma) = I - S_i(\gamma) S_i^{-1} S_i'.
$$

Using this notation, we obtain

$$
M_i(\gamma) \Sigma_i^{-1/2}(\gamma)(Y_i - X_i \rho - Z_i \alpha) = M_i(\gamma) S_i(\gamma) \lambda_i + M_i(\gamma) \Sigma_i^{-1/2}(\gamma) U_i = M_i(\gamma) V_i.
$$
This leads to the conditional moment condition
\[
\mathbb{E}[M_i(\tilde{\gamma}) \Sigma_i^{-1/2}(\tilde{\gamma}) (Y_i - X_i\rho - Z_i\alpha)(Y_i - X_i\rho - Z_i\alpha)' \Sigma_i^{-1/2}(\tilde{\gamma}) M_i'(\tilde{\gamma}) - M_i(\tilde{\gamma}) | H_i] = 0,
\]
which can be rewritten as
\[
M_i(\tilde{\gamma}) (\Sigma_i^{-1/2}(\tilde{\gamma}) \Sigma_i(\gamma) \Sigma_i^{-1/2}(\tilde{\gamma}) - I) M_i'(\tilde{\gamma}) = 0. \tag{A.56}
\]
for each \(h_i\). Taking expectations with respect to \(H_i\) and using Assumption 5.2(ii), we deduce that
\[
\mathbb{E}[M_i(\tilde{\gamma}) (\Sigma_i^{-1/2}(\tilde{\gamma}) \Sigma_i(\gamma) \Sigma_i^{-1/2}(\tilde{\gamma}) - I) M_i'(\tilde{\gamma})] = 0. \tag{A.57}
\]
if and only if \(\tilde{\gamma} = \gamma\).

(iii) The subsequent argument is similar to the proof of Theorem 2 in Arellano and Bonhomme (2012). Conditional on \(\rho, \alpha,\) and \(\gamma\) we can remove the effect of \(X_i\) and \(Z_i\) from \(Y_i\) and define
\[
\tilde{Y}_i = \Sigma_i^{-1/2}(\gamma)(Y_i - X_i\rho - Z_i\alpha) = S_i(\gamma)\lambda_i + V_i. \tag{A.58}
\]
To simplify the notation, we will omit the \(i\) subscripts and the \(\gamma\) argument in the remainder of the proof.

Because \(S(\gamma), \lambda\) and \(V\) are independent conditional on \(H\) (and \(\gamma\)), we have
\[
\ln \Psi_{\tilde{Y}}(\tau | h) = \ln \Psi_{\lambda}(S'\tau|h) + \ln \Psi_{V}(\tau) \tag{A.59}
\]
Taking the second derivative with respect to \(\tau\) leads to
\[
\frac{\partial^2}{\partial \tau \partial \tau'} \ln \Psi_{\tilde{Y}}(\tau | h) = \frac{\partial^2}{\partial \tau \partial \tau'} (\ln \Psi_{\lambda}(S'\tau|h)) + \frac{\partial^2}{\partial \tau \partial \tau'} \ln \Psi_{V}(\tau)
= S \left( \frac{\partial^2}{\partial \xi \partial \xi'} \ln \Psi_{\lambda}(S'\tau|h) \right) S' + \frac{\partial^2}{\partial \tau \partial \tau'} \ln \Psi_{V}(\tau). \tag{A.60}
\]
Using the assumption that the \(V_t\)s are independent over \(t\), we can write
\[
\ln \Psi_{V}(\tau) = \sum_{t=1}^{T} \ln \Psi_{V_t}(\tau_t),
\]
where $\Psi_{V_t}$ is the characteristic function of $V_t$. Then,

$$\text{vec} \left( \frac{\partial^2}{\partial \tau \partial \tau'} \ln \Psi_{V} (\tau) \right) = \text{vec} \left( \text{diag} \left( \frac{\partial^2}{\partial \tau^2} \ln \Psi_{V_1} (\tau_1), ..., \frac{\partial^2}{\partial \tau^2} \ln \Psi_{V_T} (\tau_T) \right) \right)$$  (A.61)

$$= B \left( \frac{\partial^2}{\partial \tau^2} \ln \Psi_{V_1} (\tau_1), ..., \frac{\partial^2}{\partial \tau^2} \ln \Psi_{V_T} (\tau_T) \right)'$$

for a suitably chosen matrix $B$. Let

$$M_{S \otimes S} = I - S(S'S)^{-1}S' \otimes S(S'S)^{-1}S'.$$

Then,

$$M_{S \otimes S} \text{vec} (\ln \Psi_{\tilde{Y}} (\tau|h)) = M_{S \otimes S} B \left( \frac{\partial^2}{\partial \tau^2} \ln \Psi_{V_1} (\tau_1), ..., \frac{\partial^2}{\partial \tau^2} \ln \Psi_{V_T} (\tau_T) \right)' .$$  (A.62)

Because $\Sigma(\gamma)$ is of full rank $T$ (Assumption 5.2(iii)) and $W$ is of full rank of $k_w$ (Assumption 5.2(iv)), $S(\gamma)$ has full rank $k_w$. Notice that $T \geq k_w + 1$. Then, according to Lemma A.9, $M_{S \otimes S} B$ is also full rank. In turn, from (A.62), we can identify $\ln \Psi_{V_t}(\tau_t)$ uniquely for $t = 1, ..., T$. Also using the restrictions that $\frac{\partial}{\partial \tau_t} \ln \Psi_{V_t}(0) = 0$ ($E(V_{it}) = 0$) and $\ln \Psi_{V_t}(0) = 0$, we can deduce that the characteristic function of $V_t$ is uniquely identified.

Next, we show how to identify $\ln \Psi_{\lambda}(\tau|h)$. Because $\ln \Psi_{\tilde{Y}} (\tau|h)$ and $\ln \Psi_{V}(\tau)$ are identified, from (A.59) we obtain

$$\ln \Psi_{\tilde{Y}}(\tau|h) - \ln \Psi_{V}(\tau) = \ln \Psi_{\lambda}(S'S'h).$$  (A.63)

Taking second derivatives, we obtain

$$\frac{\partial^2}{\partial \tau \partial \tau'} \left( \ln \Psi_{\tilde{Y}}(\tau|h) - \sum_{t=1}^{T} \ln \Psi_{V}(\tau_t) \right) = S \left( \frac{\partial^2}{\partial \xi \partial \xi'} \ln \Psi_{\lambda}(S'S|h) \right) S'.$$  (A.64)

Because $S$ is of full rank, we can identify

$$\frac{\partial^2}{\partial \xi \partial \xi'} \ln \Psi_{\lambda}(S'S|h) = (S'S)^{-1}S' \left[ \frac{\partial^2}{\partial \tau \partial \tau'} \left( \ln \Psi_{\tilde{Y}}(\tau|h) - \sum_{t=1}^{T} \ln \Psi_{V}(\tau_t) \right) \right] S(S'S)^{-1}. $$  (A.65)

The mean $E(\lambda|h)$ can be identified as follows. Note that

$$\hat{\lambda} = (S'S)^{-1}S'\tilde{Y} = \lambda + (S'S)^{-1}S'V.$$  (A.66)
Taking expectations yields
\[ \mathbb{E}(\lambda|h) = \mathbb{E}[^\hat{\lambda}|h], \] (A.67)
because \( \mathbb{E}[(S'S)^{-1}S'V|h] = (S'S)^{-1}S'\mathbb{E}[V|h] = 0. \) Once the mean has been determined, we can identify \( \ln \Psi_\lambda(\xi|h) \) using \( \frac{\partial}{\partial \xi} \ln \Psi_\lambda(0|h) = \mathbb{E}(\lambda|h) \) and \( \ln \Psi_\lambda(0|h) = 0. \)

**Discussion of Assumption 5.2(i).** We discuss an example of how to identify \( \alpha \) and \( \rho \) based on moment conditions in the general model (1). Under the model (1) we can remove the effect of \( \lambda_i \) with the following within projections:

\[
Y_{it}^* = Y_{it} - \left( \sum_{s=t+1}^{T} Y_{is} W_{is-1}' \right) \left( \sum_{s=t+1}^{T} W_{is-1} W_{is-1}' \right)^{-1} W_{it-1}
\]

\[
X_{it-1}^* = X_{it-1} - \left( \sum_{s=t+1}^{T} X_{is-1} W_{is-1}' \right) \left( \sum_{s=t+1}^{T} W_{is-1} W_{is-1}' \right)^{-1} W_{it-1}
\]

\[
Z_{it-1}^* = Z_{it-1} - \left( \sum_{s=t+1}^{T} Z_{is-1} W_{is-1}' \right) \left( \sum_{s=t+1}^{T} W_{is-1} W_{is-1}' \right)^{-1} W_{it-1}
\]

for \( t = 1, \ldots, T - k_w. \) Because \( \mathbb{E}[U_{it}|Y_{i1:t-1}, H_i, \lambda_i] = 0, \) we obtain the moment condition

\[
\mathbb{E} \left[ \begin{bmatrix} Y_{it}^* - [\tilde{\rho}' \tilde{\alpha}'] & X_{it-1}^* & Z_{it-1}^* \end{bmatrix} \begin{bmatrix} X_{it-s-1}^* & X_{it-s-1}^* & Z_{it-s-1}^* \end{bmatrix} \right] = 0 \] (A.68)

for \( s \geq 0. \) To simplify the exposition, suppose that we choose \([X_{it-1}, Z_{it-1}]\) as instrumental variables. In this case, for the moment conditions to be only satisfied only at \( \tilde{\rho} = \rho \) and \( \tilde{\alpha} = \alpha \) it is necessary that the matrix

\[
\mathbb{E} \left[ \begin{bmatrix} X_{it-1}^* X_{it-1}^* & X_{it-1}^* Z_{it-1}^* \\ Z_{it-1}^* X_{it-1}^* & Z_{it-1}^* Z_{it-1}^* \end{bmatrix} \right] \] (A.69)
is full rank. Consider, for instance, the upper-left element. We can write

\[
\mathbb{E}[X'_{it-1}X'_{it-1}]
\]

\[
= \mathbb{E} \left[ \left( X'_{it-1} - \left( \sum_{s=t+1}^{T} X_{is-1}W'_{is-1} \right) \left( \sum_{s=t+1}^{T} W_{is-1}W'_{is-1} \right)^{-1} W_{it-1} \right) X'_{it-1} \right]
\]

\[
= \mathbb{E} \left[ \mathbb{E} \left[ X'_{it-1} - \left( \sum_{s=t+1}^{T} X_{is-1}W'_{is-1} \right) \left( \sum_{s=t+1}^{T} W_{is-1}W'_{is-1} \right)^{-1} W_{it-1} \right) \left| W_{i}^{t:T-1} \right] \right]
\]

\[
= \mathbb{E}[X_{it-1}X'_{it-1}] - \frac{1}{T-h} \left( \sum_{s=t+1}^{T} \mathbb{E} \left[ X_{is-1}X_{it-1} \left| W_{i}^{t:T-1} \right. \right] \right)
\]

\[
\times W'_{is-1} \left( \frac{1}{T-h} \sum_{s=t+1}^{T} W_{is-1}W'_{is-1} \right)^{-1} W_{it-1} \right)
\]

\[
= \mathbb{E}[X_{it-1}X'_{it-1}] - \frac{1}{T-h} \sum_{s=t+1}^{T} \kappa_s \mathbb{E}[X_{is-1}X'_{it-1}] = I + II, \text{ say.}
\]

The fourth equality is based on the assumption that the \( W_{it}'s \) are strictly exogenous. The completion of the identification argument requires a moment bound for

\[
\kappa_s = \mathbb{E} \left[ W'_{is-1} \left( \frac{1}{T-h} \sum_{s=t+1}^{T} W_{is-1}W'_{is-1} \right)^{-1} W_{it-1} \right],
\]

a full rank condition on \( \mathbb{E}[X_{it-1}X'_{it-1}] \), and a condition that ensures that term \( II \) does not induce a rank deficiency in term \( I \). Similar conditions need to be imposed on the terms that appear in the other submatrices of (A.69).
\section*{B Data Set}

The construction of our data is based on Covas, Rump, and Zakrajsek (2014). We downloaded FR Y-9C BHC financial statements for the quarters 2002Q1 to 2014Q4 using the web portal of the Federal Reserve Bank of Chicago. We define PPNR (relative to assets) as follows

\[ \text{PPNR} = \frac{400(NII + ONII - ONIE)}{\text{ASSETS}}, \]

where

- \( NII = \) Net Interest Income \( \text{BHCK 4074} \)
- \( ONII = \) Total Non-Interest Income \( \text{BHCK 4079} \)
- \( ONIE = \) Total Non-Interest Expenses \( \text{BHCK 4093 - C216 - C232} \)
- \( \text{ASSETS} = \) Consolidated Assets \( \text{BHCK 3368} \)

Here net interest income is the difference between total interest income and expenses. It excludes provisions for loan and lease losses. Non-interest income includes various types of fees, trading revenue, as well as net gains on asset sales. Non-interest expenses include, for instance, salaries and employee benefits and expenses of premises and fixed assets. As in Covas, Rump, and Zakrajsek (2014), we exclude impairment losses (C216 and C232). We divide the net revenues by the amount of consolidated assets. This ratio is multiplied by 400 to annualize the flow variables and convert the ratio into percentages.

The raw data take the form of an unbalanced panel of BHCs. The appearance and disappearance of specific institutions in the data set is affected by entry and exit, mergers and acquisitions, as well as changes in reporting requirements for the FR Y-9C form. Note that NII, ONII, and ONIE are reported as year-to-date values. Thus, in order to obtain quarterly data, we take differences: \( Q1 \mapsto Q1, (Q2 - Q1) \mapsto Q2, (Q3 - Q2) \mapsto Q3, \) and \( (Q4 - Q3) \mapsto Q4. \) ASSETS is a stock variable and no further transformation is needed.

Our goal is to construct rolling samples that consist of \( T + 2 \) observations, where \( T \) is the size of the estimation sample and varies between \( T = 3 \) and \( T = 11. \) The additional two observations in each rolling sample are used, respectively, to initialize the lag in the first period of the estimation sample and to compute the error of the one-step-ahead forecast. We index each rolling sample by the forecast origin \( t = \tau. \) For instance, taking the time period \( t \) to be a quarter, with data from 2002Q1 to 2014Q4 we can construct \( M = 45 \) samples of size \( T = 6 \) with forecast origins running from \( \tau = 2003Q3 \) to \( \tau = 2014Q3. \) Each rolling sample is
indexed by the pair \((\tau, T)\). The following adjustment procedure that eliminates BHCs with missing observations and outliers is applied to each rolling sample \((\tau, T)\) separately:

1. Eliminate BCHs for which total assets are missing for all time periods in the sample.
2. Compute average non-missing total assets and eliminate BCHs with average assets below 500 million dollars.
3. Eliminate BCHs for which one or more PPNR components are missing for at least one period of the sample.
4. Eliminate BCHs for which the absolute difference between the temporal mean and the temporal median exceeds 10.
5. Define deviations from temporal means as \(\delta_{it} = y_{it} - \bar{y}_i\). Pooling the \(\delta_{it}\)'s across institutions and time periods, compute the median \(q_{0.5}\) and the 0.025 and 0.975 quantiles, \(q_{0.025}\) and \(q_{0.975}\). We delete institutions for which at least one \(\delta_{it}\) falls outside of the range \(q_{0.5} \pm (q_{0.975} - q_{0.025})\).

The effect of the sample-adjustment procedure on the size of the rolling samples is summarized in Table A-1. Here we are focusing on samples with \(T = 6\) as in the main text. The column labeled \(N_0\) provides the number of raw data for each sample. In columns \(N_j\), \(j = 1, \ldots, 4\), we report the observations remaining after adjustment \(j\). Finally, \(N\) is the number of observations after the fifth adjustment. This is the relevant sample size for the subsequent empirical analysis. For many BCHs we do not have information on the consolidated assets, which leads to reduction of the sample size by 60% to 80%. Once we restrict average consolidated assets to be above 500 million dollars, the sample size shrinks to approximately 700 to 1,200 institutions. Roughly 10% to 25% of these institutions have missing observations for PPNR components, which leads to \(N_3\). The outlier elimination in Steps 4. and 5. have a relatively small effect on the sample size.

Descriptive statistics for the \(T = 6\) rolling samples are reported in Table A-2. For each rolling sample we pool observations across institutions and time periods. We do not weight the observations by the size of the institution. Notice that the mean PPNR falls from about 2% for the 2003 samples to 1.24% for the 2010Q2 sample, which includes observations starting in 2008Q4. Then, the mean slightly increases and levels off at around 1.3%. The means are close to the medians, suggesting that the samples are not very skewed, which is
confirmed by the skewness measures reported in the second to last column. The samples also exhibit fat tails. The kurtosis statistics range from 4 to 190.
Table A-1: Size of Adjusted Rolling Samples ($T = 6$)

<table>
<thead>
<tr>
<th>Sample</th>
<th>$N_0$</th>
<th>$N_1$</th>
<th>$N_2$</th>
<th>$N_3$</th>
<th>$N_4$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003Q3</td>
<td>6176</td>
<td>2258</td>
<td>710</td>
<td>653</td>
<td>653</td>
<td>614</td>
</tr>
<tr>
<td>2003Q4</td>
<td>6177</td>
<td>2289</td>
<td>730</td>
<td>658</td>
<td>658</td>
<td>618</td>
</tr>
<tr>
<td>2004Q1</td>
<td>6142</td>
<td>2351</td>
<td>744</td>
<td>660</td>
<td>660</td>
<td>622</td>
</tr>
<tr>
<td>2004Q2</td>
<td>6089</td>
<td>2375</td>
<td>754</td>
<td>657</td>
<td>657</td>
<td>613</td>
</tr>
<tr>
<td>2004Q3</td>
<td>6093</td>
<td>2416</td>
<td>778</td>
<td>669</td>
<td>668</td>
<td>624</td>
</tr>
<tr>
<td>2004Q4</td>
<td>6090</td>
<td>2448</td>
<td>787</td>
<td>668</td>
<td>667</td>
<td>621</td>
</tr>
<tr>
<td>2005Q1</td>
<td>6101</td>
<td>2486</td>
<td>797</td>
<td>680</td>
<td>679</td>
<td>629</td>
</tr>
<tr>
<td>2005Q2</td>
<td>6077</td>
<td>2489</td>
<td>809</td>
<td>695</td>
<td>694</td>
<td>644</td>
</tr>
<tr>
<td>2005Q3</td>
<td>6083</td>
<td>2473</td>
<td>826</td>
<td>718</td>
<td>717</td>
<td>660</td>
</tr>
<tr>
<td>2005Q4</td>
<td>6050</td>
<td>2451</td>
<td>828</td>
<td>728</td>
<td>727</td>
<td>658</td>
</tr>
<tr>
<td>2006Q1</td>
<td>6054</td>
<td>2425</td>
<td>834</td>
<td>715</td>
<td>715</td>
<td>664</td>
</tr>
<tr>
<td>2006Q2</td>
<td>6024</td>
<td>2403</td>
<td>849</td>
<td>734</td>
<td>734</td>
<td>685</td>
</tr>
<tr>
<td>2006Q3</td>
<td>6053</td>
<td>2376</td>
<td>858</td>
<td>747</td>
<td>747</td>
<td>697</td>
</tr>
<tr>
<td>2006Q4</td>
<td>6038</td>
<td>2367</td>
<td>880</td>
<td>757</td>
<td>757</td>
<td>711</td>
</tr>
<tr>
<td>2007Q1</td>
<td>6075</td>
<td>2355</td>
<td>905</td>
<td>772</td>
<td>772</td>
<td>727</td>
</tr>
<tr>
<td>2007Q2</td>
<td>6044</td>
<td>2337</td>
<td>929</td>
<td>777</td>
<td>777</td>
<td>732</td>
</tr>
<tr>
<td>2007Q3</td>
<td>6054</td>
<td>1101</td>
<td>941</td>
<td>773</td>
<td>773</td>
<td>712</td>
</tr>
<tr>
<td>2007Q4</td>
<td>6038</td>
<td>1061</td>
<td>919</td>
<td>769</td>
<td>769</td>
<td>710</td>
</tr>
<tr>
<td>2008Q1</td>
<td>6014</td>
<td>1081</td>
<td>945</td>
<td>770</td>
<td>770</td>
<td>713</td>
</tr>
<tr>
<td>2008Q2</td>
<td>5997</td>
<td>1070</td>
<td>942</td>
<td>775</td>
<td>775</td>
<td>722</td>
</tr>
<tr>
<td>2008Q3</td>
<td>5953</td>
<td>1062</td>
<td>949</td>
<td>784</td>
<td>784</td>
<td>731</td>
</tr>
<tr>
<td>2008Q4</td>
<td>5947</td>
<td>1058</td>
<td>949</td>
<td>792</td>
<td>792</td>
<td>741</td>
</tr>
<tr>
<td>2009Q1</td>
<td>5904</td>
<td>1113</td>
<td>1006</td>
<td>795</td>
<td>795</td>
<td>744</td>
</tr>
<tr>
<td>2009Q2</td>
<td>5878</td>
<td>1104</td>
<td>996</td>
<td>795</td>
<td>795</td>
<td>745</td>
</tr>
<tr>
<td>2009Q3</td>
<td>5805</td>
<td>1087</td>
<td>986</td>
<td>799</td>
<td>799</td>
<td>749</td>
</tr>
<tr>
<td>2009Q4</td>
<td>5793</td>
<td>1081</td>
<td>977</td>
<td>809</td>
<td>808</td>
<td>754</td>
</tr>
<tr>
<td>2010Q1</td>
<td>5709</td>
<td>1124</td>
<td>1015</td>
<td>800</td>
<td>799</td>
<td>744</td>
</tr>
<tr>
<td>2010Q2</td>
<td>5700</td>
<td>1116</td>
<td>1005</td>
<td>800</td>
<td>799</td>
<td>738</td>
</tr>
<tr>
<td>2010Q3</td>
<td>5665</td>
<td>1105</td>
<td>997</td>
<td>795</td>
<td>794</td>
<td>727</td>
</tr>
<tr>
<td>2010Q4</td>
<td>5652</td>
<td>1105</td>
<td>996</td>
<td>844</td>
<td>843</td>
<td>780</td>
</tr>
<tr>
<td>2011Q1</td>
<td>5586</td>
<td>1131</td>
<td>1027</td>
<td>838</td>
<td>837</td>
<td>773</td>
</tr>
<tr>
<td>2011Q2</td>
<td>5566</td>
<td>1129</td>
<td>1027</td>
<td>836</td>
<td>836</td>
<td>777</td>
</tr>
<tr>
<td>2011Q3</td>
<td>5483</td>
<td>1119</td>
<td>1018</td>
<td>833</td>
<td>833</td>
<td>770</td>
</tr>
<tr>
<td>2011Q4</td>
<td>5636</td>
<td>1115</td>
<td>1011</td>
<td>864</td>
<td>864</td>
<td>797</td>
</tr>
<tr>
<td>2012Q1</td>
<td>5876</td>
<td>1259</td>
<td>1154</td>
<td>863</td>
<td>863</td>
<td>794</td>
</tr>
<tr>
<td>2012Q2</td>
<td>5847</td>
<td>1240</td>
<td>1140</td>
<td>858</td>
<td>858</td>
<td>792</td>
</tr>
<tr>
<td>2012Q3</td>
<td>5809</td>
<td>1226</td>
<td>1135</td>
<td>849</td>
<td>849</td>
<td>789</td>
</tr>
<tr>
<td>2012Q4</td>
<td>5793</td>
<td>1216</td>
<td>1124</td>
<td>878</td>
<td>878</td>
<td>811</td>
</tr>
<tr>
<td>2013Q1</td>
<td>5749</td>
<td>1246</td>
<td>1157</td>
<td>875</td>
<td>875</td>
<td>808</td>
</tr>
<tr>
<td>2013Q2</td>
<td>5739</td>
<td>1245</td>
<td>1153</td>
<td>874</td>
<td>874</td>
<td>806</td>
</tr>
<tr>
<td>2013Q3</td>
<td>5699</td>
<td>1230</td>
<td>1142</td>
<td>874</td>
<td>874</td>
<td>805</td>
</tr>
<tr>
<td>2013Q4</td>
<td>5695</td>
<td>1233</td>
<td>1143</td>
<td>997</td>
<td>995</td>
<td>920</td>
</tr>
<tr>
<td>2014Q1</td>
<td>5603</td>
<td>1253</td>
<td>1162</td>
<td>979</td>
<td>977</td>
<td>899</td>
</tr>
<tr>
<td>2014Q2</td>
<td>5572</td>
<td>1237</td>
<td>1143</td>
<td>973</td>
<td>972</td>
<td>897</td>
</tr>
<tr>
<td>2014Q3</td>
<td>5514</td>
<td>1231</td>
<td>1140</td>
<td>966</td>
<td>965</td>
<td>898</td>
</tr>
</tbody>
</table>
Table A-2: Descriptive Statistics for Rolling Samples ($T = 6$)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Min</th>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>StdD</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003Q3</td>
<td>2.04</td>
<td>-2.10</td>
<td>2.00</td>
<td>12.01</td>
<td>0.90</td>
<td>3.00</td>
<td>29.46</td>
</tr>
<tr>
<td>2003Q4</td>
<td>2.02</td>
<td>-1.43</td>
<td>1.98</td>
<td>11.18</td>
<td>0.87</td>
<td>2.75</td>
<td>25.03</td>
</tr>
<tr>
<td>2004Q1</td>
<td>1.99</td>
<td>-2.10</td>
<td>1.95</td>
<td>11.18</td>
<td>0.91</td>
<td>3.13</td>
<td>29.90</td>
</tr>
<tr>
<td>2004Q2</td>
<td>1.96</td>
<td>-0.98</td>
<td>1.92</td>
<td>11.18</td>
<td>0.83</td>
<td>2.76</td>
<td>27.24</td>
</tr>
<tr>
<td>2004Q3</td>
<td>1.92</td>
<td>-0.98</td>
<td>1.89</td>
<td>10.80</td>
<td>0.76</td>
<td>2.06</td>
<td>22.28</td>
</tr>
<tr>
<td>2004Q4</td>
<td>1.90</td>
<td>-0.83</td>
<td>1.88</td>
<td>6.06</td>
<td>0.69</td>
<td>0.52</td>
<td>4.85</td>
</tr>
<tr>
<td>2005Q1</td>
<td>1.89</td>
<td>-0.73</td>
<td>1.87</td>
<td>6.01</td>
<td>0.70</td>
<td>0.62</td>
<td>4.94</td>
</tr>
<tr>
<td>2005Q2</td>
<td>1.90</td>
<td>-0.73</td>
<td>1.87</td>
<td>5.76</td>
<td>0.70</td>
<td>0.61</td>
<td>4.74</td>
</tr>
<tr>
<td>2005Q3</td>
<td>1.91</td>
<td>-0.60</td>
<td>1.87</td>
<td>9.99</td>
<td>0.74</td>
<td>1.56</td>
<td>13.97</td>
</tr>
<tr>
<td>2005Q4</td>
<td>1.88</td>
<td>-0.60</td>
<td>1.85</td>
<td>5.30</td>
<td>0.70</td>
<td>0.46</td>
<td>4.13</td>
</tr>
<tr>
<td>2006Q1</td>
<td>1.87</td>
<td>-0.60</td>
<td>1.84</td>
<td>5.30</td>
<td>0.69</td>
<td>0.50</td>
<td>4.09</td>
</tr>
<tr>
<td>2006Q2</td>
<td>1.86</td>
<td>-0.89</td>
<td>1.82</td>
<td>5.30</td>
<td>0.71</td>
<td>0.50</td>
<td>4.09</td>
</tr>
<tr>
<td>2006Q3</td>
<td>1.83</td>
<td>-2.05</td>
<td>1.80</td>
<td>5.30</td>
<td>0.74</td>
<td>0.30</td>
<td>4.58</td>
</tr>
<tr>
<td>2006Q4</td>
<td>1.81</td>
<td>-2.05</td>
<td>1.77</td>
<td>5.30</td>
<td>0.75</td>
<td>0.32</td>
<td>4.45</td>
</tr>
<tr>
<td>2007Q1</td>
<td>1.78</td>
<td>-2.19</td>
<td>1.73</td>
<td>5.30</td>
<td>0.76</td>
<td>0.30</td>
<td>4.46</td>
</tr>
<tr>
<td>2007Q2</td>
<td>1.75</td>
<td>-2.36</td>
<td>1.70</td>
<td>5.68</td>
<td>0.77</td>
<td>0.32</td>
<td>4.97</td>
</tr>
<tr>
<td>2007Q3</td>
<td>1.71</td>
<td>-1.67</td>
<td>1.67</td>
<td>5.68</td>
<td>0.75</td>
<td>0.40</td>
<td>4.94</td>
</tr>
<tr>
<td>2007Q4</td>
<td>1.67</td>
<td>-1.67</td>
<td>1.63</td>
<td>6.00</td>
<td>0.75</td>
<td>0.50</td>
<td>5.33</td>
</tr>
<tr>
<td>2008Q1</td>
<td>1.64</td>
<td>-2.20</td>
<td>1.59</td>
<td>15.92</td>
<td>0.88</td>
<td>4.21</td>
<td>61.22</td>
</tr>
<tr>
<td>2008Q2</td>
<td>1.59</td>
<td>-2.20</td>
<td>1.56</td>
<td>15.92</td>
<td>0.88</td>
<td>4.23</td>
<td>63.45</td>
</tr>
<tr>
<td>2008Q3</td>
<td>1.52</td>
<td>-2.61</td>
<td>1.51</td>
<td>15.92</td>
<td>0.90</td>
<td>3.69</td>
<td>57.87</td>
</tr>
<tr>
<td>2008Q4</td>
<td>1.46</td>
<td>-3.56</td>
<td>1.47</td>
<td>15.70</td>
<td>0.90</td>
<td>3.12</td>
<td>50.67</td>
</tr>
<tr>
<td>2009Q1</td>
<td>1.39</td>
<td>-2.61</td>
<td>1.42</td>
<td>6.53</td>
<td>0.81</td>
<td>-0.13</td>
<td>6.22</td>
</tr>
<tr>
<td>2009Q2</td>
<td>1.33</td>
<td>-2.61</td>
<td>1.37</td>
<td>6.53</td>
<td>0.83</td>
<td>-0.23</td>
<td>6.33</td>
</tr>
<tr>
<td>2009Q3</td>
<td>1.29</td>
<td>-4.10</td>
<td>1.35</td>
<td>7.53</td>
<td>0.89</td>
<td>-0.46</td>
<td>7.09</td>
</tr>
<tr>
<td>2009Q4</td>
<td>1.27</td>
<td>-4.10</td>
<td>1.33</td>
<td>7.53</td>
<td>0.87</td>
<td>-0.45</td>
<td>6.93</td>
</tr>
<tr>
<td>2010Q1</td>
<td>1.26</td>
<td>-3.59</td>
<td>1.32</td>
<td>7.53</td>
<td>0.86</td>
<td>-0.41</td>
<td>6.92</td>
</tr>
<tr>
<td>2010Q2</td>
<td>1.24</td>
<td>-3.59</td>
<td>1.30</td>
<td>5.83</td>
<td>0.85</td>
<td>-0.68</td>
<td>5.97</td>
</tr>
<tr>
<td>2010Q3</td>
<td>1.26</td>
<td>-3.54</td>
<td>1.32</td>
<td>5.83</td>
<td>0.85</td>
<td>-0.56</td>
<td>5.70</td>
</tr>
<tr>
<td>2010Q4</td>
<td>1.27</td>
<td>-3.78</td>
<td>1.32</td>
<td>7.29</td>
<td>0.88</td>
<td>-0.26</td>
<td>6.51</td>
</tr>
<tr>
<td>2011Q1</td>
<td>1.29</td>
<td>-3.32</td>
<td>1.34</td>
<td>7.29</td>
<td>0.87</td>
<td>-0.27</td>
<td>6.58</td>
</tr>
<tr>
<td>2011Q2</td>
<td>1.31</td>
<td>-3.32</td>
<td>1.36</td>
<td>8.65</td>
<td>0.90</td>
<td>0.10</td>
<td>8.05</td>
</tr>
<tr>
<td>2011Q3</td>
<td>1.31</td>
<td>-2.83</td>
<td>1.36</td>
<td>8.65</td>
<td>0.91</td>
<td>0.38</td>
<td>9.20</td>
</tr>
<tr>
<td>2011Q4</td>
<td>1.32</td>
<td>-2.83</td>
<td>1.36</td>
<td>7.98</td>
<td>0.88</td>
<td>0.26</td>
<td>8.57</td>
</tr>
<tr>
<td>2012Q1</td>
<td>1.31</td>
<td>-2.80</td>
<td>1.36</td>
<td>7.98</td>
<td>0.87</td>
<td>0.22</td>
<td>8.48</td>
</tr>
<tr>
<td>2012Q2</td>
<td>1.30</td>
<td>-2.87</td>
<td>1.35</td>
<td>7.98</td>
<td>0.88</td>
<td>0.24</td>
<td>8.46</td>
</tr>
<tr>
<td>2012Q3</td>
<td>1.32</td>
<td>-3.03</td>
<td>1.35</td>
<td>7.98</td>
<td>0.90</td>
<td>0.47</td>
<td>9.09</td>
</tr>
<tr>
<td>2012Q4</td>
<td>1.32</td>
<td>-3.03</td>
<td>1.35</td>
<td>7.98</td>
<td>0.89</td>
<td>0.49</td>
<td>9.36</td>
</tr>
<tr>
<td>2013Q1</td>
<td>1.33</td>
<td>-3.03</td>
<td>1.35</td>
<td>7.98</td>
<td>0.86</td>
<td>0.51</td>
<td>9.31</td>
</tr>
<tr>
<td>2013Q2</td>
<td>1.36</td>
<td>-2.87</td>
<td>1.34</td>
<td>22.32</td>
<td>1.07</td>
<td>7.15</td>
<td>125.30</td>
</tr>
<tr>
<td>2013Q3</td>
<td>1.32</td>
<td>-2.78</td>
<td>1.32</td>
<td>6.89</td>
<td>0.82</td>
<td>0.71</td>
<td>9.54</td>
</tr>
<tr>
<td>2013Q4</td>
<td>1.32</td>
<td>-2.78</td>
<td>1.29</td>
<td>22.32</td>
<td>1.03</td>
<td>7.39</td>
<td>133.39</td>
</tr>
<tr>
<td>2014Q1</td>
<td>1.31</td>
<td>-2.78</td>
<td>1.28</td>
<td>22.32</td>
<td>1.01</td>
<td>8.43</td>
<td>160.34</td>
</tr>
<tr>
<td>2014Q2</td>
<td>1.29</td>
<td>-2.78</td>
<td>1.28</td>
<td>7.75</td>
<td>0.79</td>
<td>1.38</td>
<td>13.12</td>
</tr>
<tr>
<td>2014Q3</td>
<td>1.33</td>
<td>-2.78</td>
<td>1.28</td>
<td>24.49</td>
<td>1.08</td>
<td>9.82</td>
<td>191.05</td>
</tr>
</tbody>
</table>

Notes: The descriptive statistics are computed for samples in which we pool observations across institutions and time periods. We did not weight the statistics by size of the institution.