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 ϵ 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 $\eta_N^c(\epsilon)$ such that

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

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 $\eta_N^a(\epsilon)$ and $\eta_N^b(\epsilon)$ such that

$$N^{-\epsilon}(A_N(\pi) + B_N(\pi)) \leq \eta_N^a(\epsilon) + \eta_N^b(\epsilon).$$

Thus, we can choose $\eta_N^c(\epsilon) = \eta_N^a(\epsilon) + \eta_N^b(\epsilon) \longrightarrow 0$ 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

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

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

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

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 $\epsilon_0 > 0$

$$\limsup_{N\longrightarrow\infty} \sup_{\pi\in\Pi} \frac{R_N(\widehat{Y}_{T+1}^N;\pi) - R_N^{\text{opt}}(\pi)}{N\mathbb{E}_{\theta;\pi}^{\mathcal{Y}^i,\lambda_i}\big[(\lambda_i - \mathbb{E}_{\theta,\pi,\mathcal{Y}^i}^{\lambda_i}[\lambda_i])^2\big] + N^{\epsilon_0}} \leq 0, \tag{A.1}$$

where

$$\begin{aligned} R_N(\widehat{Y}_{T+1}^N;\pi) &= N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N,\lambda_i}\left[\Big(\lambda_i + \rho Y_{iT} - \widehat{Y}_{iT+1}\Big)^2\right] + N\sigma^2 \\ R_N^{\text{opt}}(\pi) &= N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}_i,\lambda_i}\left[\Big(\lambda_i - \mathbb{E}_{\theta,\pi,\mathcal{Y}^i}^{\lambda_i}[\lambda_i]\Big)^2\right] + N\sigma^2. \end{aligned}$$

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

$$\limsup_{N\longrightarrow\infty} \sup_{\pi\in\Pi} \frac{N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N,\lambda_i}\left[\Big(\lambda_i + \rho Y_{iT} - \widehat{Y}_{iT+1}\Big)^2\right]}{N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^i,\lambda_i}\big[(\lambda_i - \mathbb{E}_{\theta,\pi,\mathcal{Y}^i}^{\lambda_i}[\lambda_i])^2\big] + N^{\epsilon_0}} \leq 1. \tag{A.2}$$

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

In the main text we asserted that

$$p_*(\hat{\lambda}_i, y_{i0}) = \mathbb{E}_{\theta,\pi,\mathcal{Y}_i}^{\mathcal{Y}^{(-i)}}[\hat{p}^{(-i)}(\hat{\lambda}_i, y_{i0})]. \tag{A.3}$$

This assertion can be verified as follows. Taking expectations with respect to $(\hat{\lambda}_j, y_{j,0})$ for $j \neq i$ yields

$$
\begin{aligned}
&\mathbb{E}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}_i;\pi}[\hat{p}^{(-i)}(\hat{\lambda}_i, y_{i0})] \\
&= \sum_{j\neq i} \int\int \frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_i - \hat{\lambda}_j}{B_N}\right)\frac{1}{B_N}\phi\left(\frac{y_{i0}-y_{j0}}{B_N}\right)p(\hat{\lambda}_j, y_{j0})d\hat{\lambda}_j dy_{j0} \\
&= \int\int \frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_i - \hat{\lambda}_j}{B_N}\right)\frac{1}{B_N}\phi\left(\frac{y_{i0}-y_{j0}}{B_N}\right)p(\hat{\lambda}_j, y_{j0})d\hat{\lambda}_j dy_{j0}.
\end{aligned}
$$

The second equality follows from the symmetry with respect to j and the fact that we integrate out $(\hat{\lambda}_j, y_{j0})$. We now substitute in

$$
p(\hat{\lambda}_j, y_{j0}) = \int p(\hat{\lambda}_j|\lambda_j)\pi(\lambda_j, y_{j0})d\lambda_j,
$$

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:

$$
\begin{aligned}
&\mathbb{E}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}_i}[\hat{p}^{(-i)}(\hat{\lambda}_i, y_{i0})] \\
&= \int\int\left[\int \frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_i - \hat{\lambda}_j}{B_N}\right)p(\hat{\lambda}_j|\lambda_j)d\hat{\lambda}_j\right]\frac{1}{B_N}\phi\left(\frac{y_{i0}-y_{j0}}{B_N}\right)\pi(\lambda_j, y_{j0})d\lambda_j dy_{j0} \\
&= \int\int \frac{1}{\sqrt{\sigma^2/T + B_N^2}}\phi\left(\frac{\hat{\lambda}_i - \lambda_j}{\sqrt{\sigma^2/T + B_N^2}}\right)\frac{1}{B_N}\phi\left(\frac{y_{i0}-y_{j0}}{B_N}\right)\pi(\lambda_j, y_{j0})d\lambda_j dy_{j0} \\
&= \int \frac{1}{\sqrt{\sigma^2/T + B_N^2}}\phi\left(\frac{\hat{\lambda}_i - \lambda_j}{\sqrt{\sigma^2/T + B_N^2}}\right)\left[\int \frac{1}{B_N}\phi\left(\frac{y_{i0}-y_{j0}}{B_N}\right)\pi(y_{j0}|\lambda_j)dy_{j0}\right]\pi(\lambda_j)d\lambda_j.
\end{aligned}
$$

Now re-label λ_j and λ_i and y_{j0} as $\tilde{y}_{i0}$ to obtain:

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

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

$$
\mu\big(\lambda, \omega^2, p(\lambda, y_0)\big) = \lambda + \omega^2\frac{\partial \ln p(\lambda, y_0)}{\partial \lambda}. \tag{A.4}
$$

Using the previously developed notation, we expand the prediction error due to parameter estimation as follows:

$$\begin{aligned}
&\widehat{Y}_{iT+1} - \lambda_i - \rho Y_{iT} \\
&= \left[\mu\big(\hat{\lambda}_i(\hat{\rho}), \hat{\sigma}^2/T + B_N^2, \hat{p}^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0})\big)\right]^{C_N} - m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) \\
&\quad + m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) - \lambda_i \\
&\quad + (\hat{\rho} - \rho) Y_{iT} \\
&= A_{1i} + A_{2i} + A_{3i}, \text{ say.}
\end{aligned}$$

Now write

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

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

$$\begin{aligned}
&\text{(i)} \quad N\mathbb{E}^{\mathcal{Y}^N}_{\theta,\pi}\left[A_{1i}^2\right] = o_{u.\pi}(N^{\epsilon_0}), \\
&\text{(ii)} \quad \limsup_{N\longrightarrow\infty} \sup_{\pi\in\Pi} \frac{N\mathbb{E}^{\mathcal{Y}^N,\lambda_i}_{\theta,\pi}\left[A_{2i}^2\right]}{N\mathbb{E}^{\mathcal{Y}^i,\lambda_i}_{\theta,\pi}\left[(\lambda_i - \mathbb{E}^{\lambda_i}_{\theta,\pi,\mathcal{Y}^i}[\lambda_i])^2\right] + N^{\epsilon_0}} \leq 1, \\
&\text{(iii)} \quad N\mathbb{E}^{\mathcal{Y}^N}_{\theta,\pi}\left[A_{3i}^2\right] = o_{u.\pi}(N^{+}).
\end{aligned}$$

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

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

Even though we do not indicate this explicitly through our notation, we also restrict the domain of (λ, y_0) arguments that appear in numerous expressions throughout the proof to the support of the distribution of the random variables (λ_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 $\mathcal{T}_\lambda = \{|\lambda| \leq C_N\}$. From Assumption 3.2 we obtain for $C_N \geq M_3$ that

$$\begin{aligned} N^{1-\epsilon}\mathbb{P}(\mathcal{T}_\lambda^c) &\leq M_1 \exp\big((1-\epsilon)\ln N - M_2(C_N - M_3)\big) \\ &= \widetilde{M}_1 \exp\left(-M_2\left[C_N - \frac{1-\epsilon}{M_2}\ln N\right]\right) \\ &= o(1), \end{aligned} \tag{A.6}$$

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

$$N\mathbb{P}(\mathcal{T}_\lambda^c) = o_{u.\pi}(N^+).$$

2. Define the truncated region $\mathcal{T}_{Y_0} = \{\max_{1\leq i\leq N}|Y_{i0}| \leq L_N\}$. Then,

$$\begin{aligned} N^{1-\epsilon}\mathbb{P}(\mathcal{T}_{Y_0}^c) &= N^{1-\epsilon}\mathbb{P}\{\max_{1\leq i\leq N}|Y_{i0}| \geq L_N\} \\ &\leq N^{1-\epsilon}\sum_{i=1}^{N}\mathbb{P}\{|Y_{i0}| \geq L_N\} \\ &= N^{2-\epsilon}\int_{|y_0|\geq L_N}\pi(y_0)dy_0 \\ &\leq \widetilde{M}_1 \exp\left(-M_2\left[L_N - \frac{2-\epsilon}{M_2}\ln N\right]\right) \\ &= o(1), \end{aligned} \tag{A.7}$$

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

$$N\mathbb{P}(\mathcal{T}_{Y_0}^c) = o_{u.\pi}(N^+).$$

3. Define the truncated region $\mathcal{T}_{\hat{\rho}} = \{|\hat{\rho} - \rho| \leq 1/L_N^2\}$. By Chebyshev's inequality, Assumption 3.6, and (A.5), we can bound

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

4. Define the truncated region $\mathcal{T}_{\hat{\sigma}^2} = \{|\hat{\sigma}^2 - \sigma^2| \leq 1/L_N\}$. By Chebyshev's inequality,

Assumption 3.6, and (A.5), we can bound

$$N\mathbb{P}(\mathcal{T}_{\hat{\sigma}^2}^c) = N\mathbb{P}\{|\hat{\sigma}^2 - \sigma^2| > 1/L_N\} \leq L_N^2 \mathbb{E}[N(\hat{\sigma}^2 - \sigma^2)^2] = o_{u.\pi}(N^+). \tag{A.9}$$

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

$$\begin{aligned} N\mathbb{P}(\mathcal{T}_{\bar{U}}^c) &= N\mathbb{P}\{\max_{1\leq i\leq N} |\bar{U}_{i,-1}(\rho)| \geq L_N\} \\ &\leq N\sum_{i=1}^{N} \mathbb{P}\{|\bar{U}_{i,-1}(\rho)| \geq L_N\} = N^2\mathbb{P}\{|\bar{U}_{i,-1}(\rho)| \geq L_N\} \\ &\leq 2N^2 \exp\left(-\frac{L_N^2}{2\sigma_{\bar{U}}^2}\right) = 2\exp\left(-\frac{L_N^2}{2\sigma_{\bar{U}}^2} + 2\ln N\right) \\ &\leq 2\exp\left(-2\left(\frac{\ln N}{M_2^2 \sigma_{\bar{U}}^2} - 1\right)\ln N\right) \\ &= o_{u.\pi}(N^+), \end{aligned} \tag{A.10}$$

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 $C_1(\rho) = \frac{1}{T}\sum_{t=1}^{T}\rho^{t-1}$, $C_2(\rho) = \frac{1}{T}\sum_{t=2}^{T}(1 + \cdots + \rho^{t-2})$. 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 $\mathcal{T}_\lambda \cap \mathcal{T}_{Y_0} \cap \mathcal{T}_{\bar{U}}$:

$$\begin{aligned} \max_{1\leq i\leq N} |\bar{Y}_{i,-1}| &\leq |C_1(\rho)| \max_{1\leq i\leq N} |\lambda_i| + |C_2(\rho)| \max_{1\leq i\leq N} |Y_{i0}| + \max_{1\leq i\leq N} |\bar{U}_{i,-1}(\rho)| \\ &\leq M(C_N + L_N + L_N), \end{aligned}$$

which leads to

$$\max_{1\leq i,j\leq N} |\bar{Y}_{j,-1} - \bar{Y}_{i,-1}| \leq 2\max_{1\leq i\leq N} |\bar{Y}_{i,-1}| \leq 2M\left(C_N + 2L_N\right) = o_{u.\pi}(N^+). \tag{A.11}$$

7. For the region $\mathcal{T}_\lambda \cap \mathcal{T}_{Y_0} \cap \mathcal{T}_{\hat{\rho}} \cap \mathcal{T}_{\bar{U}}$ and with some finite constant M, we obtain the bound

$$\max_{1\leq i,j\leq N} |(\hat{\rho} - \rho)(\bar{Y}_{j,-1} - \bar{Y}_{i,-1})| \leq \frac{M\left(C_N + L_N\right)}{L_N^2} = o_{u.\pi}(N^+). \tag{A.12}$$

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

$$\begin{aligned} N\mathbb{P}(\mathcal{T}_m^c) &\leq \frac{1}{C_N^2} N\mathbb{E}(m(\hat{\lambda}_i, Y_{i0})^2 \mathcal{T}_m^c) \leq o_{u.\pi}(N^+) \\ N\mathbb{P}(\mathcal{T}_{m_*}^c) &\leq \frac{1}{C_N^2} N\mathbb{E}(m_*(\hat{\lambda}_i, Y_{i0})^2 \mathcal{T}_{m_*}^c) \leq o_{u.\pi}(N^+). \end{aligned} \tag{A.13}$$

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,*

$$N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[\left(\left[\mu\big(\hat{\lambda}_i(\hat{\rho}), \hat{\sigma}^2/T + B_N^2, \hat{p}^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0})\big)\right]^{C_N} - m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N)\right)^2\right] = o_{u.\pi}(N^{\epsilon_0}).$$

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

$$\begin{aligned} |A_{1i}|^2 &= \left(\left[\mu\big(\hat{\lambda}_i(\hat{\rho}), \hat{\sigma}^2/T + B_N^2, \hat{p}^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0})\big)\right]^{C_N} - m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N)\right)^2 \\ &\leq 2C_N^2 + 2m_*^2(\hat{\lambda}_i, y_{i0}; \pi, B_N) \\ &= 2C_N^2 + 2m_*^2(\hat{\lambda}_i, y_{i0}; \pi, B_N)\mathbb{I}(\mathcal{T}_{m*}) + 2m_*^2(\hat{\lambda}_i, y_{i0}; \pi, B_N)\mathbb{I}(\mathcal{T}_{m*}^c) \\ &\leq 4C_N^2 + 2m_*^2(\hat{\lambda}_i, y_{i0}; \pi, B_N)\mathbb{I}(\mathcal{T}_{m*}^c). \end{aligned} \tag{A.14}$$

Then,

$$\begin{aligned} N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[A_{1i}^2] &\leq N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[A_{1i}^2\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\,\mathcal{T}_{\hat{\rho}}\,\mathcal{T}_{\bar{U}}\,\mathcal{T}_{Y_0}\,\mathcal{T}_{m*}\,\mathcal{T}_\lambda) \qquad (A.15) \\ &\quad + N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[A_{1i}^2\big(\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}^c) + \mathbb{I}(\mathcal{T}_{\hat{\rho}}^c) + \mathbb{I}(\mathcal{T}_{\bar{U}}^c) + \mathbb{I}(\mathcal{T}_{Y_0}^c) + \mathbb{I}(\mathcal{T}_{m*}^c) + \mathbb{I}(\mathcal{T}_\lambda^c)\big)\right] \\ &\leq N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[A_{1i}^2\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\,\mathcal{T}_{\hat{\rho}}\,\mathcal{T}_{\bar{U}}\,\mathcal{T}_{Y_0}\,\mathcal{T}_{m*}\,\mathcal{T}_\lambda)] \\ &\quad + 4C_N^2 N\big(\mathbb{P}(\mathcal{T}_{\hat{\sigma}^2}^c) + \mathbb{P}(\mathcal{T}_{\hat{\rho}}^c) + \mathbb{P}(\mathcal{T}_{\bar{U}}^c) + \mathbb{P}(\mathcal{T}_{Y_0}^c) + \mathbb{P}(\mathcal{T}_{m*}^c) + \mathbb{P}(\mathcal{T}_\lambda^c)\big) \\ &\quad + 12N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[m_*^2(\hat{\lambda}_i, y_{i0}; \pi, B_N)\mathbb{I}(\mathcal{T}_{m*}^c)] \\ &= N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[A_{1i}^2\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\,\mathcal{T}_{\hat{\rho}}\,\mathcal{T}_{\bar{U}}\,\mathcal{T}_{Y_0}\,\mathcal{T}_{m*}\,\mathcal{T}_\lambda)] + o_{u.\pi}(N^+) + o_{u.\pi}(N^+). \end{aligned}$$

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:

$$\begin{aligned}\mathcal{T}_{\hat{\lambda}Y_0} &= \big\{(\hat{\lambda}_i, Y_{i0}) \,\big|\, -C'_N \leq \hat{\lambda}_i \leq C'_N,\ -C'_N \leq Y_{i0} \leq C'_N\big\} \\ \mathcal{T}_{p(\cdot)} &= \left\{(\hat{\lambda}_i, Y_{i0}) \,\middle|\, p(\hat{\lambda}_i, Y_{i0}) \geq \frac{N^{\epsilon}}{N}\right\},\end{aligned}$$

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

Notice that since $C_N = o(N^+)$ and $\sqrt{\ln N} = o(N^+)$,

$$C'_N = o(N^+). \tag{A.16}$$

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

$$\begin{aligned}N\mathbb{E}^{\mathcal{Y}^N}_{\theta,\pi}[A^2_{1i}\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\ \mathcal{T}^c_{p(\cdot)})] &\leq o_{u.\pi}(N^{\epsilon_0}) &(\text{A.17})\\ N\mathbb{E}^{\mathcal{Y}^N}_{\theta,\pi}[A^2_{1i}\mathbb{I}(\mathcal{T}_{\lambda}\mathcal{T}^c_{\hat{\lambda}Y_0})] &\leq o_{u.\pi}(N^+). &(\text{A.18})\end{aligned}$$

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:

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

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 $2Nm_*^2(\hat{\lambda}_i, y_{i0}; \pi, B_N)\mathbb{I}(\mathcal{T}_{m*}^c)$ to the expected value of A_{1i}^2 by $o_{u.\pi}(N^+)$.

$$
\begin{aligned}
& N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[A_{1i}^2\mathbb{I}(\mathcal{T}_\lambda\mathcal{T}_{\hat{\lambda}Y_0}^c))\right] \\
\leq\ & 4NC_N^2\mathbb{P}(\mathcal{T}_\lambda \cap \mathcal{T}_{\hat{\lambda}Y_0}^c) + o_{u.\pi}(N^+) \\
=\ & 4NC_N^2\int_{\mathcal{T}_{\hat{\lambda}Y_0}^c}\left[\int_{\lambda_i}\frac{1}{\sigma/\sqrt{T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)\pi(\lambda_i, y_{i0})d\lambda_i\right]d(\hat{\lambda}_i, y_{i0}) + o_{u.\pi}(N^+) \\
\leq\ & 4NC_N^2\int_{\lambda_i}\int_{|\hat{\lambda}_i|>C'_N}\left[\int_{y_{i0}}\frac{1}{\sigma/\sqrt{T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)\pi(y_{i0}|\lambda_i)dy_{i0}\right]\pi(\lambda_i)d(\hat{\lambda}_i, \lambda_i) \\
& +4NC_N^2\int_{\lambda_i}\int_{|y_{i0}|>C'_N}\left[\int_{\hat{\lambda}_i}\frac{1}{\sigma/\sqrt{T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)d\hat{\lambda}_i\right]\pi(\lambda_i, y_{i0})d(\lambda_i, y_{i0}) + o_{u.\pi}(N^+) \\
=\ & 4NC_N^2\int_{|\lambda_i|\leq C_N}\left[\int_{|\hat{\lambda}_i|>C'_N}\frac{1}{\sigma/\sqrt{T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)d\hat{\lambda}_i\right]\pi(\lambda_i)d\lambda_i \\
& +4NC_N^2\int_{|y_{i0}|>C'_N}\left[\int_{\lambda_i}\pi(\lambda_i|y_{i0})d\lambda_i\right]\pi(y_{i0})dy_{i0} \\
& +o_{u.\pi}(N^+) \\
\leq\ & 4NC_N^2\int_{|\lambda_i|\leq C_N}\left[\int_{|\hat{\lambda}_i|>C'_N}\frac{1}{\sigma/\sqrt{T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)d\hat{\lambda}_i\right]\pi(\lambda_i)d\lambda_i \\
& +4NC_N^2\int_{|y_{i0}|>C'_N}\pi(y_{i0})dy_{i0} \\
& +o_{u.\pi}(N^+) \\
=\ & B_1 + o_{u.\pi}(N^+) + o_{u.\pi}(N^+), \quad \text{say.}
\end{aligned}
$$

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

$$\begin{aligned}
&\int_{|\hat{\lambda}_i|>C'_N} \frac{1}{\sigma/\sqrt{T}} \phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right) d\hat{\lambda}_i \\
&= \int_{-\infty}^{-\sqrt{T}(C'_N+\lambda_i)/\sigma} \phi(\tilde{\lambda}_i)d\tilde{\lambda}_i + \int_{\sqrt{T}(C'_N-\lambda_i)/\sigma}^{\infty} \phi(\tilde{\lambda}_i)d\tilde{\lambda}_i \\
&\leq \int_{-\infty}^{-\sqrt{T}(C'_N-|\lambda_i|)/\sigma} \phi(\tilde{\lambda}_i)d\tilde{\lambda}_i + \int_{\sqrt{T}(C'_N-|\lambda_i|)/\sigma}^{\infty} \phi(\tilde{\lambda}_i)d\tilde{\lambda}_i \\
&\leq 2\int_{\sqrt{T}(C'_N-|\lambda_i|)/\sigma}^{\infty} \phi(\tilde{\lambda}_i)d\tilde{\lambda}_i \\
&\leq 2\frac{\phi\big(\sqrt{T}(C'_N-|\lambda_i|)/\sigma\big)}{\sqrt{T}(C'_N-|\lambda_i|)/\sigma},
\end{aligned}$$

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

$$\begin{aligned}
B_1 &\leq 4NC_N^2 \int_{|\lambda_i|<C_N} \frac{\phi\big(\sqrt{T}(C'_N-|\lambda_i|)/\sigma\big)}{\sqrt{T}(C'_N-|\lambda_i|)/\sigma} \pi(\lambda_i)d\lambda_i \\
&\leq 4NC_N^2 \int_{|\lambda_i|<C_N} \phi\big(\sqrt{2\ln N}\big)\pi(\lambda_i)d\lambda_i \\
&\leq 4NC_N^2 \exp\big(-\ln N\big) \int_{|\lambda_i|<C_N} \pi(\lambda_i)d\lambda_i \\
&\leq 4C_N^2 \\
&= o_{u.\pi}(N^+).
\end{aligned}$$

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

Step 2. It remains to be shown that

$$N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[A_{1i}^2\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\ \mathcal{T}_{\hat{\rho}}\ \mathcal{T}_{\bar{U}}\ \mathcal{T}_{Y_0}\ \mathcal{T}_{m*}\ \mathcal{T}_{\lambda}\ \mathcal{T}_{\hat{\lambda}Y_0}\ \mathcal{T}_{p(\cdot)})] \leq o_{u.\pi}(N^+). \tag{A.19}$$

We introduce the following notation:

$$\begin{aligned}
\widetilde{p}_i^{(-i)} &= \hat{p}^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0}) \\
d\widetilde{p}_i^{(-i)} &= \frac{1}{\partial\hat{\lambda}_i(\hat{\rho})}\partial\hat{p}^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0}) \\
\hat{p}_i^{(-i)} &= \hat{p}^{(-i)}(\hat{\lambda}_i(\rho), Y_{i0}) \\
d\hat{p}_i^{(-i)} &= \frac{1}{\partial\hat{\lambda}_i(\rho)}\partial\hat{p}^{-i}(\hat{\lambda}_i(\rho), Y_{i0}) \\
p_i &= p(\hat{\lambda}_i(\rho), Y_{i0}) \\
p_{*i} &= p_*(\hat{\lambda}_i(\rho), Y_{i0}) \\
dp_{*i} &= \frac{1}{\partial\hat{\lambda}_i(\rho)}\partial p_*(\hat{\lambda}_i(\rho), Y_{i0}).
\end{aligned} \tag{A.20}$$

Moreover, we introduce another truncation

$$\mathcal{T}_{\tilde{p}(\cdot)} = \left\{ (\hat{\lambda}_i, Y_{i0}) \,\middle|\, \widetilde{p}_i^{(-i)} > \frac{p_{*i}}{2} \right\}. \tag{A.21}$$

On the set $\mathcal{T}_{m*}$, we have $|m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N)| \leq C_N$, and so

$$|A_{1i}| \leq 2C_N. \tag{A.22}$$

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

$$N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[A_{1i}^2 \mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\ \mathcal{T}_{\hat{\rho}}\ \mathcal{T}_{\bar{U}}\ \mathcal{T}_{Y_0}\ \mathcal{T}_{m*}\ \mathcal{T}_{\lambda}\ \mathcal{T}_{\hat{\lambda}Y_0}\ \mathcal{T}_{p(\cdot)})\ \mathcal{T}_{\tilde{p}(\cdot)}] \leq o_{u.\pi}(N^+) \tag{A.23}$$

$$N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}[A_{1i}^2 \mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\ \mathcal{T}_{\hat{\rho}}\ \mathcal{T}_{\bar{U}}\ \mathcal{T}_{Y_0}\ \mathcal{T}_{m*}\ \mathcal{T}_{\lambda}\ \mathcal{T}_{\hat{\lambda}Y_0}\ \mathcal{T}_{p(\cdot)})\ \mathcal{T}_{\tilde{p}(\cdot)}^c] \leq o_{u.\pi}(N^+). \tag{A.24}$$

Step 2.1. Using the triangle inequality, we obtain

$$\begin{aligned} |A_{1i}| &= \left| \left[\mu\big(\hat{\lambda}_i(\hat{\rho}), Y_{i0}, \hat{\sigma}^2/T + B_N^2, \hat{p}^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0})\big)\right]^{C_N} - m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) \right| \\ &\leq \left| \mu\big(\hat{\lambda}_i(\hat{\rho}), Y_{i0}, \hat{\sigma}^2/T + B_N^2, \hat{p}^{(-i)}(\hat{\lambda}_i(\hat{\rho}), Y_{i0})\big) - m_*(\hat{\lambda}_i, y_{i0}; \pi, B_N) \right| \\ &= \left| \hat{\lambda}_i(\hat{\rho}) - \lambda_i(\rho) + \left(\frac{\hat{\sigma}^2}{T} - \frac{\sigma^2}{T}\right)\frac{dp_{*i}}{p_{*i}} + \left(\frac{\hat{\sigma}^2}{T} + B_N^2\right)\left(\frac{d\widetilde{p}_i^{(-i)}}{\widetilde{p}_i^{(-i)}} - \frac{dp_{*i}}{p_{*i}}\right) \right| \\ &\leq |\hat{\rho} - \rho| \big|\bar{Y}_{i,-1}\big| + \left|\frac{\hat{\sigma}^2}{T} - \frac{\sigma^2}{T}\right| \left|\frac{dp_{*i}}{p_{*i}}\right| + \left(\frac{\hat{\sigma}^2}{T} + B_N^2\right)\left|\frac{d\widetilde{p}_i^{(-i)}}{\widetilde{p}_i^{(-i)}} - \frac{dp_{*i}}{p_{*i}}\right|, \\ &= A_{11i} + A_{12i} + A_{13i}, \quad \text{say.} \end{aligned}$$

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

$$N\mathbb{E}_\theta^{\mathcal{Y}^N}\left[A_{1ji}^2 \mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\ \mathcal{T}_{\hat{\rho}}\ \mathcal{T}_{\bar{U}}\ \mathcal{T}_{Y_0}\ \mathcal{T}_{m*}\ \mathcal{T}_{\lambda}\ \mathcal{T}_{\hat{\lambda}Y_0}\ \mathcal{T}_{p(\cdot)}\ \mathcal{T}_{\tilde{p}(\cdot)})\right] \leq o_{u.\pi}(N^+), \quad j = 1, 2, 3.$$

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

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

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

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

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

$$N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[A_{12i}^2 \mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0})\mathbb{I}(\mathcal{T}_{m*})\right] \leq \frac{1}{(\sigma^2/T + B_N^2)^2}\mathbb{E}_\theta^{\mathcal{Y}^N}\left[N(\hat{\sigma}^2 - \sigma^2)^2\right](C_N' + C_N)^2 = o_{u.\pi}(N^+).$$

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

$$A_{13i}^2 \mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}) \leq \left(\frac{\sigma^2}{T} + B_N^2 + \frac{1}{T}\frac{1}{L_N}\right)^2 \left(\frac{d\widetilde{p}_i^{(-i)}}{\widetilde{p}_i^{(-i)}} - \frac{dp_{*i}}{p_{*i}}\right)^2.$$

Thus, the desired result follows if we show

$$N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[\left(\frac{d\widetilde{p}_i^{(-i)}}{\widetilde{p}_i^{(-i)}}-\frac{dp_{*i}}{p_{*i}}\right)^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\ \mathcal{T}_{\bar{U}}\ \mathcal{T}_{Y_0}\ \mathcal{T}_{m*}\ \mathcal{T}_{\lambda}\ \mathcal{T}_{\hat{\lambda}Y_0}\ \mathcal{T}_{p(\cdot)})\right]=o_{u.\pi}(N^+) \tag{A.26}$$

Decompose

$$\frac{d\widetilde{p}_i^{(-i)}}{\widetilde{p}_i^{(-i)}}-\frac{dp_{*i}}{p_{*i}}=\frac{d\widetilde{p}_i^{(-i)}-dp_{*i}}{\widetilde{p}_i^{(-i)}-p_{*i}+p_{*i}}-\frac{dp_{*i}}{p_{*i}}\left(\frac{\widetilde{p}_i^{(-i)}-p_{*i}}{\widetilde{p}_i^{(-i)}-p_{*i}+p_{*i}}\right).$$

Using the C_r inequality, we obtain

$$\begin{aligned}
&N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[\left(\frac{d\widetilde{p}_i^{(-i)}}{\widetilde{p}_i^{(-i)}}-\frac{dp_{*i}}{p_{*i}}\right)^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\\
\leq\ &2N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[\left(\frac{d\widetilde{p}_i^{(-i)}-dp_{*i}}{\widetilde{p}_i^{(-i)}-p_{*i}+p_{*i}}\right)^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\\
&+2N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[\left(\frac{dp_{*i}}{p_{*i}}\right)^2\left(\frac{\widetilde{p}_i^{(-i)}-p_{*i}}{\widetilde{p}_i^{(-i)}-p_{*i}+p_{*i}}\right)^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\\
=\ &2\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[\left(\frac{\sqrt{N}(d\widetilde{p}_i^{(-i)}-dp_{*i})}{\widetilde{p}_i^{(-i)}-p_{*i}+p_{*i}}\right)^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\\
&+2o_{u.\pi}(N^+)\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^N}\left[\left(\frac{\sqrt{N}(\widetilde{p}_i^{(-i)}-p_{*i})}{\widetilde{p}_i^{(-i)}-p_{*i}+p_{*i}}\right)^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\\
=\ &2B_{1i}+2o_{u.\pi}(N^+)B_{2i},
\end{aligned}$$

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

$$\begin{aligned}
\sqrt{N}(d\widetilde{p}_i^{(-i)}-dp_{*i}) &= \sqrt{N}(d\hat{p}_i^{(-i)}-dp_{*i})+\sqrt{N}(\hat{\rho}-\rho)R_{1i}(\widetilde{\rho})\\
\sqrt{N}(\widetilde{p}_i^{(-i)}-p_{*i}) &= \sqrt{N}(\hat{p}_i^{(-i)}-p_{*i})+\sqrt{N}(\hat{\rho}-\rho)R_{2i}(\widetilde{\rho}),
\end{aligned}$$

where

$$\begin{aligned}
R_{1i}(\rho) &= -\frac{1}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)^2\left(\bar{Y}_{j,-1}-\bar{Y}_{i,-1}\right)\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right)\\
&\quad +\frac{1}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N^2}\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\left(\bar{Y}_{j,-1}-\bar{Y}_{i,-1}\right)\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right),\\
R_{2i}(\rho) &= \frac{1}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\left(\bar{Y}_{j,-1}-\bar{Y}_{i,-1}\right)\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right),
\end{aligned}$$

and $\widetilde{\rho}$ is located between $\hat{\rho}$ and ρ.

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

$$\begin{aligned}
B_{2i} &\leq 4\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\frac{1}{p_{*i}^2}\mathbb{E}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left[N(\hat{p}_i^{(-i)}-p_{*i})^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\right]\\
&\quad +4\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\frac{1}{p_{*i}^2}\mathbb{E}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left[N(\hat{\rho}-\rho)^2R_{2i}^2(\tilde{\rho})\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\right]\\
&= 4\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}[B_{21i}+B_{22i}],
\end{aligned}$$

say.

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

$$\mathbb{E}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left[N(\hat{p}_i^{(-i)}-p_{*i})^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\leq\frac{M}{B_N^2}p_i\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}).$$

This leads to

$$\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}[B_{21i}]\leq\frac{M}{B_N^2}\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\frac{p_i}{p_{*i}^2}\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right]=\frac{M}{B_N^2}\int_{\mathcal{T}_{\hat{\lambda}Y_0}\cap\mathcal{T}_{p(\cdot)}}\frac{p_i^2}{p_{*i}^2}d\hat{\lambda}_i dy_{i0}.$$

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

$$\int_{\mathcal{T}_{\hat{\lambda}Y_0}\cap\mathcal{T}_{p(\cdot)}}\frac{p_i^2}{p_{*i}^2}d\hat{\lambda}_i dy_{i0}=o_{u.\pi}(N^+).$$

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

$$\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}[B_{21i}]\leq o_{u.\pi}(N^+).$$

Using the Cauchy-Schwarz Inequality, we obtain

$$B_{22i} \leq \frac{1}{p_{*i}^2}\sqrt{\mathbb{E}^{\mathcal{Y}(-i)}_{\theta,\pi,\mathcal{Y}^i}\left[N^2(\hat{\rho}-\rho)^4\right]}\sqrt{\mathbb{E}^{\mathcal{Y}(-i)}_{\theta,\pi,\mathcal{Y}^i}\left[R^4_{2i}(\tilde{\rho})\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]}.$$

Using the inequality once more leads to

$$\begin{aligned}
\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}[B_{22i}] &\leq \sqrt{\mathbb{E}^{\mathcal{Y}^N}_{\theta,\pi}\left[N^2(\hat{\rho}-\rho)^4\right]}\sqrt{\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\frac{1}{p_{*i}^4}\mathbb{E}^{\mathcal{Y}(-i)}_{\theta,\mathcal{Y}^i}\left[R^4_{2i}(\tilde{\rho})\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\right]} \\
&\leq o_{u.\pi}(N^+)\sqrt{\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\frac{1}{p_{*i}^4}\mathbb{E}^{\mathcal{Y}(-i)}_{\theta,\pi,\mathcal{Y}^i}\left[R^4_{2i}(\tilde{\rho})\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right]\right]}.
\end{aligned}$$

The second inequality follows from Assumption 3.6.

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

$$\mathbb{E}^{\mathcal{Y}(-i)}_{\theta,\pi,\mathcal{Y}^i}\left[R^4_{2i}(\tilde{\rho})\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{m*}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\right] \leq ML^4_N p_i^4\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}),$$

where $L_N = o(N^+)$ was defined in (A.5). This leads to the bound

$$\begin{aligned}
\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}[B_{22i}] &\leq o_{u.\pi}(N^+)ML^2_N\sqrt{\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\left(\frac{p_i}{p_{*i}}\right)^4\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right]} \\
&= o_{u.\pi}(N^+)ML^2_N\sqrt{\int_{\mathcal{T}_{\hat{\lambda}Y_0}\cap\mathcal{T}_{p(\cdot)}}\left(\frac{p_i}{p_{*i}}\right)^4 p_i d\hat{\lambda}_i dy_{i0}} \\
&\leq o_{u.\pi}(N^+)M_*L^2_N\sqrt{\int_{\mathcal{T}_{\hat{\lambda}Y_0}\cap\mathcal{T}_{p(\cdot)}}\left(\frac{p_i}{p_{*i}}\right)^4 d\hat{\lambda}_i dy_{i0}} \\
&\leq o_{u.\pi}(N^+).
\end{aligned}$$

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 $\mathcal{T}_{m^*}$,

$$|A_{1i}| \leq 2C_N = o_{u.\pi}(N^+).$$

Then,

$$
\begin{aligned}
& N\mathbb{E}^{\mathcal{Y}^N}_{\theta,\pi}[A^2_{1i}\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\ \mathcal{T}_{\hat{\rho}}\ \mathcal{T}_{\bar{U}}\ \mathcal{T}_{Y_0}\ \mathcal{T}_{m*}\ \mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\ \mathcal{T}_{p(\cdot)})\ \mathcal{T}^c_{\tilde{p}(\cdot)}] \\
& \quad \leq \ o_{u.\pi}(N^+)N\mathbb{P}^{\mathcal{Y}^N}_{\theta,\pi}(\mathcal{T}_{\hat{\sigma}^2}\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}^c_{\tilde{p}(\cdot)}).
\end{aligned}
$$

Notice that

$$
\begin{aligned}
\mathcal{T}^c_{\tilde{p}(\cdot)} &= \left\{\hat{p}_i^{(-i)} - p_{*i} + (\hat{\rho}-\rho)R_{2i}(\widetilde{\rho}) < -\frac{p_{*i}}{2}\right\} \\
&\subset \left\{\hat{p}_i^{(-i)} - p_{*i} - |\hat{\rho}-\rho||R_{2i}(\widetilde{\rho})| < -\frac{p_{*i}}{2}\right\} \\
&\subset \left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\} \cup \left\{|\hat{\rho}-\rho||R_{2i}(\widetilde{\rho})| > \frac{p_{*i}}{4}\right\}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
& N\mathbb{P}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}(\mathcal{T}_{\hat{\sigma}^2}\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}^c_{\tilde{p}(\cdot)}) \\
& \leq \ N\mathbb{P}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\} \\
& \qquad + N\mathbb{P}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left[\left\{|\hat{\rho}-\rho||R_{2i}(\widetilde{\rho})| > \frac{p_{*i}}{4}\right\}\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right] \\
& \leq \ N\mathbb{P}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\} + \frac{16L_N^4}{p_{*i}^2}\mathbb{E}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left[R_{2i}(\widetilde{\rho})^2\mathbb{I}(\mathcal{T}_{\hat{\sigma}^2}\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\lambda}\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right] \\
& \leq \ N\mathbb{P}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\} + \frac{ML_N^6}{p_{*i}^2}p_i^2\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}).
\end{aligned}
$$

The first inequality is based on the superset of $\mathcal{T}^c_{\tilde{p}(\cdot)}$ from above. The second inequality is based on Chebychev's inequality and truncation $\mathcal{T}_{\hat{\rho}}$. 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:

$$
p_i = \int p(\hat{\lambda}|\lambda)\pi(Y_{i0}|\lambda)\pi(\lambda)d\lambda \leq \widetilde{M}\int \pi(\lambda)d\lambda = \widetilde{M} < \infty, \tag{A.27}
$$

because $p(\hat{\lambda}|\lambda)$ 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 probability $\mathbb{P}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\}$ uniformly over $(\hat{\lambda}_i, Y_{i0})$ in the region $\mathcal{T}_{\hat{\lambda}Y_0}$, showing that

$$N\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\mathbb{P}^{\mathcal{Y}^{(-i)}}_{\theta,\pi,\mathcal{Y}^i}\left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\}\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right] = o_{u.\pi}(N^+),$$

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

$$\mathbb{E}^{\mathcal{Y}^i}_{\theta,\pi}\left[\frac{p_i}{p_{*i}^2}\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right] = \int_{\mathcal{T}_{\hat{\lambda}Y_0}\cap\mathcal{T}_{p(\cdot)}}\left(\frac{p_i}{p_{*i}}\right)^2 d\hat{\lambda}_i dy_{i0} = o_{u.\pi}(N^+),$$

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 $\epsilon_0 > 0$,*

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

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

$$\mathbb{E}^{\mathcal{Y}^i,\lambda_i}_{*,\theta,\pi}[\cdot] = \int[\cdot]p_*(\hat{\lambda}|\lambda)\pi_*(y_0|\lambda)\pi(\lambda)d(\hat{\lambda},\lambda,y_0),$$

where

$$p_*(\hat{\lambda}|\lambda) = \frac{1}{\sqrt{\sigma^2/T + B_N^2}}\phi\left(\frac{\hat{\lambda}-\lambda}{\sqrt{\sigma^2/T + B_N^2}}\right)$$
$$\pi_*(y_0|\lambda) = \int\frac{1}{B_N}\phi\left(\frac{y_0 - \tilde{y}_0}{B_N}\right)\pi(\tilde{y}_0|\lambda)d\tilde{y}_0.$$

The desired result follows if we can show that

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

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

Part (i): Notice that the denominator is bounded below by N^{ϵ_0}. We will proceed by constructing an upper bound for the numerator. Using the fact that the posterior mean minimizes the integrated risk, we obtain

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

say.

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

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

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

$$\begin{aligned}\Big(m(\hat{\lambda}_i, y_{i0}; \pi) - \lambda_i\Big)^2 &\leq 2m^2(\hat{\lambda}_i, y_{i0}; \pi)\big(\mathbb{I}(\mathcal{T}_m) + \mathbb{I}(\mathcal{T}_m^c)\big) + 2\lambda_i^2\big(\mathbb{I}(\mathcal{T}_\lambda) + \mathbb{I}(\mathcal{T}_\lambda^c)\big) \\ &\leq 4C_N^2 + 2m(\hat{\lambda}_i, y_{i0}; \pi)^2\mathbb{I}(\mathcal{T}_m^c) + 2\lambda_i^2\mathbb{I}(\mathcal{T}_\lambda^c).\end{aligned}$$

Thus,

$$\begin{aligned}B_{2i} &\leq 4C_N^2 N\big(\mathbb{P}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}^c) + \mathbb{P}(\mathcal{T}_{\hat{\lambda}Y_0}^c) + \mathbb{P}(\mathcal{T}_m^c) + \mathbb{P}(\mathcal{T}_\lambda^c)\big) \\ &\quad + 8N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^i,\lambda_i}\big[m(\hat{\lambda}_i, y_{i0}; \pi)^2\mathbb{I}(\mathcal{T}_m^c)\big] + 8N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^i,\lambda_i}\big[\lambda_i^2\mathbb{I}(\mathcal{T}_\lambda^c)\big].\end{aligned}$$

Notice that $C_N^2 = o_{u.\pi}(N^+)$, $N\mathbb{P}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}^c) = o_{u.\pi}(N^{\epsilon_0})$ (see Step 1.1), $N\mathbb{P}(\mathcal{T}_{\hat{\lambda}Y_0}^c) = o_{u.\pi}(N^+)$ (see Step 1.2), and $N\mathbb{P}(\mathcal{T}_m^c) = o_{u.\pi}(N^+)$ (see Truncation 1). Also, notice that

$$\begin{aligned}N\mathbb{E}_{\theta,\pi}^{\mathcal{Y}^i,\lambda_i}\big[\lambda_i^2\mathbb{I}(\mathcal{T}_\lambda^c)\big] &= N\int_{|\lambda|>C_N}\lambda^2\pi(\lambda)d\lambda \\ &\leq \sqrt{\int_{|\lambda|>C_N}\lambda^4\pi(\lambda)d\lambda}\sqrt{N^2\int_{|\lambda|>C_N}\pi(\lambda)d\lambda} \\ &\leq M\sqrt{N^2\int_{|\lambda|>C_N}\pi(\lambda)d\lambda} \\ &= o_{u.\pi}(N^+).\end{aligned}$$

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,

$$B_{2i} \leq o_{u.\pi}(N^{\epsilon_0}).$$

Combining the bounds for B_{1i} and B_{2i}, we have

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

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

$$\begin{aligned}
&N\mathbb{E}^{\mathcal{Y}_i,\lambda_i}_{\theta,\pi}\left[\left(m_*(\hat{\lambda}_i, y_{i0};\pi,B_N)-\lambda_i\right)^2\right]\\
&\leq \quad N\mathbb{E}^{\mathcal{Y}_i,\lambda_i}_{\theta,\pi}\left[\left(m_*(\hat{\lambda}_i, y_{i0};\pi,B_N)-\lambda_i\right)^2 \mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{m*}\mathcal{T}_{\lambda})\right]\\
&\quad\quad +N\mathbb{E}^{\mathcal{Y}_i,\lambda_i}_{\theta}\left[\left(m_*(\hat{\lambda}_i, y_{i0};\pi,B_N)-\lambda_i\right)^2\left(\mathbb{I}(\mathcal{T}^c_{\hat{\lambda}Y_0})+\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}^c_{p(\cdot)})+\mathbb{I}(\mathcal{T}^c_{m*})+\mathbb{I}(\mathcal{T}^c_{\lambda})\right)\right]\\
&= \quad B_{1i}+B_{2i},
\end{aligned}$$

say. Now consider the term B_{1i}:

$$\begin{aligned}
B_{1i} &= N\int\int\int\left(m_*(\hat{\lambda}_i, y_{i0};\pi,B_N)-\lambda_i\right)^2 p_*(\hat{\lambda}_i|\lambda_i)\pi_*(y_{i0}|\lambda_i)\frac{p(\hat{\lambda}_i|\lambda_i)\pi(y_{i0}|\lambda_i)}{p_*(\hat{\lambda}_i|\lambda_i)\pi_*(y_{i0}|\lambda_i)}\pi(\lambda_i)\\
&\qquad \times\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{m*}\mathcal{T}_{\lambda})d(\hat{\lambda}_i,\lambda_i,y_{i0})\\
&= (1+o(1))N\int\int\int\left(m_*(\hat{\lambda}_i, y_{i0};\pi,B_N)-\lambda_i\right)^2 p_*(\hat{\lambda}_i|\lambda_i)\pi_*(y_{i0}|\lambda_i)\pi(\lambda_i)\\
&\qquad \times\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{m*}\mathcal{T}_{\lambda})d(\hat{\lambda}_i,d\lambda_i,dy_{i0})\\
&\leq (1+o(1))N\mathbb{E}^{\mathcal{Y}_i,\lambda_i}_{*,\theta,\pi}\left[\left(m_*(\hat{\lambda}_i, y_{i0};\pi,B_N)-\lambda_i\right)^2\right],
\end{aligned}$$

where the $o(1)$ term is uniform in $\pi \in \Pi$. Using the a similar argument as in Part (i) we can establish that $B_{2i} = o_{u.\pi}(N^{\epsilon_0})$, 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,*

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

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

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

By Assumption 3.6, we have

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

For the second term, write

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

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

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

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

$$\begin{aligned} R_{1i}(\rho) &= -\frac{1}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)^2(\bar{Y}_{j,-1}-\bar{Y}_{i,-1})\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right) \\ &\quad +\frac{1}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N^2}\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)(\bar{Y}_{j,-1}-\bar{Y}_{i,-1})\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right) \\ R_{2i}(\rho) &= \frac{1}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)(\bar{Y}_{j,-1}-\bar{Y}_{i,-1})\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right) \end{aligned}$$

For expositional purposes, our analysis focuses on the slightly simpler term $R_{2i}(\widetilde{\rho})$. The extension to $R_{1i}(\widetilde{\rho})$ is fairly straightforward. By definition,

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

Therefore,

$$\begin{aligned} R_{2i}(\widetilde{\rho}) &= \frac{1}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}-(\widetilde{\rho}-\rho)\left(\frac{\bar{Y}_{j,-1}-\bar{Y}_{i,-1}}{B_N}\right)\right) \\ &\quad \times\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}-(\widetilde{\rho}-\rho)\left(\frac{\bar{Y}_{j,-1}-\bar{Y}_{i,-1}}{B_N}\right)\right) \\ &\quad \times(\bar{Y}_{j,-1}-\bar{Y}_{i,-1})\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right). \end{aligned}$$

Consider the region $\mathcal{T}_{\hat{\rho}}\cap\mathcal{T}_{\bar{U}}\cap\mathcal{T}_{Y_0}$. First, using (A.12) we can bound

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

Thus,

$$\begin{aligned}
&\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}-(\widetilde{\rho}-\rho)\left(\frac{\bar{Y}_{j,-1}-\bar{Y}_{i,-1}}{B_N}\right)\right)\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}) \\
\leq\ & \phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}+\left(\frac{M}{L_N B_N}\right)\right)\mathbb{I}\left\{\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\leq-\frac{M}{L_N B_N}\right\} \\
&+\phi(0)\mathbb{I}\left\{\left|\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right|\leq\frac{M}{L_N B_N}\right\} \\
&+\phi\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}-\left(\frac{M}{L_N B_N}\right)\right)\mathbb{I}\left\{\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\geq\frac{M}{L_N B_N}\right\} \\
=\ & \bar{\phi}\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right),
\end{aligned}$$

say. The function $\bar{\phi}(x)$ 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

$$\left|\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}-(\widetilde{\rho}-\rho)\left(\frac{\bar{Y}_{j,-1}-\bar{Y}_{i,-1}}{B_N}\right)\right|\leq\left|\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right|+\frac{M}{L_N B_N}.$$

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

$$\max_{1\leq i,j\leq N}|\bar{Y}_{j,-1}-\bar{Y}_{i,-1}|\leq ML_N.$$

Therefore,

$$\left|\bar{Y}_{j,-1}-\bar{Y}_{i,-1}\right|\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right)\leq\frac{ML_N}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right).$$

Now, define the function

$$\bar{\phi}_*(x)=\bar{\phi}\left(x\right)\left(|x|+\frac{M}{L_N B_N}\right).$$

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 $\bar{\phi}_*(x)$ has the property that for any finite positive integer m there is a finite constant M such that

$$\int\bar{\phi}_*(x)^m dx\leq M.$$

Combining the previous results we obtain the following bound for $R_{2i}(\widetilde{\rho})$:

$$\left|R_{2i}(\widetilde{\rho})\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0})\right| \leq \frac{ML_N}{N-1}\sum_{j\neq i}^{N}\frac{1}{B_N}\bar{\phi}_*\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\frac{1}{B_N}\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right). \tag{A.28}$$

For the subsequent analysis it is convenient define the function

$$f(\hat{\lambda}_j-\hat{\lambda}_i, Y_{j0}-Y_{i0}) = \frac{1}{B_N^2}\bar{\phi}_*\left(\frac{\hat{\lambda}_j(\rho)-\hat{\lambda}_i(\rho)}{B_N}\right)\phi\left(\frac{Y_{j0}-Y_{i0}}{B_N}\right). \tag{A.29}$$

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

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 $\mathcal{T}_{p(\cdot)}$, there exists a finite constant M that does not depend on π such that*

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

Proof of Lemma A.6. We have

$$\begin{aligned}
&\mathbb{E}_i\big[f^m(\hat{\lambda}_j-\hat{\lambda}_i, Y_{j0}-Y_{i0})\big]\\
&= \frac{1}{B_N^{2(m-1)}}\int \frac{1}{B_N}\bar{\phi}_*\left(\frac{\hat{\lambda}-\hat{\lambda}_i}{B_N}\right)^m \frac{1}{B_N}\phi\left(\frac{y_0-Y_{i0}}{B_N}\right)^m p(\hat{\lambda}, y_0)d(\hat{\lambda}, y_0)\\
&= \frac{1}{B_N^{2(m-1)}}\int \left[\int \frac{1}{B_N}\bar{\phi}_*\left(\frac{\hat{\lambda}-\hat{\lambda}_i}{B_N}\right)^m \frac{1}{B_N}\phi\left(\frac{y_0-Y_{i0}}{B_N}\right)^m p(\hat{\lambda}, y_0|\lambda_i)d(\hat{\lambda}, y_0)\right]\pi(\lambda_i)d\lambda_i\\
&= \frac{1}{B_N^{2(m-1)}}\int_{\mathcal{T}_\lambda} \left[\int \frac{1}{B_N}\bar{\phi}_*\left(\frac{\hat{\lambda}-\hat{\lambda}_i}{B_N}\right)^m \frac{1}{B_N}\phi\left(\frac{y_0-Y_{i0}}{B_N}\right)^m p(\hat{\lambda}, y_0|\lambda_i)d(\hat{\lambda}, y_0)\right]\pi(\lambda_i)d\lambda_i\\
&\quad + \frac{1}{B_N^{2(m-1)}}\int_{\mathcal{T}_\lambda^c} \left[\int \frac{1}{B_N}\bar{\phi}_*\left(\frac{\hat{\lambda}-\hat{\lambda}_i}{B_N}\right)^m \frac{1}{B_N}\phi\left(\frac{y_0-Y_{i0}}{B_N}\right)^m p(\hat{\lambda}, y_0|\lambda_i)d(\hat{\lambda}, y_0)\right]\pi(\lambda_i)d\lambda_i.
\end{aligned}$$

The required result of the lemma follows if we show

$$
\begin{aligned}
I &= \frac{1}{p(\hat{\lambda}_i, Y_{i0})} \int_{\mathcal{T}_\lambda} \left[\int \frac{1}{B_N} \bar{\phi}_* \left(\frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{B_N} \phi \left(\frac{y_0 - Y_{i0}}{B_N} \right)^m p(\hat{\lambda}, y_0 | \lambda_i) d(\hat{\lambda}, y_0) \right] \pi(\lambda_i) d\lambda_i \\
&\leq M \qquad \text{(A.30)} \\
II &= \frac{1}{p(\hat{\lambda}_i, Y_{i0})} \int_{\mathcal{T}_\lambda^c} \left[\int \frac{1}{B_N} \bar{\phi}_* \left(\frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{B_N} \phi \left(\frac{y_0 - Y_{i0}}{B_N} \right)^m p(\hat{\lambda}, y_0 | \lambda_i) d(\hat{\lambda}, y_0) \right] \pi(\lambda_i) d\lambda_i \\
&\leq M \qquad \text{(A.31)}
\end{aligned}
$$

over the regions $\mathcal{T}_{\hat{\lambda} Y_0}$ and $\mathcal{T}_{p(\cdot)}$ and uniformly in π.

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

$$
\begin{aligned}
& \int \frac{1}{B_N} \bar{\phi}_* \left(\frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{B_N} \phi \left(\frac{y_0 - Y_{i0}}{B_N} \right)^m p(\hat{\lambda}, y_0 | \lambda) d(\hat{\lambda}, y_0) \\
&= \int \frac{1}{B_N} \bar{\phi}_* \left(\frac{\hat{\lambda} - \hat{\lambda}_i}{B_N} \right)^m \frac{1}{\sigma/\sqrt{T}} \exp \left(-\frac{1}{2} \left(\frac{\hat{\lambda} - \lambda_i}{\sigma/\sqrt{T}} \right)^2 \right) d\hat{\lambda} \\
&\quad \times \int \frac{1}{B_N} \phi \left(\frac{y_0 - Y_{i0}}{B_N} \right)^m \pi(y_0 | \lambda) dy_0 \\
&= I_1 \times I_2,
\end{aligned}
$$

say.

Notice that

$$
\begin{aligned}
I_1 &= \int \frac{1}{B_N}\bar{\phi}_*\left(\frac{\hat{\lambda}-\hat{\lambda}_i}{B_N}\right)^m \frac{1}{\sigma/\sqrt{T}}\exp\left(-\frac{1}{2}\left(\frac{\hat{\lambda}-\lambda_i}{\sigma/\sqrt{T}}\right)^2\right)d\hat{\lambda} \\
&= \int \bar{\phi}_*(\lambda^*)^m \frac{1}{\sigma/\sqrt{T}}\exp\left(-\frac{1}{2}\left(\frac{\hat{\lambda}_i-\lambda_i+B_N\lambda^*}{\sigma/\sqrt{T}}\right)^2\right)d\lambda^* \\
&= \int \bar{\phi}_*(\lambda^*)^m \exp\left(-\left((\hat{\lambda}_i-\lambda_i)B_N\lambda^*\right)\frac{1}{\sigma^2/T}\right)\exp\left(-\frac{1}{2}\left(\frac{B_N\lambda^*}{\sigma/\sqrt{T}}\right)^2\right)d\lambda^* \\
&\quad \times\left[\frac{1}{\sigma/\sqrt{T}}\exp\left(-\frac{1}{2}\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)^2\right)\right] \\
&\leq \int \bar{\phi}_*(\lambda^*)^m \exp\left(-\left(\frac{(\hat{\lambda}_i-\lambda_i)B_N}{\sigma^2/T}\right)\lambda^*\right)d\lambda^* \\
&\quad \times\left[\frac{1}{\sigma/\sqrt{T}}\exp\left(-\frac{1}{2}\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)^2\right)\right] \\
&\leq M\left(\int_0^\infty \bar{\phi}_*(\lambda^*)^m \exp\left(v_N\lambda^*\right)d\lambda^*\right)\left[\frac{1}{\sigma/\sqrt{T}}\exp\left(-\frac{1}{2}\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)^2\right)\right] \\
&\leq M\left[\frac{1}{\sigma/\sqrt{T}}\exp\left(-\frac{1}{2}\left(\frac{\hat{\lambda}_i-\lambda_i}{\sigma/\sqrt{T}}\right)^2\right)\right] \\
&= Mp(\hat{\lambda}_i|\lambda_i),
\end{aligned}
$$

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 $\mathcal{T}_{\hat{\lambda}Y_0}$ and $\mathcal{T}_{\lambda}$, $|\hat{\lambda}_i| \leq C'_N$ 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 $\int_0^\infty \bar{\phi}_*(\lambda^*)^m \exp\left(v_N\lambda^*\right)d\lambda^*$ 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

$$\begin{aligned} I_2 &= \int \frac{1}{B_N}\phi\left(\frac{y_0 - Y_{i0}}{B_N}\right)^m \pi(y_0|\lambda)dy_0 \\ &\leq M\int \frac{1}{B_N}\phi\left(\frac{y_0 - Y_{i0}}{B_N}\right)\pi(y_0|\lambda)dy_0 \\ &= M\big(1+o(1)\big)\pi(Y_{i0}|\lambda), \end{aligned}$$

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 λ, we obtain

$$\begin{aligned} I &\leq M\frac{1}{p(\hat{\lambda}_i, Y_{i0})}\int_{\mathcal{T}_\lambda}\Big[p(\hat{\lambda}_i|\lambda_i)\pi(Y_{i0}|\lambda_i)\Big]\pi(\lambda_i)d\lambda_i \\ &\leq M\frac{1}{p(\hat{\lambda}_i, Y_{i0})}\int p(\hat{\lambda}_i|\lambda_i)\pi(Y_{i0}|\lambda_i)\pi(\lambda_i)d\lambda_i = M, \end{aligned}$$

as required for (A.30).

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

$$\begin{aligned} II &\leq \frac{M}{p(\hat{\lambda}_i, Y_{i0})B_N^2}\int_{\mathcal{T}_\lambda^c}\pi(\lambda_i)d\lambda_i \\ &\leq MN^{-\epsilon}\left(\frac{1}{B_N^2}\right)\left(N\int_{\mathcal{T}_\lambda^c}\pi(\lambda_i)d\lambda_i\right) \\ &\leq MN^{-\epsilon}o_{u.\pi}(N^+)o_{u.\pi}(N^+) \\ &\leq M, \end{aligned}$$

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,*

$$\begin{aligned} &(a)\ \sup_{\pi\in\Pi}\ \sup_{(\hat{\lambda}_i, Y_{i0})\in\mathcal{T}_{\hat{\lambda}Y_0}\cap\mathcal{T}_{p(\cdot)}}\left|\frac{p_{*i}}{p_i} - 1\right| = o(1), \\ &(b)\ \sup_{\pi\in\Pi}\ \sup_{(\hat{\lambda}_i, Y_{i0})\in\mathcal{T}_{\hat{\lambda}Y_0}\cap\mathcal{T}_{p(\cdot)}}\left|\frac{p_i}{p_{*i}} - 1\right| = o(1). \end{aligned}$$

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

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

Notice that

$$\begin{aligned}
\left|\frac{p_{*i}}{p_i} - 1\right| &= \left|\frac{p_{*i} - p_i}{p_i}\right| \\
&\leq \frac{1}{p_i}\int \left|p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i)\right|\pi(\lambda_i)d\lambda_i \\
&= \frac{1}{p_i}\int_{\mathcal{T}_\lambda} \left|p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i)\right|\pi(\lambda_i)d\lambda_i \\
&\quad + \frac{1}{p_i}\int_{\mathcal{T}_\lambda^c} \left|p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i)\right|\pi(\lambda_i)d\lambda_i \\
&= I + II, \text{ say.}
\end{aligned}$$

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 $\mathcal{T}_{\hat{\lambda}Y_0}$, we can choose a constant M that does not depend on π such that for N sufficiently large

$$\begin{aligned}
\frac{1}{\sqrt{B_N^2 + \sigma^2/T}}\phi\left(\frac{\hat{\lambda}_i - \lambda_i}{\sqrt{B_N^2 + \sigma^2/T}}\right) &= \frac{1}{\sqrt{\sigma^2/T}}\phi\left(\frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}}\right) \\
&\quad \times \frac{\sqrt{\sigma^2/T}}{\sqrt{B_N^2 + \sigma^2/T}}\exp\left\{\frac{1}{2}\left(\frac{\hat{\lambda}_i - \lambda_i}{\sqrt{B_N^2 + \sigma^2/T}}\right)^2 \frac{B_N^2}{\sigma^2/T}\right\} \\
&\leq \frac{1}{\sqrt{\sigma^2/T}}\phi\left(\frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}}\right)\exp(M(C_N' + C_N)^2 B_N^2) \\
&= \frac{1}{\sqrt{\sigma^2/T}}\phi\left(\frac{\hat{\lambda}_i - \lambda_i}{\sqrt{\sigma^2/T}}\right)(1 + o(1)),
\end{aligned}$$

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 $(\hat{\lambda}_i, Y_{i0}) \in \mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}$ and in $\pi \in \Pi$.

According to Assumption 3.4,

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

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

Then,

$$\begin{aligned}
I &= \frac{1}{p_i}\int_{\mathcal{T}_\lambda}\left|p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i)\right|\pi(\lambda_i)d\lambda_i \\
&= \frac{1}{p_i}\int_{\mathcal{T}_\lambda}\left|\frac{1}{\sqrt{B_N^2+\sigma^2/T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sqrt{B_N^2+\sigma^2/T}}\right)\int \frac{1}{B_N}\phi\left(\frac{y_{i0}-\tilde{y}_{i0}}{B_N}\right)\pi(\tilde{y}_{i0}|\lambda_i)d\tilde{y}_{i0}\right. \\
&\qquad \left. - \frac{1}{\sqrt{\sigma^2/T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sqrt{\sigma^2/T}}\right)\pi(y_{i0}|\lambda_i)\right|\pi(\lambda_i)d\lambda_i \\
&\leq |(1+o(1))^2-1|\frac{1}{p_i}\int_{\mathcal{T}_\lambda}\frac{1}{\sqrt{\sigma^2/T}}\phi\left(\frac{\hat{\lambda}_i-\lambda_i}{\sqrt{\sigma^2/T}}\right)\pi(y_{i0}|\lambda_i)\pi(\lambda_i)d\lambda_i \\
&\leq |(1+o(1))^2-1| = o(1).
\end{aligned}$$

Note that the $o(1)$ term does not depend on $(\hat{\lambda}_i, Y_{i0}) \in \mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}$ 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

$$\begin{aligned}
II &= \frac{1}{p_i}\int_{\mathcal{T}_\lambda^c}\left|p_*(\hat{\lambda}_i, y_{i0}|\lambda_i) - p(\hat{\lambda}_i, y_{i0}|\lambda_i)\right|\pi(\lambda_i)d\lambda_i \\
&\leq \frac{2M}{p_i}\int_{\mathcal{T}_\lambda^c}\pi(\lambda_i)d\lambda_i \\
&\leq 2M\sup_{\pi\in\Pi} N^{1-\epsilon}\int_{\mathcal{T}_\lambda^c}\pi(\lambda_i)d\lambda_i \\
&= o(1),
\end{aligned}$$

where the second inequality holds since $p_i > \frac{N^\epsilon}{N}$ under the truncation $\mathcal{T}_{p(\cdot)}$ 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),

$$p_{*i} = p_i(1+o(1))$$

uniformly in $(\hat{\lambda}_i, Y_{i0}) \in \mathcal{T}_{\hat{\lambda} Y_0} \cap \mathcal{T}_{p(\cdot)}$ and in $\pi \in \Pi$. Then, for some finite constant M that does not depend on $(\hat{\lambda}_i, Y_{i0})$ and π,

$$\begin{aligned}
\sup_{\pi\in\Pi} \sup_{(\hat{\lambda}_i, Y_{i0})\in \mathcal{T}_{\hat{\lambda} Y_0}\cap \mathcal{T}_{p(\cdot)}} \frac{|p_i - p_{*i}|}{p_{*i}} &= \sup_{\pi\in\Pi} \sup_{(\hat{\lambda}_i, Y_{i0})\in \mathcal{T}_{\hat{\lambda} Y_0}\cap \mathcal{T}_{p(\cdot)}} \frac{|p_i - p_{*i}|}{p_i}\frac{p_i}{p_{*i}} \\
&= \sup_{\pi\in\Pi} \sup_{(\hat{\lambda}_i, Y_{i0})\in \mathcal{T}_{\hat{\lambda} Y_0}\cap \mathcal{T}_{p(\cdot)}} \frac{|p_i - p_{*i}|}{p_i}\frac{p_i}{p_i(1+o(1))} \\
&\le M \sup_{\pi\in\Pi} \sup_{(\hat{\lambda}_i, Y_{i0})\in \mathcal{T}_{\hat{\lambda} Y_0}\cap \mathcal{T}_{p(\cdot)}} \frac{|p_i - p_{*i}|}{p_i} \\
&= o(1),
\end{aligned}$$

as required for Part (b). ■

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

(a) $\mathbb{E}_i\big[R_{2i}^4(\widetilde{\rho})\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\big] \le M L_N^4 p_i^4 \mathbb{I}(\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)})$,

(b) $\mathbb{E}_i\big[R_{1i}^4\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\big] \le M \frac{L_N^4}{B_N^4} p_i^4 \mathbb{I}(\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)})$,

(c) $\mathbb{E}_i\Big[N(\hat{p}_i^{(-i)} - p_{*i})^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\Big] \le \frac{M}{B_N^2} p_i \mathbb{I}(\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)})$,

(d) $\mathbb{E}_i\Big[N(d\hat{p}_i^{(-i)} - dp_{*i})^2\mathbb{I}(\mathcal{T}_{\hat{\rho}}\mathcal{T}_{\bar{U}}\mathcal{T}_{Y_0}\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)}\mathcal{T}_{\tilde{p}(\cdot)})\Big] \le \frac{M}{B_N^2} p_i \mathbb{I}(\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)})$, *where* M *is a finite constant that does not depend on* $\pi \in \Pi$.

(e) For any finite $m > 1$, $\int_{\mathcal{T}_{\hat{\lambda} Y_0}\cap \mathcal{T}_{p(\cdot)}} \left(\frac{p_i}{p_{*i}}\right)^m d\hat{\lambda}_i dy_{i0} = o_{u.\pi}(N^+)$.

(f) $N\mathbb{E}\big[\mathbb{P}_i\{\hat{p}_i^{(-i)} - p_{*i} < -p_{*i}/4\}\mathbb{I}(\mathcal{T}_{\hat{\lambda} Y_0}\mathcal{T}_{p(\cdot)})\big] = o_{u.\pi}(N^+)$.

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

$$\begin{aligned}
\bar{\phi}(x) &= \phi\left(x + \frac{M}{L_N B_N}\right)\mathbb{I}\left\{x \le -\frac{M}{L_N B_N}\right\} + \phi(0)\mathbb{I}\left\{|x| \le \frac{M}{L_N B_N}\right\} \\
&\quad +\phi\left(x - \frac{M}{L_N B_N}\right)\mathbb{I}\left\{x \ge \frac{M}{L_N B_N}\right\} \\
\bar{\phi}_*(x) &= \bar{\phi}(x)\left(|x| + \frac{M}{L_N B_N}\right).
\end{aligned}$$

First, recall that according to (A.28) and (A.29), in the region $\mathcal{T}_{\hat{\rho}} \cap \mathcal{T}_{\bar{U}} \cap \mathcal{T}_{Y_0}$

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

Then,

$$\begin{aligned}
|R_{2i}(\widetilde{\rho})|^4 &\leq \left[\frac{ML_N}{N-1} \sum_{j \neq i}^{N} f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) \right]^4 \\
&= \left[\frac{ML_N}{N-1} \sum_{j \neq i}^{N} \Big\{ f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) - \mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})] \right. \\
&\quad \left. + \mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})] \Big\} \right]^4 \\
&\leq ML_N^4 \left[\frac{1}{N-1} \sum_{j \neq i}^{N} \Big(f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0}) - \mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})] \Big) \right]^4 \\
&\quad + ML_N^4 \Big[\mathbb{E}_i[f(\hat{\lambda}_j - \hat{\lambda}_i, Y_{j0} - Y_{i0})] \Big]^4 \\
&= ML_N^4 \big(A_1 + A_2 \big),
\end{aligned}$$

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

$$\begin{aligned}
\left(\sum a_j \right)^4 &= \left(\sum a_j^2 + 2 \sum_j \sum_{i>j} a_j a_i \right)^2 \\
&= \left(\sum a_j^2 \right)^2 + 4 \left(\sum a_j^2 \right) \left(\sum_j \sum_{i>j} a_j a_i \right) + 4 \left(\sum_j \sum_{i>j} a_j a_i \right)^2 \\
&= \sum a_j^4 + 6 \sum_j \sum_{i>j} a_j^2 a_i^2 \\
&\quad + 4 \left(\sum a_j^2 \right) \left(\sum_j \sum_{i>j} a_j a_i \right) + 4 \sum_j \sum_{i>j} \sum_{l \neq j} \sum_{k>l} a_j a_i a_l a_k,
\end{aligned}$$

where

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

Notice that conditional on $(\hat{\lambda}_i(\rho), Y_{i0})$, the random variables a_j have mean zero and are *iid* across $j \neq i$. This implies that

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

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 $\mathcal{T}_{\hat{\rho}} \cap \mathcal{T}_{\bar{U}} \cap \mathcal{T}_{Y_0} \cap \mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)} \cap \mathcal{T}_{\tilde{p}(\cdot)}$,

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

The second inequality holds because over $\mathcal{T}_{p(\cdot)}$, $p_i \geq \frac{N^\epsilon}{N} \geq \frac{M}{NB_N^2}$ 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

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

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 π such that

$$p_i - p_{*i} \leq cp_{*i}$$

over the region $\mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}$. Thus,

$$\left(\frac{p_i}{p_{*i}}\right)^m = \left(1 + \frac{p_i - p_{*i}}{p_{*i}}\right)^m \le (1+c)^m.$$

We deduce that

$$\int_{\mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}} \left(\frac{p_i}{p_{*i}}\right)^m d\hat{\lambda}_i dy_{i0} \le (1+c)^m \int_{\mathcal{T}_{\hat{\lambda}Y_0} \cap \mathcal{T}_{p(\cdot)}} d\hat{\lambda}_i dy_{i0} \le (1+c)^m \big(2C'_N\big)^2 = o_{u.\pi}(N^+),$$

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

$$\begin{aligned} \hat{p}_i^{(-i)} - p_{*i} &= \frac{1}{N-1}\sum_{j\neq i}^{N}\left\{\frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_j - \hat{\lambda}_i}{B_N}\right)\frac{1}{B_N}\phi\left(\frac{Y_{j0} - Y_{i0}}{B_N}\right)\right. \\ &\qquad \left. - \mathbb{E}_i\left[\frac{1}{B_N}\phi\left(\frac{\hat{\lambda}_j - \hat{\lambda}_i}{B_N}\right)\frac{1}{B_N}\phi\left(\frac{Y_{j0} - Y_{i0}}{B_N}\right)\right]\right\} \\ &= \frac{1}{B_N^2(N-1)}\sum_{j\neq i}^{N}\Big(\psi_i(\hat{\lambda}_j, Y_{j0}) - \mathbb{E}_i[\psi_i(\hat{\lambda}_j, Y_{j0})]\Big). \end{aligned}$$

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

$$\begin{aligned} & N\mathbb{P}_i\left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\}\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}) \\ = \;& N\mathbb{P}_i\left\{\frac{1}{B_N^2(N-1)}\sum_{j\neq i}^{N}\Big(\psi_i(\hat{\lambda}_j, Y_{j0}) - \mathbb{E}_i[\psi_i(\hat{\lambda}_j, Y_{j0})]\Big) < -\frac{p_{*i}}{4}\right\}\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}) \\ \le \;& 2N\exp\left(-\frac{1}{4}\frac{B_N^4(N-1)p_{*i}^2/16}{\mathbb{E}_i[\psi_i(\hat{\lambda}_j, Y_{j0})^2] + MB_N^2 p_{i*}/4}\right)\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)}). \end{aligned}$$

[15] For a bounded function f and a sequence of iid random variables X_i,

$$\mathbb{P}\left\{\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(f(X_i) - \mathbb{E}[f(X_i)])\right| > x\right\} \le 2\exp\left(-\frac{1}{4}\frac{x^2}{\mathbb{E}[f(X_i)^2] + \frac{1}{\sqrt{N}}x\sup_x|f(x)|}\right).$$

Using an argument similar to the proof of Lemma A.6 one can show that

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

In turn

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

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

$$\frac{p_{*i}^2}{p_i + p_{*i}} \geq \frac{p_i}{(2+c)(1+c)^2}.$$

Then, on the region $\mathcal{T}_{p(\cdot)}$

$$\begin{aligned}
&N\mathbb{E}\left[\mathbb{P}_i\left\{\hat{p}_i^{(-i)} - p_{*i} < -\frac{p_{*i}}{4}\right\}\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right] \\
&\leq 2\mathbb{E}\left[\exp\left(-MNB_N^2\frac{p_{*i}^2}{p_i + p_{*i}} + \ln N\right)\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\right] \\
&\leq 2\mathbb{E}\big[\exp\big(-MNB_N^2 p_i + \ln N\big)\mathbb{I}(\mathcal{T}_{\hat{\lambda}Y_0}\mathcal{T}_{p(\cdot)})\big] \\
&\leq 2\exp\big(-MB_N^2 N^{\epsilon} + \ln N\big) \\
&= o(1),
\end{aligned}$$

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

$$\mathbb{P}\big(|\lambda| \geq C\big) \leq \mathbb{P}\Big(|\lambda - \mu_\lambda| \geq C - |\mu_\lambda|\Big) \leq 2\exp\left(-\frac{C - |\mu_\lambda|}{2\sigma_\lambda^2}\right) \tag{A.32}$$

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 $\sigma_y^2 = \sigma_{y|\lambda}^2 + \alpha_1^2 \sigma_\lambda^2$.

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

$$\sup_{\pi \in \Pi} \sup_{|y| \leq C_N', |\lambda| < C_N} \left| \frac{\int \frac{1}{B_N} \phi\left(\frac{\tilde{y}-y}{B_N}\right) \pi(\tilde{y}|\lambda) d\tilde{y}}{\pi(y|\lambda)} - 1 \right| = \sup_{\pi \in \Pi} \sup_{|y| \leq C_N', |\lambda| < C_N} \big| \mathcal{R}_{1,N} \cdot \mathcal{R}_{2,N} - 1 \big|,$$

where

$$\mathcal{R}_{1,N} = \sqrt{\frac{\sigma^2_{y|\lambda}}{\sigma^2_{y|\lambda} + B_N^2}} \leq 1, \quad \mathcal{R}_{2,N} = \exp\left\{ -\frac{1}{2}(y - \mu_y(\lambda))^2 \left(\frac{1}{\sigma^2_{y|\lambda} + B_N^2} - \frac{1}{\sigma^2_{y|\lambda}} \right) \right\} \geq 1.$$

$\mathcal{R}_{1,N}$ can be bounded from below by replacing $\sigma^2_{y|\lambda}$ with $\delta_{\sigma^2_{y|\lambda}}$. Because $B_N \longrightarrow 0$ as $N \longrightarrow \infty$, $\mathcal{R}_{1,N} \longrightarrow 1$ uniformly. For the term $\mathcal{R}_{2,N}$ notice that

$$\begin{aligned} \big(y - \mu_y(\lambda)\big)^2 \left(\frac{1}{\sigma^2_{y|\lambda}} - \frac{1}{\sigma^2_{y|\lambda} + B_N^2} \right) &= \big(y - \alpha_0 - \alpha_1 \lambda\big)^2 \frac{B_N^2}{\sigma^2_{y|\lambda}\big(\sigma^2_{y|\lambda} + B_N^2\big)} \\ &\leq 3\big((C_N')^2 + M_{\alpha_0}^2 + M_{\alpha_1}^2 C_N^2\big) \frac{B_N^2}{\big(\delta_{\sigma^2_{y|\lambda}}\big)^2} \longrightarrow 0 \end{aligned}$$

as $N \longrightarrow \infty$ because $B_N C_N = o(1)$ and $B_N C_N' = o(1)$ according to Assumption 3.3. Thus, $\mathcal{R}_{2,N} \longrightarrow 1$ uniformly which delivers the desired result.

Assumption 3.5. The first step is to derive the conditional prior distribution $\pi(\lambda|y)$ which is of the form $\lambda|y \sim N(\mu_{\lambda|y}, \sigma^2_{y|\lambda})$. The prior mean function is of the form $\mu_{\lambda|y} = \gamma_0 + \gamma_1 y$. If the prior for λ is a point mass, i.e., $\sigma^2_\lambda = 0$, then the distribution of $(\lambda|y)$ is also a point mass with $\mu_{\lambda|y} = \mu_\lambda$ and $\sigma^2_{\lambda|y} = 0$. It can be verified that the coefficients γ_0, γ_1, and the variance $\sigma^2_{y|\lambda}$ are bounded from above in absolute value.

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

$$m(\hat{\lambda}, y; \pi) = \left(\frac{1}{\sigma^2_{y|\lambda}} + \frac{1}{\sigma^2/T} \right)^{-1} \left(\frac{1}{\sigma^2_{y|\lambda}}(\gamma_0 + \gamma_1 y_0) + \frac{1}{\sigma^2/T}\hat{\lambda} \right) = \bar{\gamma}_0 + \bar{\gamma}_1 y_0 + \bar{\gamma}_2 \hat{\lambda}. \quad \text{(A.33)}$$

The $\bar{\gamma}$ 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

$$\mu_{\hat{\lambda},y} = \begin{bmatrix} \mu_\lambda \\ \alpha_0 + \alpha_1 \mu_\lambda \end{bmatrix}, \quad \Sigma_{\hat{\lambda},y} = \begin{bmatrix} \sigma^2_\lambda + \sigma^2/T & \gamma_1\big(\sigma^2_{y|\lambda} + \alpha_1^2 \sigma^2_\lambda\big) \\ \gamma_1\big(\sigma^2_{y|\lambda} + \alpha_1^2\sigma^2_\lambda\big) & \sigma^2_{y|\lambda} + \alpha_1^2 \sigma^2_\lambda \end{bmatrix}. \tag{A.34}$$

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 $\hat{\lambda} = \lambda + (\sum_{t=1}^T u_t)/T$. 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 $\pi_{mix} \in \Pi^{(K)}_{mix}$.

Assumption 3.2. Consider the marginal density of λ given by $\pi_{mix}(\lambda) = \sum_{k=1}^K \omega_k \pi_k(\lambda)$. Thus, for any non-negative integrable function $h(\cdot)$ we can use the crude bound

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

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

Assumption 3.4 The key insight is that we can express

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

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

$$\begin{aligned} &\sup_{\pi_{mix} \in \Pi^{(K)}_{mix}} \sup_{|y| \leq C'_N, |\lambda| < C_N} \left| \frac{\int \frac{1}{B_N} \phi\left(\frac{\tilde{y}-y}{B_N}\right) \big[\pi_{mix}(\tilde{y}|\lambda) - \pi_{mix}(y|\lambda)\big] d\tilde{y}}{\pi_{mix}(y|\lambda)} \right| \\ \leq &\sum_{k=1}^K \sup_{\pi_k \in \Pi_k} \sup_{|y| \leq C'_N, |\lambda| < C_N} \left| \frac{\int \frac{1}{B_N} \phi\left(\frac{\tilde{y}-y}{B_N}\right) \big[\pi_k(\tilde{y}|\lambda) - \pi_k(y|\lambda)\big] d\tilde{y}}{\pi_k(y|\lambda)} \right| \\ = &\; o(1). \end{aligned}$$

Assumption 3.5. The prior distribution of λ given y is a mixture of normals with weights that are a function of y:

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

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

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

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

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

where the $\bar{\gamma}$ 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 θ

corresponds to the "true" θ_0. We do so, by calculating

$$(\theta_*, \xi_*) = \text{argmax}_{\theta,\xi}\, \mathbb{E}^{\mathcal{Y}}_{\theta_0}\left[\ln p(Y, X_2|H, \theta, \xi)\right], \tag{A.39}$$

and verifying that $\theta_* = \theta_0$. 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 Σ only depends on γ, but not on (ρ, α). Moreover, we will split ξ into the parameters that characterize the conditional mean of λ, denoted by Φ, and ω, which are the non-redundant elements of the prior covariance matrix $\underline{\Omega}$. Finally, we define

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

with the understanding that $\theta_1 = (\rho, \alpha)$ and excludes γ. 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

$$\begin{aligned}
&\ln p(y, x_2|h, \theta_1, \gamma, \phi, \omega) \qquad\qquad (A.40)\\
&= C - \frac{1}{2}\ln|\Sigma(\gamma)| - \frac{1}{2}\big(\tilde{y}(\theta_1) - w\hat{\lambda}(\theta)\big)'\Sigma^{-1}(\gamma)\big(\tilde{y}(\theta_1) - w\hat{\lambda}(\theta)\big)\\
&\quad -\frac{1}{2}\ln\big|\underline{\Omega}\big| + \frac{1}{2}\ln\big|\bar{\Omega}(\gamma,\omega)\big|\\
&\quad -\frac{1}{2}\bigg(\hat{\lambda}(\theta)'w'\Sigma^{-1}(\gamma)w\hat{\lambda}(\theta) + \phi'\tilde{h}\underline{\Omega}^{-1}\tilde{h}'\phi - \bar{\lambda}'(\theta,\xi)\bar{\Omega}^{-1}(\gamma,\omega)\bar{\lambda}(\theta,\xi)\bigg),
\end{aligned}$$

where

$$\begin{aligned}
\hat{\lambda}(\theta) &= (w'\Sigma^{-1}(\gamma)w)^{-1}w'\Sigma^{-1}(\gamma)\tilde{y}(\theta_1)\\
\bar{\Omega}^{-1}(\gamma,\omega) &= \underline{\Omega}^{-1} + w'\Sigma^{-1}(\gamma)w, \quad \bar{\lambda}(\theta,\xi) = \bar{\Omega}(\gamma,\omega)\big(\underline{\Omega}^{-1}\tilde{h}'\phi + w'\Sigma^{-1}(\gamma)w\hat{\lambda}(\theta)\big).
\end{aligned}$$

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

$$\begin{aligned}
\ln p(y|h, \rho, \gamma, \phi, \omega) &= C - \frac{T-1}{2}\ln|\gamma^2| - \frac{1}{2\gamma^2}\big(\tilde{y}(\rho) - \iota\hat{\lambda}(\rho)\big)'\big(\tilde{y}(\rho) - \iota\hat{\lambda}(\rho)\big)\\
&\quad -\frac{1}{2}\ln\big|\omega^2\big| - \frac{1}{2}\ln\big|\gamma^2/T\big| + \frac{1}{2}\ln(1/T) + \frac{1}{2}\ln\big|\bar{\Omega}(\gamma,\omega)\big|\\
&\quad -\frac{1}{2}\bigg(\frac{T}{\gamma^2}\hat{\lambda}^2(\rho) + \frac{1}{\omega^2}\phi'\tilde{h}\tilde{h}'\phi - \frac{1}{\bar{\Omega}(\gamma,\omega)}\bar{\lambda}^2(\theta,\xi)\bigg),
\end{aligned}$$

where

$$
\begin{aligned}
\hat{\lambda}(\rho) &= \frac{1}{T}\iota'\tilde{y}(\rho) \\
\bar{\Omega}^{-1}(\gamma,\omega) &= \frac{1}{\omega^2} + \frac{1}{\gamma^2/T}, \quad \bar{\lambda}(\theta,\xi) = \bar{\Omega}(\gamma,\omega)\left(\frac{1}{\omega^2}\tilde{h}'\phi + \frac{T}{\gamma^2}\hat{\lambda}(\rho)\right).
\end{aligned}
$$

Note that

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

In turn, we can write

$$
\begin{aligned}
&\ln p(y|h,\rho,\gamma,\phi,\omega) \\
&= C - \frac{T-1}{2}\ln|\gamma^2| - \frac{1}{2\gamma^2}\tilde{y}(\rho)'(I-\iota\iota'/T)\tilde{y}(\rho) - \frac{1}{2}\ln\left|\omega^2+\gamma^2/T\right| \\
&\quad -\frac{1}{2}\left(\frac{T}{\gamma^2}\hat{\lambda}^2(\rho) + \frac{1}{\omega^2}\phi'\tilde{h}\tilde{h}'\phi - \frac{\omega^2\gamma^2/T}{\omega^2+\gamma^2/T}\left(\frac{1}{\omega^2}\tilde{h}'\phi + \frac{T}{\gamma^2}\hat{\lambda}(\rho)\right)^2\right) \\
&= C - \frac{T-1}{2}\ln|\gamma^2| - \frac{1}{2\gamma^2}\tilde{y}(\rho)'(I-\iota\iota'/T)\tilde{y}(\rho) - \frac{1}{2}\ln\left|\omega^2+\gamma^2/T\right| \\
&\quad -\frac{1}{2(\omega^2+\gamma^2/T)}\left(\phi'\tilde{h}\tilde{h}'\phi - 2\hat{\lambda}(\rho)\tilde{h}'\phi + \hat{\lambda}^2(\rho)\right).
\end{aligned}
$$

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

$$
\begin{aligned}
&\mathbb{E}\big[\ln p(Y|H,\rho,\gamma,\phi,\omega)\big] \qquad \text{(A.41)} \\
&= C - \frac{T-1}{2}\ln|\gamma^2| - \frac{1}{2\gamma^2}\mathbb{E}\big[\tilde{Y}(\rho)'(I-\iota\iota'/T)\tilde{Y}(\rho)\big] - \frac{1}{2}\ln\left|\omega^2+\gamma^2/T\right| \\
&\quad -\frac{1}{2(\omega^2+\gamma^2/T)}\Bigg(\Big(\phi - \big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\hat{\lambda}(\rho)]\Big)'\mathbb{E}[\tilde{H}\tilde{H}']\Big(\phi - \big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\hat{\lambda}(\rho)]\Big) \\
&\quad -\mathbb{E}[\hat{\lambda}(\rho)\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\hat{\lambda}(\rho)] + \mathbb{E}[\hat{\lambda}^2(\rho)]\Bigg).
\end{aligned}
$$

We deduce that

$$
\phi_*(\rho) = \big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\hat{\lambda}(\rho)]. \qquad \text{(A.42)}
$$

To evaluate $\phi_*(\rho_0)$, note that $\hat{\lambda}(\rho_0) = \lambda + \iota'u/T$. 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,

$$
\phi_*(\rho_0) = \big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\lambda]. \qquad \text{(A.43)}
$$

The pseudo-true value is obtained through a population regression of λ on H.

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

$$\begin{aligned}
&\mathbb{E}\big[\ln p(Y|H,\rho,\gamma,\phi_*(\rho),\omega)\big] \qquad \text{(A.44)}\\
&= C - \frac{T-1}{2}\ln|\gamma^2| - \frac{1}{2\gamma^2}\mathbb{E}\big[\tilde{Y}(\rho)'(I-\iota\iota'/T)\tilde{Y}(\rho)\big]\\
&\quad -\frac{1}{2}\ln\big|\omega^2+\gamma^2/T\big| - \frac{1}{2(\omega^2+\gamma^2/T)}\big(\mathbb{E}[\hat{\lambda}^2(\rho)] - \mathbb{E}[\hat{\lambda}(\rho)\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\hat{\lambda}(\rho)]\big).
\end{aligned}$$

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

$$\begin{aligned}
\gamma_*^2(\rho) &= \frac{1}{T-1}\mathbb{E}\big[\tilde{Y}(\rho)'(I-\iota\iota'/T)\tilde{Y}(\rho)\big] \qquad \text{(A.45)}\\
\omega_*^2(\rho)+\gamma_*^2(\rho)/T &= \big(\mathbb{E}[\hat{\lambda}^2(\rho)] - \mathbb{E}[\hat{\lambda}(\rho)\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\hat{\lambda}(\rho)]\big).
\end{aligned}$$

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

$$\gamma_*^2(\rho_0)=\gamma_0^2, \quad \omega_*^2(\rho_0)=\mathbb{E}[\lambda^2]-\mathbb{E}[\lambda\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\lambda]. \qquad \text{(A.46)}$$

Given $\rho=\rho_0$ the pseudo-true value for γ^2 is the "true" γ_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 λ onto H.

Using (A.45), we can now concentrate out γ^2 and ω^2 from the objective function (A.44):

$$\begin{aligned}
&\mathbb{E}\big[\ln p(Y|H,\rho,\gamma_*(\rho),\phi_*(\rho),\omega_*(\rho)\big] \qquad \text{(A.47)}\\
&= C - \frac{T-1}{2}\ln\big|\mathbb{E}\big[\tilde{Y}(\rho)'(I-\iota\iota'/T)\tilde{Y}(\rho)\big]\big|\\
&\quad -\frac{1}{2}\ln\big|\mathbb{E}[\tilde{Y}'(\rho)\iota\iota'\tilde{Y}(\rho)] - \mathbb{E}[\tilde{Y}'(\rho)\iota\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\iota'\tilde{Y}(\rho)]\big|.
\end{aligned}$$

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

$$\begin{aligned}
\text{F.O.C.}(\rho) &= (T-1)\frac{\mathbb{E}\big[X'(I-\iota\iota'/T)\tilde{Y}(\rho)\big]}{\mathbb{E}\big[\tilde{Y}(\rho)'(I-\iota\iota'/T)\tilde{Y}(\rho)\big]}\\
&\quad +\frac{\mathbb{E}[X'\iota\iota'\tilde{Y}(\rho)] - \mathbb{E}[X'\iota\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\iota'\tilde{Y}(\rho)]}{\mathbb{E}[\tilde{Y}'(\rho)\iota\iota'\tilde{Y}(\rho)] - \mathbb{E}[\tilde{Y}'(\rho)\iota\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\iota'\tilde{Y}(\rho)]}.
\end{aligned}$$

We will now verify that F.O.C.$(\rho_0) = 0$. Because both denominators are strictly positive, we can rewrite the condition as

$$\begin{aligned}
\text{F.O.C.}(\rho_0) \;=\; & (T-1)\mathbb{E}\big[X'(I-\iota\iota'/T)\tilde{Y}(\rho_0)\big] \\
& \times\Big(\mathbb{E}[\tilde{Y}'(\rho_0)\iota\iota'\tilde{Y}(\rho_0)] - \mathbb{E}[\tilde{Y}'(\rho_0)\iota\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\iota'\tilde{Y}(\rho_0)]\Big) \\
& + \mathbb{E}\big[\tilde{Y}(\rho_0)'(I-\iota\iota'/T)\tilde{Y}(\rho_0)\big] \\
& \times\Big(\mathbb{E}[X'\iota\iota'\tilde{Y}(\rho_0)] - \mathbb{E}[X'\iota\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\iota'\tilde{Y}(\rho_0)]\Big).
\end{aligned} \tag{A.48}$$

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:

$$\begin{aligned}
\mathbb{E}\big[X'(I-\iota\iota'/T)\tilde{Y}(\rho_0)\big] &= \mathbb{E}\big[X'(I-\iota\iota'/T)u\big] = \mathbb{E}[X'u] - \mathbb{E}[X'\iota\iota'u]/T = -\mathbb{E}[X'\iota\iota'u]/T \\
\mathbb{E}[\tilde{Y}'(\rho_0)\iota\iota'\tilde{Y}(\rho)] &= \mathbb{E}\big[(\lambda\iota' + u')\iota\iota'(\iota\lambda + u)\big] = T^2\mathbb{E}[\lambda^2] + \mathbb{E}[u'\iota\iota'u] = T^2\mathbb{E}[\lambda^2] + T\gamma_0^2 \\
\mathbb{E}[\tilde{H}\iota'\tilde{Y}(\rho_0)] &= \mathbb{E}[\tilde{H}\iota'(\iota\lambda + u)] = T\mathbb{E}[\tilde{H}\lambda] \\
\mathbb{E}\big[\tilde{Y}(\rho_0)'(I-\iota\iota'/T)\tilde{Y}(\rho_0)\big] &= \mathbb{E}\big[u'(I-\iota\iota'/T)u\big] = (T-1)\gamma^2 \\
\mathbb{E}[X'\iota\iota'\tilde{Y}(\rho_0)] &= \mathbb{E}[X'\iota\iota'(\iota\lambda + u)] = T\mathbb{E}[X'\iota\lambda] + \mathbb{E}[X'\iota\iota'u].
\end{aligned}$$

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:

$$\begin{aligned}
& \text{F.O.C.}(\rho_0) \\
&= -(T-1)\big(\mathbb{E}[X'\iota\iota'u]\big)\Big(\gamma_0^2 + T\big(\mathbb{E}[\lambda^2] - \mathbb{E}[\lambda\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\lambda]\big)\Big) \\
&\quad + \Big(\mathbb{E}[X'\iota\iota'u] + T\big(\mathbb{E}[X'\iota\lambda] - \mathbb{E}[X'\iota\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\lambda]\big)\Big)(T-1)\gamma_0^2 \\
&= T(T-1)\Big[\gamma_0^2\Big(\mathbb{E}[X'\iota\lambda] - \mathbb{E}[X'\iota\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\lambda]\Big) \\
&\quad - \mathbb{E}[X'\iota\iota'u]\Big(\mathbb{E}[\lambda^2] - \mathbb{E}[\lambda\tilde{H}']\big(\mathbb{E}[\tilde{H}\tilde{H}']\big)^{-1}\mathbb{E}[\tilde{H}\lambda]\Big)\Big].
\end{aligned} \tag{A.49}$$

We now have to analyze the terms involving $X'\iota$. 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 $a_t = \sum_{\tau=0}^{t-1} \rho_0^\tau$ and $b = \sum_{t=1}^{T-1} a_t$. Thus, we can write

$$Y_t = \rho_0^t Y_0 + \lambda a_t + \sum_{\tau=0}^{t-1} \rho_0^\tau U_{t-\tau}, \quad t > 0.$$

Consequently,

$$X'\iota = \sum_{t=0}^{T-1} Y_t = Y_0 \left(\sum_{t=0}^{T-1} \rho_0^t \right) + \lambda \left(\sum_{t=1}^{T-1} a_t \right) + \sum_{t=1}^{T-1} \sum_{\tau=0}^{t-1} \rho_0^\tau U_{t-\tau} = a_T y_0 + b\lambda + \sum_{t=1}^{T-1} a_t U_{T-t}.$$

Thus, we obtain

$$\begin{aligned}
\mathbb{E}[X'\iota\iota' u] &= \mathbb{E}\left[\left(a_T Y_0 + b\lambda + \sum_{t=1}^{T-1} a_t U_{T-t} \right) \left(\sum_{t=1}^{T} U_t \right) \right] = b\gamma_0^2 \\
\mathbb{E}[X'\iota\lambda] &= \mathbb{E}\left[\left(a_T Y_0 + b\lambda + \sum_{t=1}^{T-1} a_t U_{T-t} \right) \lambda \right] = a_T \mathbb{E}[Y_0\lambda] + b\mathbb{E}[\lambda^2] \\
\mathbb{E}[X'\iota\tilde{H}'] &= \mathbb{E}\left[\left(a_T Y_0 + b\lambda + \sum_{t=1}^{T-1} a_t U_{T-t} \right) \tilde{H}' \right] = a_T \mathbb{E}[Y_0\tilde{H}'] + b\mathbb{E}[\lambda\tilde{H}'].
\end{aligned}$$

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

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

Now consider

$$\begin{aligned}
&\mathbb{E}[Y_0\tilde{H}'] \big(\mathbb{E}[\tilde{H}\tilde{H}'] \big)^{-1} \mathbb{E}[\tilde{H}\lambda] \\
&= \frac{1}{\mathbb{E}[Y_0^2] - (\mathbb{E}[Y_0])} \begin{bmatrix} \mathbb{E}[Y_0] & \mathbb{E}[Y_0^2] \end{bmatrix} \begin{bmatrix} \mathbb{E}[Y_0^2] & -\mathbb{E}[Y_0] \\ -\mathbb{E}[Y_0] & 1 \end{bmatrix} \begin{bmatrix} \mathbb{E}[Y_0] \\ \mathbb{E}[Y_0^2] \end{bmatrix} \\
&= \mathbb{E}[Y_0\lambda].
\end{aligned}$$

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

$$\begin{aligned}
\rho_* &= \rho_0, \quad \gamma_*^2 = \gamma_0, \quad \phi_* = \big(\mathbb{E}[\tilde{H}\tilde{H}'] \big)^{-1} \mathbb{E}[\tilde{H}\lambda], \\
\omega_*^2 &= \mathbb{E}[\lambda^2] - \mathbb{E}[\lambda\tilde{H}'] \big(\mathbb{E}[\tilde{H}\tilde{H}'] \big)^{-1} \mathbb{E}[\tilde{H}\lambda]. \quad \blacksquare
\end{aligned} \tag{A.51}$$

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 λ_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 μ_+ and variance γ_+^2, and $a_{it} = 0$ if U_{it} is generated from the mixture component with mean μ_- variance γ_-^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

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

Conditional on b, the prior distribution is

$$\lambda|(Y_0, b) \sim N(\phi(Y_0, b), \underline{\Omega}),$$

and we obtain a posterior distribution of the form

$$\lambda|(Y_{0:T}, a_{1:T}, b) \sim N\big(\bar{\lambda}(a_{1:T}, b), \bar{\Omega}(a_{1:T})\big), \tag{A.52}$$

where

$$\begin{aligned} \bar{\Omega}(a_{1:T}) &= \Big(\underline{\Omega}^{-1} + \sum_{t=1}^{T}(\gamma^2(a_t))^{-1}\Big)^{-1} \\ \bar{\lambda}(a_{1:T}, b) &= \bar{\Omega}(a_{1:T})\Big(\underline{\Omega}^{-1}\phi(Y_0, b) + \sum_{t=1}^{T}(\gamma^2(a_t))^{-1}\tilde{Y}_t(a_t)\Big). \end{aligned}$$

The posterior probability of $a_t = 1$ conditional on $(\lambda, Y_{0:T})$ is given by

$$\begin{aligned} & \mathbb{P}\big(a_t = 1 | \lambda, Y_{0:T}) \\ = & \frac{p_u(\gamma_+)^{-1} \exp\left\{-\frac{1}{2\gamma_+^2}(\tilde{Y}_t(1) - \lambda)^2\right\}}{p_u(\gamma_+)^{-1} \exp\left\{-\frac{1}{2\gamma_+^2}(\tilde{Y}_t(1) - \lambda)^2\right\} + (1-p_u)(\gamma_-)^{-1} \exp\left\{-\frac{1}{2\gamma_-^2}(\tilde{Y}_t(0) - \lambda)^2\right\}}, \end{aligned} \tag{A.53}$$

And the posterior probability of $b = 1$ conditional on $(\lambda, Y_{0:T}, a_{1:T})$ is given by

$$\begin{aligned} & \mathbb{P}\big(b = 1 | \lambda, Y_{0:T}, a_{1:T}) \\ = & \frac{p_\lambda \exp\left\{-\frac{1}{2}\sum_{t=1}^T \frac{(\tilde{Y}_t(a_t) - \phi_+(Y_0))^2}{\underline{\Omega} + \gamma^2(a_t)}\right\}}{p_\lambda \exp\left\{-\frac{1}{2}\sum_{t=1}^T \frac{(\tilde{Y}_t(a_t) - \phi_+(Y_0))^2}{\underline{\Omega} + \gamma^2(a_t)}\right\} + (1-p_\lambda) \exp\left\{-\frac{1}{2}\sum_{t=1}^T \frac{(\tilde{Y}_t(a_t) - \phi_-(Y_0))^2}{\underline{\Omega} + \gamma^2(a_t)}\right\}}. \end{aligned} \tag{A.54}$$

The posterior mean $\mathbb{E}[\lambda | Y_{0:T}]$ can be approximated with the following Gibbs sampler. Generate a sequence of draws $\{\lambda^s, a_{1:T}^s, b^s\}_{s=1}^{N_{sim}}$ by iterating over the conditional distributions given in (A.52), (A.53), and (A.54). Denote $\bar{p}(a_{1:T}, b) = \mathbb{P}\big(b | \lambda, Y_{0:T}, a_{1:T})$, then,

$$\begin{aligned} \widehat{\mathbb{E}}[\lambda | Y_{0:T}] &= \frac{1}{N_{sim}} \sum_{s=1}^{N_{sim}} \sum_{b=0}^{1} \bar{p}(a_{1:T}^s, b) \bar{\lambda}(a_{1:T}^s, b), \qquad \text{(A.55)} \\ \widehat{\mathbb{V}}[\lambda | Y_{0:T}] &= \frac{1}{N_{sim}} \sum_{s=1}^{N_{sim}} \left(\bar{\Omega}(a_{1:T}^s) + \sum_{b=0}^{1} \bar{p}(a_{1:T}^s, b) \bar{\lambda}^2(a_{1:T}^s, b)\right) - \left(\widehat{\mathbb{E}}[\lambda | Y_{0:T}]\right)^2. \end{aligned}$$

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* $\mathrm{rank(W) = k_w}$. *Let Σ be a $T \times T$ matrix of rank T. Let $S = \Sigma W$. Then,* $\mathrm{rank(M_{S\otimes S}B) = T}$, *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 $(1,1),(T+2,2),(2T+3,3),...,(T^2,T)$ and zeros at the other positions. Notice that since Σ is full rank, $\mathrm{rank}(S) = \mathrm{rank}(\Sigma W) = \mathrm{rank}(W) = k_w$. If $\mathrm{rank}(S) = k_w$, then $\mathrm{rank}(S \otimes S) = k_w^2$. Since the rank of the projection matrix is the same as its trace, we have $\mathrm{rank}(M_{S\otimes S}) = tr(M_{S\otimes S}) = T^2 - k_w^2$.

By the spectral decomposition, we can decompose $M_{S\otimes S} = F\Lambda F'$, where F is a $T^2 \times T^2$ orthogonal matrix and Λ 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, $\mathrm{rank}(M_{S\otimes S}B) = \mathrm{rank}(F\Lambda F'B) = \mathrm{rank}(\Lambda F'B)$. Notice that $F'B$ is a $T^2 \times T$ matrix that collects the columns of F' in the positions of $1, T+2, 2T+3, ..., T^2$. Since the columns of F' are linearly independent, $\mathrm{rank}(F'B) = T$. 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 $\mathbb{E}\big[(W_{it}', X_{it}', Z_{it}')'(W_{it}', X_{it}', Z_{it}')\big]$ 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 α and ρ 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}', and U_{it}, respectively, for $t = 1, \ldots, T$. Define

$$\begin{aligned}\Sigma_i^{1/2}(\gamma) &= \mathrm{diag}\big(\sigma_1(h_i,\gamma_1), \ldots, \sigma_T(h_i,\gamma_T)\big),\\ S_i(\gamma) &= \Sigma_i^{-1/2}(\gamma)W_i, \quad M_i(\gamma) = I - S_i(S_i'S_i)^{-1}S_i'.\end{aligned}$$

Using this notation, we obtain

$$M_i(\tilde\gamma)\Sigma_i^{-1/2}(\tilde\gamma)\big(Y_i - X_i\rho - Z_i\alpha\big) = M_i(\tilde\gamma)S_i(\tilde\gamma)\lambda_i + M_i(\tilde\gamma)\Sigma_i^{-1/2}(\tilde\gamma)U_i = M_i(\tilde\gamma)V_i.$$

This leads to the conditional moment condition

$$\mathbb{E}\big[M_i(\tilde{\gamma})\Sigma_i^{-1/2}(\tilde{\gamma})\big(Y_i - X_i\rho - Z_i\alpha\big)\big(Y_i - X_i\rho - Z_i\alpha\big)'\Sigma_i^{-1/2}(\tilde{\gamma})M_i'(\tilde{\gamma}) - M_i(\tilde{\gamma})\big|H_i\big] = 0,$$

which can be rewritten as

$$M_i(\tilde{\gamma})\big(\Sigma_i^{-1/2}(\tilde{\gamma})\Sigma_i(\gamma)\Sigma_i^{-1/2}(\tilde{\gamma}) - I\big)M_i'(\tilde{\gamma}) = 0. \tag{A.56}$$

for each h_i. Taking expectations with respect to H_i and using Assumption 5.2(ii), we deduce that

$$\mathbb{E}\big[M_i(\tilde{\gamma})\big(\Sigma_i^{-1/2}(\tilde{\gamma})\Sigma_i(\gamma)\Sigma_i^{-1/2}(\tilde{\gamma}) - I\big)M_i'(\tilde{\gamma})\big] = 0. \tag{A.57}$$

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 ρ, α, and γ 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 γ argument in the remainder of the proof.

Because $S(\gamma)$, λ and V are independent conditional on H (and γ), 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 τ leads to

$$\begin{aligned} \frac{\partial^2}{\partial\tau\partial\tau'} \ln \Psi_{\tilde{Y}}(\tau|h) &= \frac{\partial^2}{\partial\tau\partial\tau'} (\ln \Psi_\lambda(S'\tau|h)) + \frac{\partial^2}{\partial\tau\partial\tau'} \ln \Psi_V(\tau) \\ &= S\left(\frac{\partial^2}{\partial\xi\partial\xi'} \ln \Psi_\lambda(S'\tau|h)\right) S' + \frac{\partial^2}{\partial\tau\partial\tau'} \ln \Psi_V(\tau). \end{aligned} \tag{A.60}$$

Using the assumption that the V_ts are independent over t, we can write

$$\ln \Psi_V(\tau) = \sum_{t=1}^{T} \ln \Psi_{V_t}(\tau_t),$$

where Ψ_{V_t} is the characteristic function of V_t. Then,

$$\begin{aligned} \operatorname{vec}\left(\frac{\partial^2}{\partial\tau\partial\tau'}\ln\Psi_V(\tau)\right) &= \operatorname{vec}\left(\operatorname{diag}\left(\frac{\partial^2}{\partial\tau_1^2}\ln\Psi_{V_1}(\tau_1),...,\frac{\partial^2}{\partial\tau_T^2}\ln\Psi_{V_T}(\tau_T)\right)\right) \\ &= B\left(\frac{\partial^2}{\partial\tau_1^2}\ln\Psi_{V_1}(\tau_1),...,\frac{\partial^2}{\partial\tau_T^2}\ln\Psi_{V_T}(\tau_T)\right)' \end{aligned} \quad \text{(A.61)}$$

for a suitably chosen matrix B. Let

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

Then,

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

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 $\frac{\partial}{\partial\tau_t}\ln\Psi_{V_t}(0) = 0$ ($\mathbb{E}(V_{it}) = 0$) 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

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

Taking second derivatives, we obtain

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

Because S is of full rank, we can identify

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

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

$$\hat{\lambda} = (S'S)^{-1}S'\tilde{Y} = \lambda + (S'S)^{-1}S'V. \quad \text{(A.66)}$$

Taking expectations yields

$$\mathbb{E}(\lambda|h) = \mathbb{E}[\hat{\lambda}|h], \tag{A.67}$$

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 α and ρ based on moment conditions in the general model (1). Under the model (1) we can remove the effect of λ_i with the following within projections:

$$\begin{aligned}
Y^*_{it} &= Y_{it} - \left(\sum_{s=t+1}^{T} Y_{is}W'_{is-1}\right)\left(\sum_{s=t+1}^{T} W_{is-1}W'_{is-1}\right)^{-1} W_{it-1} \\
X^*_{it-1} &= X_{it-1} - \left(\sum_{s=t+1}^{T} X_{is-1}W'_{is-1}\right)\left(\sum_{s=t+1}^{T} W_{is-1}W'_{is-1}\right)^{-1} W_{it-1} \\
Z^*_{it-1} &= Z_{it-1} - \left(\sum_{s=t+1}^{T} Z_{is-1}W'_{is-1}\right)\left(\sum_{s=t+1}^{T} W_{is-1}W'_{is-1}\right)^{-1} W_{it-1}
\end{aligned}$$

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

$$\mathbb{E}\left[\left(Y^*_{it} - \begin{bmatrix} \tilde{\rho}' & \tilde{\alpha}' \end{bmatrix}\begin{bmatrix} X^*_{it-1} \\ Z^*_{it-1} \end{bmatrix}\right)\begin{bmatrix} X'_{it-s-1} & Z'_{it-s-1} \end{bmatrix}\right] = 0 \tag{A.68}$$

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

$$\mathbb{E}\begin{bmatrix} X^*_{it-1}X'_{it-1} & X^*_{it-1}Z'_{it-1} \\ Z^*_{it-1}X'_{it-1} & Z^*_{it-1}Z'_{it-1} \end{bmatrix} \tag{A.69}$$

is full rank. Consider, for instance, the upper-left element. We can write

$$\begin{aligned}
&\mathbb{E}[X^*_{it-1}X'_{it-1}] \\
&= \mathbb{E}\left[\left(X_{it-1} - \left(\sum_{s=t+1}^{T} X_{is-1}W'_{is-1}\right)\left(\sum_{s=t+1}^{T} W_{is-1}W'_{is-1}\right)^{-1} W_{it-1}\right) X'_{it-1}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left(X_{it-1} - \left(\sum_{s=t+1}^{T} X_{is-1}W'_{is-1}\right)\left(\sum_{s=t+1}^{T} W_{is-1}W'_{is-1}\right)^{-1} W_{it-1}\right) X'_{it-1} \,\middle|\, W_i^{t:T-1}\right]\right] \\
&= \mathbb{E}[X_{it-1}X'_{it-1}] - \frac{1}{T-h}\left(\sum_{s=t+1}^{T} \mathbb{E}\Big[\mathbb{E}[X_{is-1}X_{it-1}|W_i^{t:T-1}] \right. \\
&\left. \times W'_{is-1}\left(\frac{1}{T-h}\sum_{s=t+1}^{T} W_{is-1}W'_{is-1}\right)^{-1} W_{it-1}\Big]\right) \\
&= \mathbb{E}[X_{it-1}X'_{it-1}] - \frac{1}{T-h}\sum_{s=t+1}^{T} \kappa_s \mathbb{E}[X_{is-1}X'_{it-1}] = I + II, \text{ say.}
\end{aligned}$$

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).

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

$$\text{PPNR} = 400\big(\text{NII} + \text{ONII} - \text{ONIE}\big)/\text{ASSETS},$$

where

NII	= Net Interest Income	BHCK 4074
ONII	= Total Non-Interest Income	BHCK 4079
ONIE	= Total Non-Interest Expenses	BHCK 4093 - C216 - C232
ASSETS	= Consolidated Assets	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 (τ, T). The following adjustment procedure that eliminates BHCs with missing observations and outliers is applied to each rolling sample (τ, 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 δ_{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 δ_{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 ($T = 6$)

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

Table A-2: Descriptive Statistics for Rolling Samples ($T = 6$)

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

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.