Supplement to “Semiparametric Estimation of Probability Weighting Functions Implicit in Option Prices”

H. Peter Boswijk
Amsterdam School of Economics, University of Amsterdam, and Tinbergen Institute
Jeroen Dalderop
Department of Economics, University of Notre Dame
Roger J. A. Laeven
Amsterdam School of Economics, University of Amsterdam, EURANDOM, and CentER
and
Niels Marijnen
Amsterdam School of Economics, University of Amsterdam, and Tinbergen Institute

September 23, 2025

Abstract

This text serves as an appendix to the paper “Semiparametric Estimation of Probability Weighting Functions Implicit in Option Prices”. For context, notation, and definitions, see that paper. This text provides all the proofs, technical results for the kernel estimator, a verification of the required technical conditions for common models, and additional simulation results.

A Proofs

The following lemma is used in several of the proofs that follow.

Lemma D. *Suppose the following conditions hold for some non-negative integer i and multi-index j:*

(i) $u'(r;\gamma)$ *is positive and* (i,j)*-times (continuously) differentiable on* $\mathbb{R}_{++} \times \bar{\Theta}$*, with* $\bar{\Theta} \subseteq \Theta$*;*

(ii) $Z'(P)$ *is positive and* $i + |j|$ *times (continuously) differentiable on* $(0,1)$*;*

(iii) $f_t(r)$ *is positive and* $\max\{i + |j| - 1, 0\}$ *times (continuously) differentiable a.s. on* $\mathbb{R}_{++}$*;*

(iv) $c_t(\gamma)$ *is finite and* j *times (continuously) differentiable a.s. on* $\bar{\Theta}$*.*

Then $g(v;\gamma)$ *is positive and (continuously) differentiable up to order* (i,j) *on* $(0,1) \times \bar{\Theta}$*.*

Proof. The function $U_t(r;\gamma)$ is differentiable w.r.t. r a.s., with derivative

$$U_t'(r;\gamma) = \frac{\partial}{\partial r}\left[c_t(\gamma)\int_0^r \frac{q_t(s)}{u'(s;\gamma)}\mathrm{d}s\right] = c_t(\gamma)\frac{q_t(r)}{u'(r;\gamma)} = \frac{c_t(\gamma)}{c_t(\gamma_0)}\frac{u'(r;\gamma_0)}{u'(r;\gamma)}Z'(F_t(r))f_t(r). \quad \text{(A.1)}$$

This derivative is $\mathbb{P}$-a.s. positive since all components on the RHS of (A.1) are. Moreover, it is (continuously) differentiable w.r.t. (v,γ) up to order (i,j). Therefore, $r \mapsto U_t(r;\gamma)$ has a strictly monotonic inverse function $v \mapsto U_t^{-1}(v;\gamma)$, whose derivative $\frac{\partial}{\partial v}U_t^{-1}(v;\gamma) = 1/U_t'(U_t^{-1}(v;\gamma);\gamma) > 0$ is (i,j) times (continuously) differentiable.

Using the law of iterated expectations, the CDF of the utility-adjusted PITs equals

$$G_\gamma(v) = \mathbb{P}(U_{t+1}(\gamma) \leq v) = \mathbb{P}(R_{t+1} \leq U_t^{-1}(v;\gamma)) = \mathbb{E}(F_t(U_t^{-1}(v;\gamma))). \quad \text{(A.2)}$$

Its density $g_\gamma(v) = \frac{\partial}{\partial v} G_\gamma(v)$ equals

$$\begin{aligned} g_\gamma(v) &= \mathbb{E}(f_t(U_t^{-1}(v;\gamma)) \frac{\partial}{\partial v} U_t^{-1}(v;\gamma)) \\ &= \mathbb{E}\left(\frac{c_t(\gamma_0)}{c_t(\gamma)} \frac{u'(U_t^{-1}(v;\gamma);\gamma)}{u'(U_t^{-1}(v;\gamma);\gamma_0)} \frac{1}{Z'(F_t(U_t^{-1}(v;\gamma)))} \right), \end{aligned} \tag{A.3}$$

inheriting positivity and order of (continuous) differentiability from c_t, u', Z', F_t, and $\frac{\partial}{\partial v} U_t^{-1}$. □

Proof of Lemma E. Let $\tilde{g}_\gamma$ be a family of densities on $(0,1)$ indexed by γ. Jensen's inequality yields that

$$\begin{aligned} \ell(\gamma, \tilde{g}_\gamma) - \ell(\gamma, g_\gamma) &= \mathbb{E}\left(\log \frac{\tilde{g}_\gamma(U_{t+1}(\gamma))}{g_\gamma(U_{t+1}(\gamma))} \right) \\ &\leq \log \mathbb{E}\left(\frac{\tilde{g}_\gamma(U_{t+1}(\gamma))}{g_\gamma(U_{t+1}(\gamma))} \right) = \log \int_0^1 \frac{\tilde{g}_\gamma(u)}{g_\gamma(u)} g_\gamma(u) \mathrm{d}u = 0, \end{aligned}$$

with equality if and only if $\tilde{g}_\gamma(U_{t+1}(\gamma)) = g_\gamma(U_{t+1}(\gamma))$ *a.s.* This requires that the densities are identical on any set in the support of $U_{t+1}(\gamma)$, and equal zero outside this support in order to integrate to one. Hence, $\tilde{g}_\gamma$ and g_γ are identical everywhere, so that g_γ is the unique maximizing function. □

Proof of Lemma I. Denote $\mathbb{E}_t(\cdot) := \mathbb{E}(\cdot|\mathcal{F}_t)$. For any γ, the law of iterated expectations and Jensen's inequality yield

$$\begin{aligned} \ell(\gamma) - \ell(\gamma_0) = \mathbb{E}\left(\log \frac{f_t(R_{t+1};\gamma, g_\gamma)}{f_t(R_{t+1};\gamma_0, g_{\gamma_0})} \right) &= \mathbb{E}\left(\mathbb{E}_t\left(\log \frac{f_t(R_{t+1};\gamma, g_\gamma)}{f_t(R_{t+1};\gamma_0, g_{\gamma_0})} \right) \right) \\ &\leq \mathbb{E}\left(\log \mathbb{E}_t\left(\frac{f_t(R_{t+1};\gamma, g_\gamma)}{f_t(R_{t+1};\gamma_0, g_{\gamma_0})} \right) \right) = 0, \end{aligned}$$

where the inequality holds with equality if and only if (3.9) holds.

By Lemma D, the density $g_\gamma(v)$ is positive so that $G_\gamma(u)$ is strictly monotonic for all

$\gamma \in \Theta$. Equation (3.9) is therefore equivalent to any of the following statements:

$$\begin{aligned}
f_t(r;\gamma,g_\gamma) = f_t(r;\gamma_0,g_{\gamma_0}) \text{ a.s. } \forall r \;&\Leftrightarrow\; F_t(r;\gamma,g_\gamma) = F_t(r;\gamma_0,g_{\gamma_0}) \text{ a.s. } \forall r \\
&\Leftrightarrow\; F_t(F_t^{-1}(v);\gamma,g_\gamma) = v \text{ a.s. } \forall v \in (0,1) \\
&\Leftrightarrow\; G_\gamma\left(U_t(F_t^{-1}(v);\gamma)\right) = v \text{ a.s. } \forall v \in (0,1) \\
&\Leftrightarrow\; U_t(F_t^{-1}(v);\gamma) = G_\gamma^{-1}(v) \text{ a.s. } \forall v \in (0,1).
\end{aligned}$$

The LHS of the final equation describes the conditional quantile function of $U_{t+1}(\gamma)$. For the final statement to hold, it should be constant over time for any u. However, the quantile density function equals

$$\frac{\partial}{\partial v}U_t(F_t^{-1}(v);\gamma) = c_t(\gamma)\frac{\partial}{\partial v}\int_0^{F_t^{-1}(v)} \frac{q_t(r)}{u'(r;\gamma)}\mathrm{d}r = \frac{c_t(\gamma)}{c_t(\gamma_0)}\frac{u'(F_t^{-1}(v);\gamma_0)}{u'(F_t^{-1}(v);\gamma)}Z'(v),$$

using (2.3) in the second equation, and its log-derivative equals

$$\frac{\partial}{\partial v}\log(\frac{\partial}{\partial v}U_t(F_t^{-1}(v);\gamma)) = \frac{1}{f_t(F_t^{-1}(v))}\left(\frac{u''(F_t^{-1}(v);\gamma_0)}{u'(F_t^{-1}(v);\gamma_0)} - \frac{u''(F_t^{-1}(v);\gamma)}{u'(F_t^{-1}(v);\gamma)}\right) + \frac{Z''(v)}{Z'(v)}. \quad \text{(A.4)}$$

The RHS of (A.4) can only be constant over time for a given $v \in (0,1)$ if (i) $\frac{u''(F_t^{-1}(v);\gamma_0)}{u'(F_t^{-1}(v);\gamma_0)} = \frac{u''(F_t^{-1}(v);\gamma)}{u'(F_t^{-1}(v);\gamma)}$ a.s. or if (ii) $f_t(F_t^{-1}(v)) = a(v)\left(\frac{u''(F_t^{-1}(v);\gamma_0)}{u'(F_t^{-1}(v);\gamma_0)} - \frac{u''(F_t^{-1}(v);\gamma)}{u'(F_t^{-1}(v);\gamma)}\right)$ a.s. for some non-zero function $a(v)$.

Let $\mathcal{V}(\bar{r})$ be the support of $F_t(\bar{r})$. Then case (i) cannot hold for any $v \in \mathcal{V}(\bar{r})$ since I*(ii)* implies that $ARA(r;\gamma) \neq ARA(r;\gamma_0)$ in a neighborhood of $\bar{r}$. Case (ii) implies for any $(v_l, v_h) \in \mathcal{V}(\bar{r})$ with $v_l < v_h$ that

$$\log\frac{\frac{\partial}{\partial v}U_t(F_t^{-1}(v_h);\gamma)}{\frac{\partial}{\partial v}U_t(F_t^{-1}(v_l);\gamma)} = \int_{v_l}^{v_h}\left(\frac{1}{a(v)} + \frac{Z''(v)}{Z'(v)}\right)\mathrm{d}v. \quad \text{(A.5)}$$

The constant RHS term implies a one-to-one relation between $F_t^{-1}(v_l)$ and $F_t^{-1}(v_h)$, which is ruled out by I*(iii)*. Therefore, for some $v \in \mathcal{V}(\bar{r})$ neither (i) nor (ii) holds, so that (A.4) is non-deterministic. As a result, none of the statements equivalent to (3.9) hold, and $\ell(\gamma) < \ell(\gamma_0)$ for $\gamma \neq \gamma_0$. □

Proof of Lemma I.* The censored log-likelihood population criterion is given by

$$\ell^*(\gamma) := \mathbb{E}\left(\log\left(G_\gamma(U_t^l(\gamma))\right)^{1_{t+1}^l}\left(c_t(\gamma)/u'(R_{t+1};\gamma)g_\gamma\left(U_{t+1}(\gamma)\right)\right)^{1_{t+1}^m}\left(1-G_\gamma(U_t^u(\gamma))\right)^{1_{t+1}^u}\right).$$

It holds that

$$\begin{aligned}
&\ell^*(\gamma) - \ell^*(\gamma_0) \\
&\leq \mathbb{E}\left(\log \mathbb{E}_t\left(\left(\frac{F_t(R_{t,l}^*;\gamma)}{F_t(R_{t,l}^*;\gamma_0)}\right)^{1_{t+1}^l}\left(\frac{f_t(R_{t+1};\gamma,g_\gamma)}{f_t(R_{t+1};\gamma_0,g_{\gamma_0})}\right)^{1_{t+1}^m}\left(\frac{1-F_t(R_{t,u}^*;\gamma)}{1-F_t(R_{t,u}^*;\gamma_0)}\right)^{1_{t+1}^u}\right)\right) \\
&= \mathbb{E}\left(\log\left(F_t(R_{t,l}^*;\gamma) + \int_{R_{t,l}^*}^{R_{t,u}^*} f_t(r;\gamma,g_\gamma)\mathrm{d}r + (1-F_t(R_{t,u}^*;\gamma))\right)\right) = 0,
\end{aligned}$$

where the Jensen's inequality holds with equality if and only if

$$F_t(R_{t,i}^*;\gamma) = F_t(R_{t,i}^*;\gamma_0) \text{ for } i \in \{l,u\} \text{ and } 1_{t+1}^m\left(f_t(R_{t+1};\gamma,g_\gamma) - f_t(R_{t+1};\gamma_0,g_{\gamma_0})\right) = 0 \quad \text{a.s.} \tag{A.6}$$

By the equivalence statements in the proof of Lemma I, this requires that $F_t(F_t^{-1}(u);\gamma,g_\gamma) = u$ for all $u \in (F_t(R_{t,l}^*), F_t(R_{t,u}^*))$ a.s. Since $(R_{t,l}^*, R_{t,u}^*)$ contains $\bar{r}$ with positive probability, case (ii) in the proof of Lemma I then implies that (A.5) holds for any pair $(v_l, v_h) \in \operatorname{supp}(F_t(\bar{r}) \mid R_{t,l}^* \leq \bar{r} \leq R_{t,u}^*)$. However, the resulting one-to-one relation between $F_t^{-1}(v_l)$ and $F_t^{-1}(v_h)$ cannot hold for the pair (v_1, v_2) in the statement of Lemma I*. As a result, (A.6) cannot hold, and $\ell^*(\gamma) < \ell^*(\gamma_0)$ for $\gamma \neq \gamma_0$. □

Proof of Proposition C. If (i) for every neighborhood Θ_0 of γ_0, $\max_{\gamma\in\Theta\setminus\Theta_0} \ell^*(\gamma) < \ell^*(\gamma_0)$ and (ii) $\sup_{\gamma\in\Theta}\left|\ell_T^*(\gamma) - \ell^*(\gamma)\right| \xrightarrow{p} 0$, then $\hat{\gamma} \xrightarrow{p} \gamma_0$ (e.g., Andrews, 1994, Lemma A-1). The identification condition (i) follows from Lemma I* given Assumption I. To establish the uniform convergence condition (ii), write

$$\sup_{\gamma\in\Theta}\left|\ell_T^*(\gamma)-\ell^*(\gamma)\right| \leq \sup_{\gamma\in\Theta}\left|\ell_T^*(\gamma)-\ell_T^*(\gamma,g_\gamma)\right| + \sup_{\gamma\in\Theta}\left|\ell_T^*(\gamma,g_\gamma)-\ell^*(\gamma)\right|. \tag{A.7}$$

The first term on the RHS of (A.7) is bounded by

$$\begin{aligned}
&|\ell_T^*(\gamma)-\ell_T^*(\gamma,g_\gamma)| \\
&\quad \leq \frac{1}{T}\sum_{t=1}^{T} 1_{t+1}^{l}\left|\log\frac{\widehat{G}_\gamma(U_t^l(\gamma))}{G_\gamma(U_t^l(\gamma))}\right| + 1_{t+1}^{m}\left|\log\frac{\widehat{g}_\gamma(U_{t+1}(\gamma))}{g_\gamma(U_{t+1}(\gamma))}\right| + 1_{t+1}^{u}\left|\log\frac{1-\widehat{G}_\gamma(U_t^u(\gamma))}{1-G_\gamma(U_t^u(\gamma))}\right|.
\end{aligned}$$

Since g_γ is positive and continuous by Lemma D, the constant $c_g := \min_{\gamma\in\Theta}\min_{v\in[v^*,1-v^*]} g_\gamma(v)$ is positive. Using $\left|\log\frac{x}{y}\right| \leq \frac{1}{x\wedge y}|y-x|$ for any $x,y>0$, uniform convergence of $\widehat{g}$ therefore implies that for any constant $0<\delta<1$, with probability approaching (w.p.a.) 1

$$\sup_{\gamma\in\Theta}\frac{1}{T}\sum_{t=1}^{T} 1_{t+1}^{m}\left|\log\frac{\widehat{g}_\gamma(U_{t+1}(\gamma))}{g_\gamma(U_{t+1}(\gamma))}\right| \leq \frac{1}{\delta c_g}\sup_{\gamma\in\Theta}\left\|\widehat{g}_\gamma-g_\gamma\right\|_{\infty,v^*} \xrightarrow{p} 0.$$

Similarly, consider the positive constant $c_G^l := \min_{\gamma\in\Theta} G_\gamma(v^*)$. Then w.p.a. 1

$$\sup_{\gamma\in\Theta}\frac{1}{T}\sum_{t=1}^{T} 1_{t+1}^{l}\left|\log\frac{\widehat{G}_\gamma(U_t^l(\gamma))}{G_\gamma(U_t^l(\gamma))}\right| \leq \frac{1}{\delta c_G^l}\sup_{\gamma\in\Theta}\left\|\widehat{G}_\gamma-G_\gamma\right\|_{\infty,v^*} \xrightarrow{p} 0,$$

where the final step follows from the uniform convergence of $\widehat{g}$ and the uniform in γ consistency of $\widehat{G}$ at v^*. The same result holds for the third term for the positive constant $c_G^u := 1-\max_{\gamma\in\Theta} G_\gamma(1-v^*)$.

For the second term on the RHS of (A.7), pointwise convergence $\ell_T^*(\gamma,g_\gamma) \xrightarrow{p} \ell^*(\gamma)$ follows from the ergodic theorem (Davidson, 1994, Theorem 13.12), as the summands of $\ell_T^*(\gamma,g_\gamma)$ are measurable in the stationary and ergodic sequence $\{R_{t+1},X_t\}$ and integrable for all $\gamma\in\Theta$. Moreover, the summands are continuous in γ a.s. by Lemma D and dominated by an integrable function, since

$$
\begin{aligned}
&1^m_{t+1}\left|\log c_t(\gamma)-\log u'(R_{t+1};\gamma)+\log g_\gamma(U_{t+1}(\gamma))\right| \\
&\qquad\le \sup_{\gamma\in\Theta}\Big(\left|\log c_t(\gamma)\right|+\left|\log u'(R_{t+1};\gamma)\right|+1^m_{t+1}\left|\log g_\gamma(U_{t+1}(\gamma))\right|\Big).
\end{aligned}
$$

The first two of these terms are integrable by assumption, while continuity of g_γ implies that the third term is bounded on $[v^*,1-v^*]$, uniformly over γ. Since Θ is compact, $\ell^*_T(\gamma)$ thus satisfies a uniform weak law of large numbers, e.g. Andrews (1992, Theorem 4). □

Proof of Lemma A. The remainder term equals

$$
\begin{aligned}
\xi_T &:= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(s(R_{t+1},q_t,\gamma_0,\widehat{g})-s(R_{t+1},q_t,\gamma_0,g_0)\right)\\
&=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Bigg[1^m_{t+1}\left(S(R_{t+1},q_t,\widehat{g})-S(R_{t+1},q_t,g_0)\right)\\
&\quad+1^l_{t+1}\left(S(R^*_{t,l},q_t,\widehat{G})-S(R^*_{t,l},q_t,G_0)\right)+1^u_{t+1}\left(S(R^*_{t,u},q_t,1-\widehat{G})-S(R^*_{t,u},q_t,1-G_0)\right)\Bigg].
\end{aligned}
$$

First, we show that for all g with $\|g_{\gamma_0}-g_{0,\gamma_0}\|_{\infty,v^*}$ small enough and (r,q) such that $R^*_l(q,v^*)\le r\le R^*_u(q,v^*)$,

$$
\|S(r,q,g)-S(r,q,g_0)-D^m_0(r,q,g-g_0)\|\le b(r,q)\|g-g_0\|^2, \tag{A.8}
$$

where $\mathbb{E}\left(1^m_{t+1}b_j(R_{t+1},q_t)\right)<\infty$ for $j=1,\dots,k$. Here $D^m_0(r,q,g):=D_S(r,q,g_0)[g]$, with D_S the pathwise derivative of S in g, defined for any direction $\bar{g}$ as

$$
\begin{aligned}
D_S(r,q,g)[\bar{g}] &:= \lim_{\tau\to 0}\frac{1}{\tau}\left(S(r,q,g+\tau\bar{g})-S(r,q,g)\right)\\
&=\frac{\partial}{\partial\tau}\frac{\dot{g}_{\gamma_0}(U_0)+\tau\dot{\bar{g}}_{\gamma_0}(U_0)+\left(g'_{\gamma_0}(U_0)+\tau\bar{g}'_{\gamma_0}(U_0)\right)\dot{U}_0}{g_{\gamma_0}(U_0)+\tau\bar{g}_{\gamma_0}(U_0)}\bigg|_{\tau=0}\\
&=\frac{1}{g_{\gamma_0}(U)}\dot{\bar{g}}_{\gamma_0}(U)+\frac{\dot{U}}{g_{\gamma_0}(U)}\bar{g}'_{\gamma_0}(U)-\frac{\dot{g}_{\gamma_0}(U)+g'_{\gamma_0}(U)\dot{U}}{g_{\gamma_0}(U)^2}\bar{g}_{\gamma_0}(U),
\end{aligned}
$$

which is linear in $\bar{g}$.

Using the relation $\frac{\tilde{a}}{\tilde{b}} - \frac{a}{b} = \frac{1}{b}\left(1 - \frac{1}{\tilde{b}}(\tilde{b} - b)\right)\left(\tilde{a} - a - \frac{a}{b}(\tilde{b} - b)\right)$, the remainder of the linearization of $\frac{\tilde{a}}{\tilde{b}}$ around $\frac{a}{b}$ takes the form $\frac{1}{b\tilde{b}}\left(\tilde{b} - b\right)\left(\tilde{a} - a - \frac{a}{b}(\tilde{b} - b)\right)$. Therefore,

$$
\begin{aligned}
&\|S(r,q,g) - S(r,q,g_0) - D_0(r,q,g-g_0)\| \\
&= \left\| \frac{g_{\gamma_0}(U_0) - g_{0,\gamma_0}(U_0)}{g_{\gamma_0}(U_0) g_{0,\gamma_0}(U_0)} \left(\dot{g}_{\gamma_0}(U_0) - \dot{g}_{0,\gamma_0}(U_0) + (g'_{\gamma_0}(U_0) - g'_{0,\gamma_0}(U_0))\dot{U}_0 - S(R,q,g_0)\left(g_{\gamma_0}(U_0) - g_{0,\gamma_0}(U_0)\right)\right) \right\| \\
&\leq \frac{2 + \|S(R,q,g_0)\| + \|\dot{U}_0\|}{g_{\gamma_0}(U_0) g_{0,\gamma_0}(U_0)} \left(\left(g_{\gamma_0}(U_0) - g_{0,\gamma_0}(U_0)\right)^2 + \|\dot{g}_{\gamma_0}(U_0) - \dot{g}_{0,\gamma_0}(U_0)\|^2 + \left(g'_{\gamma_0}(U_0) - g'_{0,\gamma_0}(U_0)\right)^2 \right).
\end{aligned}
$$

As a result, (A.8) holds with $b(r,q) := \frac{2 + \|S(r,q,g_0)\| + \|\dot{U}_0\|}{\delta c_g g_{0,\gamma_0}(U_0)} \leq \frac{2 + \|\dot{g}_{0,\gamma_0}(U_0) + g'_{0,\gamma_0}(U_0)\dot{U}_0\|/c_g + \|\dot{U}_0\|}{\delta c_g^2}$, with δ and c_g as in the proof of Proposition C. The conditions in A*(i)* imply that g_{γ_0}, g'_{γ_0}, and $\dot{g}_{\gamma_0}$ are continuous and g_{γ_0} is positive on $(0,1)$ by Lemma D. Therefore $\dot{g}_{0,\gamma_0}$ and g'_{0,γ_0} are bounded on $[v^*, 1-v^*]$, so that the integrability of $b(r,q)$ follows from that of $\dot{U}_0$.

Similar bounds on the scores of the trimming terms can be obtained by replacing g by G in the above linearization. In particular, for all G with $\|G - G_0\|$ small enough,

$$
\|S(R_l^*(q,v^*), q, G) - S(R_l^*(q,v^*), q, G_0) - D_0^l(q, G - G_0)\| \leq B^l(q)\|G - G_0\|^2, \tag{A.9}
$$

where

$$
D_0^l(q,\bar{G}) := D_S(R_l^*(q,v^*), q, G_0)[\bar{G}], B^l(q) := \frac{1}{\delta c_G^l}\left(2 + \|\dot{G}_{0,\gamma_0}(U_0^l)\| + |1 + g_{0,\gamma_0}(U_0^l)|\|\dot{U}_0^l\|\right),
$$

with $U_0^i := U^i(\gamma_0)$ where $U^i(\gamma) := U(R_i^*(q,v^*); q, \gamma)$ and $\dot{U}_0^i$ defined analogously for $i \in \{l, u\}$. The bound for the right tail follows symmetrically by replacing G with $1-G$ and superscript l with u. Since U_0^l and U_0^u are in $[v^*, 1-v^*]$ a.s., the integrability of $B^l(q_t)$ and $B^u(q_t)$ follows from that of $\dot{U}_0$.

Define the linear score approximator

$$D_0(R_{t+1}, q_t, \bar{g}) = 1^m_{t+1} D^m_0(R_{t+1}, q_t, \bar{g}) + 1^l_{t+1} D^l_0(q_t, \bar{G}) + 1^u_{t+1} D^u_0(q_t, \bar{G}).$$

Then, for $j = 1, \ldots, k$ w.p.a. 1

$$\begin{aligned}\|\xi_{Tj}\| \leq & \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} D_{0,j}(R_{t+1}, q_t, \widehat{g} - g_0) \right\| + \left| \frac{1}{T} \sum_{t=1}^{T} 1^m_{t+1} b_j(R_{t+1}, q_t) \right| \sqrt{T} \|\widehat{g} - g_0\|^2 \\ & + \left| \frac{1}{T} \sum_{t=1}^{T} \left(1^l_{t+1} B^l_j(q_t) + 1^u_{t+1} B^u_j(q_t) \right) \right| \sqrt{T} \|\widehat{G} - G_0\|^2.\end{aligned}$$

By Assumption A*(ii)*, $\|\widehat{g} - g_0\|$ and $\|\widehat{G} - G_0\| \leq \|\widehat{g}_{\gamma_0} - g_{0,\gamma_0}\|_{\infty, v^*} + |\widehat{G}_{\gamma_0}(v^*) - G_{0,\gamma_0}(v^*)| + \|\widehat{\dot{G}}_{\gamma_0}(v^*) - \dot{G}_{0,\gamma_0}(v^*)\|$ are $o_p(T^{-\frac{1}{4}})$. By Assumption C*(iv)* and the continuity of $u'(\cdot; \gamma_0)$ and g_{γ_0}, $q_t(r)$ is a measurable function of (R, S). The ergodic theorem and Assumption A*(iii)* therefore imply that $\xi_T = o_p(1)$. □

Proof of Proposition N. This follows by combining Proposition C, Lemma A, Assumption N, and the mean-value expansion in (3.12). We show that $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} s(R_{t+1}, q_t, \gamma_0, g_0) \xrightarrow{d} N(0, V)$. The score of each observation equals

$$\begin{aligned}& s(R_{t+1}, q_t, \gamma, g) \\ & = 1^m_{t+1} \frac{\partial}{\partial \gamma} \log f_t(R_{t+1}; \gamma, g_\gamma) + 1^l_{t+1} \frac{\partial}{\partial \gamma} \log G_\gamma(U^l_t(\gamma)) + 1^u_{t+1} \frac{\partial}{\partial \gamma} \log(1 - G_\gamma(U^u_t(\gamma))).\end{aligned}$$

Under specification (3.2), $\mathbb{E}_t s(R_{t+1}, q_t, \gamma_0, g_0) = 0$ for all t. Therefore, applying a central limit theorem for martingale difference sequences to the scores at the true parameter yields the asymptotic normality. Here the covariance matrix V exists due to the moment conditions in N*(i)* and the boundedness of $\frac{\dot{G}_{\gamma_0}(v)}{G_{\gamma_0}(v)}$ on the trimmed domain. Furthermore, the uniform convergence of the Jacobian to the family of matrices $M(\gamma) := \partial \mathbb{E}^{\mathbb{P}} \left(s(R_{t+1}, q_t, \gamma, g_0) \right) / \partial \gamma$

follows from the triangle inequality

$$\begin{aligned}
&\sup_{\gamma\in\mathcal{N}_0}\left\|\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\ell_T^*(\gamma)-M(\gamma)\right\| \\
&\leq \sup_{\gamma\in\mathcal{N}_0}\left\|\frac{\partial^2}{\partial\gamma\partial\gamma^\top}(\ell_T^*(\gamma)-\ell_T^*(\gamma,g_{0,\gamma}))\right\|+\sup_{\gamma\in\mathcal{N}_0}\left\|\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\ell_T^*(\gamma,g_{0,\gamma})-M(\gamma)\right\|.
\end{aligned} \tag{A.10}$$

Define the norm $\|g\|_{\mathcal{N}_0,v^*,l}=\max_{i+|j|\leq l}\sup_{\gamma\in\mathcal{N}_0}\|\frac{\partial^{i+|j|}}{\partial v^i\partial\gamma^j}g\|_{\infty,v^*}$ for non-negative integers (i,l) and multi-index j with sum $|j|$. To control the first term on the RHS of (A.10), we show there is some integrable $\tilde{b}$ such that $\sup_{\gamma\in\mathcal{N}_0}\|\frac{\partial}{\partial\gamma}s(R_{t+1},q_t,\gamma,g)-\frac{\partial}{\partial\gamma}s(R_{t+1},q_t,\gamma,g_0)\|\leq \tilde{b}(R_{t+1},q_t)\|g-g_0\|_{\mathcal{N}_0,v^*,2}$ for small enough $\|g-g_0\|_{\mathcal{N}_0,v^*,2}$. The partial derivative of the score, using the twice continuous differentiability of $g_\gamma(v)$ by Lemma D and Assumption N*(ii)*, equals

$$\begin{aligned}
&\frac{\partial}{\partial\gamma}s(R_{t+1},q_t,\gamma,g) \\
&=1_{t+1}^m\left(\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\log c_t(\gamma)-\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\log u'(R_{t+1};\gamma)+\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\log g_\gamma(U_{t+1}(\gamma))\right) \\
&\quad+1_{t+1}^l\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\log G_\gamma(U_t^l(\gamma))+1_{t+1}^u\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\log(1-G_\gamma(U_t^u(\gamma))).
\end{aligned}$$

For $R_l^*(q,v^*)\leq r\leq R_u^*(q,v^*)$, we have $\left\|\frac{\partial}{\partial\gamma}s(r,q,\gamma,g)-\frac{\partial}{\partial\gamma}s(r,q,\gamma,g_0)\right\|\leq(*)+(**)$, where

$$\begin{aligned}
(*)&=\left\|\frac{\frac{\partial^2}{\partial\gamma\partial\gamma^\top}g_\gamma\left(U(\gamma)\right)}{g_\gamma\left(U(\gamma)\right)}-\frac{\frac{\partial^2}{\partial\gamma\partial\gamma^\top}g_{0,\gamma}\left(U(\gamma)\right)}{g_{0,\gamma}\left(U(\gamma)\right)}\right\| \\
&\leq\frac{1}{g_\gamma(U(\gamma))}\left\|\frac{\partial^2}{\partial\gamma\partial\gamma^\top}\left(g_\gamma\left(U(\gamma)\right)-g_{0,\gamma}\left(U(\gamma)\right)\right)\right\| \\
&\qquad+\frac{\|\frac{\partial^2}{\partial\gamma\partial\gamma^\top}g_{0,\gamma}\left(U(\gamma)\right)\|}{g_\gamma(U(\gamma))g_{0,\gamma}(U(\gamma))}\left|g_\gamma\left(U(\gamma)\right)-g_{0,\gamma}\left(U(\gamma)\right)\right| \\
&\leq\frac{1+2\|\dot{U}(\gamma)\|+\|\ddot{U}(\gamma)\|}{g_\gamma(U(\gamma))}\|g-g_0\|_{\mathcal{N},v^*,2}+\frac{\|\frac{\partial^2}{\partial\gamma\partial\gamma^\top}g_{0,\gamma}\left(U(\gamma)\right)\|}{g_\gamma(U(\gamma))g_{0,\gamma}(U(\gamma))}\|g-g_0\|_{\mathcal{N},v^*,0},
\end{aligned}$$

and

$$
\begin{aligned}
(**) &= \left\| \frac{\frac{\partial}{\partial\gamma} g_\gamma(U(\gamma)) \frac{\partial}{\partial\gamma} g_\gamma(U(\gamma))^\top}{g_\gamma(U(\gamma))^2} - \frac{\frac{\partial}{\partial\gamma} g_{0,\gamma}(U(\gamma)) \frac{\partial}{\partial\gamma} g_{0,\gamma}(U(\gamma))^\top}{g_{0,\gamma}(U(\gamma))^2} \right\| \\
&\leq \left\| \frac{\frac{\partial}{\partial\gamma} g_\gamma\left(U(\gamma)\right)}{g_\gamma\left(U(\gamma)\right)} - \frac{\frac{\partial}{\partial\gamma} g_{0,\gamma}\left(U(\gamma)\right)}{g_{0,\gamma}\left(U(\gamma)\right)} \right\| \cdot \left\| \frac{\frac{\partial}{\partial\gamma} g_\gamma\left(U(\gamma)\right)}{g_\gamma\left(U(\gamma)\right)} + \frac{\frac{\partial}{\partial\gamma} g_{0,\gamma}\left(U(\gamma)\right)}{g_{0,\gamma}\left(U(\gamma)\right)} \right\|^\top \\
&\leq \left(\frac{1 + \|\dot{U}(\gamma)\| + \frac{1}{g_\gamma(U(\gamma))} \|\dot{g}_\gamma(U(\gamma)) + g'_\gamma(U(\gamma))\dot{U}(\gamma)\|}{g_\gamma(U(\gamma))} \|g - g_0\|_{\mathcal{N}_0, v^*, 1} + o_P\left(\|g - g_0\|_{\mathcal{N}_0, v^*, 1}\right) \right) \\
&\quad \cdot 2 \sup_{\gamma \in \mathcal{N}_0} \frac{\|\dot{g}_\gamma(U(\gamma)) + g'_\gamma(U(\gamma))\dot{U}(\gamma)\|}{g_\gamma(U(\gamma))},
\end{aligned}
$$

using the same asymptotic linearization as in the proof of Lemma A. The first two derivatives of g in both arguments are continuous and therefore bounded on the trimmed domain, so that integrability of $\tilde{b}$ follows from that of $\dot{U}$ and $\ddot{U}$.

When $r < R_l^*(q, v^*)$ or $r > R_u^*(q, v^*)$, $\left\| \frac{\partial}{\partial\gamma} s(r, q, \gamma, g) - \frac{\partial}{\partial\gamma} s(r, q, \gamma, g_0) \right\|$ is bound using the same expansion in the order of $\|G - G_0\|_{\mathcal{N}_0, v^*, 2}$, which vanishes when $\|g - g_0\|_{\mathcal{N}_0, v^*, 2}$ and $\sup_{\gamma \in \mathcal{N}_0} |\frac{\partial^{|j|}}{\partial\gamma^j}(G_\gamma(v^*) - G_{0,\gamma}(v^*))|$ vanish for all multi-indices j with $|j| \leq 2$.

The second term on the RHS of (A.10) vanishes by dominated convergence. In particular, the continuous differentiability assumptions ensure that $\frac{\partial}{\partial\gamma} s(R_{t+1}, q_t, \gamma, g_0)$ is continuous in γ at γ_0 a.s., while the uniform integrability assumptions assure that $\mathbb{E}\left(\sup_{\gamma \in \mathcal{N}_{\gamma_0}} \|\frac{\partial}{\partial\gamma} s(R_{t+1}, q_t, \gamma, g_0)\|\right) < \infty$. $\square$

B Convergence of the Kernel Estimator

The following lemma establishes the stochastic equicontinuity condition A*(iii)* for the kernel estimator (3.4) using a result for dependent U-statistics.

Lemma SE. *Suppose A(i), C(iv), and the following conditions hold for some $\delta' > 0$ and $\delta > \delta'$:*

(i) (R_{t+1}, X_t) *is* β*-mixing with* $\beta(\tau) = O\left(\tau^{-\frac{2+\delta'}{\delta'}}\right)$ *when* $\tau \to \infty$*;*

(ii) $\mathbb{E}(\|\dot{U}_0\|^{2+\delta} \mid U_0 = v)$ *is continuous for* $v \in (0,1)$*,* $\mathbb{E}\left(\|\dot{U}_{0,s} - \dot{U}_{0,t}\|^{2+\delta} \mid U_{0,s} = u, U_{0,t} = v\right)$ *is continuously differentiable at* $u = v \in (0,1)$ *for all* $s \neq t$*;*

(iii) $\int K(z)\mathrm{d}z = 1,\ \int K(z)z\mathrm{d}z = 0,\ K(0) < \infty,\ \int |K(z)|^{2+\delta}\mathrm{d}z + \int |zK'(z)|^{2+\delta}\mathrm{d}z < \infty$*;*

(iv) $T^{\frac{\delta-\delta'}{\delta'}} h^{1+\delta} \to \infty$ *and* $Th^2 \to \infty$*.*

Then Assumption A(iii) holds for the kernel density (3.4).

Assumption SE*(i)* stipulates that the mixing coefficients vanish at least as fast as $1/\tau$ when the time between observations τ becomes large. A larger value of δ' allows for a slower rate within this class, though by SE*(ii)*-*(iii)* requires higher moments of the PIT-derivative $\dot{U}_0$ and kernel function K to exist. Meanwhile, Assumption SE*(iv)* ensures that the bandwidth does not vanish too quickly when the sample period increases.

Proof of Proposition SE. First, we conjecture and verify that $\int D_0(r, q, \widehat{g} - g_0)\mathrm{d}F(r, q) = 0$, where F is the joint physical distribution of (R_{t+1}, q_t). The score approximator is equal to

$$D_0(r, q, \bar{g}) = \begin{cases} \frac{\dot{\bar{G}}_{\gamma_0}(U_0^l)}{G_{0,\gamma_0}(U_0^l)} + \frac{\partial}{\partial\gamma} \frac{\bar{G}_{\gamma_0}(U^l(\gamma))}{G_{0,\gamma}(U^l(\gamma))}\Big|_{\gamma=\gamma_0} & \text{for } r < R_l^*(q, v^*); \\ \frac{\dot{\bar{g}}_{\gamma_0}(U_0)}{g_{0,\gamma_0}(U_0)} + \frac{\partial}{\partial\gamma} \frac{\bar{g}_{\gamma_0}(U(\gamma))}{g_{0,\gamma}(U(\gamma))}\Big|_{\gamma=\gamma_0} & \text{for } R_l^*(q, v^*) \leq r \leq R_u^*(q, v^*); \\ \frac{-\dot{\bar{G}}_{\gamma_0}(U_0^u)}{1-G_{0,\gamma_0}(U_0^u)} + \frac{\partial}{\partial\gamma} \frac{1-\bar{G}_{\gamma_0}(U^u(\gamma))}{1-G_{0,\gamma}(U^u(\gamma))}\Big|_{\gamma=\gamma_0} & \text{for } r > R_u^*(q, v^*), \end{cases}$$

where $U(\gamma) := U(r; q, \gamma)$. For $\bar{g} = \widehat{g} - g_0$, integrating the lower, middle, and upper parts with respect to the probability distribution cancels out the terms $\frac{1}{G_0}$, $\frac{1}{g_0}$ and $\frac{1}{1-G_0}$, respectively, to yield

$$
\begin{aligned}
\int D_0(r,q,\widehat{g}-g_0)\mathrm{d}F(r,q) &= \frac{1}{T}\sum_{t=1}^{T}\dot{U}_{0,t+1}\int \left(K_h'(U_{0,t+1}-v)-\dot{g}_0(v)\right)\mathrm{d}v \\
&\quad + \frac{1}{T}\sum_{t=1}^{T}\frac{\partial}{\partial\gamma}\int \left(K_h(U_{t+1}(\gamma)-v)-g_0(v;\gamma)\right)\mathrm{d}v\Big|_{\gamma=\gamma_0} \\
&= \frac{1}{Th}\sum_{t=1}^{T}\dot{U}_{0,t+1}\int K'(z)\mathrm{d}z-\frac{\partial}{\partial\gamma}\int g_0(v;\gamma)\mathrm{d}v\Big|_{\gamma=\gamma_0}-0=0,
\end{aligned}
$$

using that K and g_0 integrate to one.

Since D_0 is linear in $\bar{g}$, decompose $D_0(R_{t+1},q_t,\widehat{g}-g_0) = D_0(R_{t+1},q_t,\widehat{g}-\bar{g}_h)+D_0(R_{t+1},q_t,\bar{g}_h-g_0)$, where $\bar{g}_h(v;\gamma):=\mathbb{E}\left(\widehat{g}(v;\gamma)\right)$. The second component is a small bias term whose scaled and centered time average

$$
\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}D_0(R_{t+1},q_t,\bar{g}_h-g_0)-\int D_0(r,q,\bar{g}_h-g_0)\mathrm{d}F(r,q)\right)=o_p(1)
$$

by Chebyshev's inequality, as $\operatorname{var}\left(D_0(R_{t+1},q_t,\bar{g}_h-g_0)\right) = O\left(\|\bar{g}_h-g_0\|^2\right) = O(h^2)$ for second or higher order kernels. For the first component, define $K_{h,s+1}(u,\gamma):=K_h(U_{s+1}(\gamma)-u)$, $d_T(R_{t+1},q_t,R_{s+1},q_s):=D_0(R_{t+1},q_t,K_{h,s+1})$, and the related marginals $d_{T1}(R_{t+1},q_t):=\int d_T(R_{t+1},q_t,r,q)\mathrm{d}F(r,q)$ and $d_{T2,s}(R_{s+1},q_s):=\int d_T(r,q,R_{s+1},q_s)\mathrm{d}F(r,q)$. Define also the shorthand $d_{T,ts}:=d_T(R_{t+1},q_t,R_{s+1},q_s)$, $d_{T1,t}:=d_{T1}(R_{t+1},q_t)$, and $d_{T2,s}:=d_{T2}(R_{s+1},q_s)$ for $t,s=1,\ldots,T$. This allows expressing

$$
\begin{aligned}
&\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}D_0(R_{t+1},q_t,\widehat{g}-\bar{g}_h)-\int D_0(r,q,\widehat{g}-\bar{g}_h)\mathrm{d}F(r,q)\right) \\
&\quad=\sqrt{T}\left(\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}d_{T,ts}-\frac{1}{T}\sum_{t=1}^{T}d_{T1,t}-\frac{1}{T}\sum_{t=1}^{T}d_{T2,t}+\mathbb{E}\left(d_{T1,t}\right)\right) \\
&\quad=\sqrt{T}\left(\frac{1}{T^2}\sum_{t=1}^{T}\tilde{d}_{T,t}+\frac{T-1}{T}U_T\right)
\end{aligned}
$$

in terms of the second-order, centered U-statistic with symmetric kernel $u_T(R_{t+1},q_t,R_{s+1},q_s)$, or $u_{T,ts}$ for short:

$$U_T = \frac{1}{T(T-1)} \sum_{t=1}^{T} \sum_{s<t} \left(u_{T,ts} - u_{T1,t} - u_{T1,s} + \mathbb{E}\left(u_{T1,t}\right) \right),$$

where $u_{T,ts} = d_{T,ts} + d_{T,st}$, $u_{T1,t} = d_{T1,t} + d_{T2,t}$, and $\tilde{d}_{T,t} = d_{T,tt} - u_{T1,t} + \mathbb{E}\left(d_{T1,t}\right)$.

First, we analyze these terms for $R^*_{t,l} \leq R_{t+1} \leq R^*_{t,u}$, in which case

$$\begin{aligned}
d_T(R_{t+1}, q_t, R_{s+1}, q_s) &= \frac{1}{g_{\gamma_0}(U_{0,t+1})} \left(\dot{U}_{0,s+1} - \dot{U}_{0,t+1} \right) K_h'(U_{0,s+1} - U_{0,t+1}) \\
&\quad - \frac{\dot{g}_{\gamma_0}(U_{0,t+1}) + g'_{\gamma_0}(U_{0,t+1})\dot{U}_{0,t+1}}{g_{\gamma_0}(U_{0,t+1})^2} K_h(U_{0,s+1} - U_{0,t+1}) \\
&= \frac{1}{g_{\gamma_0}(U_{0,t+1})} \frac{\partial}{\partial \gamma} K_h\left(U_{t+1}(\gamma) - U_{s+1}(\gamma)\right)\Big|_{\gamma=\gamma_0} \\
&\quad - \frac{\frac{\partial}{\partial \gamma} g_\gamma(U_{t+1}(\gamma))|_{\gamma=\gamma_0}}{g_{\gamma_0}(U_{0,t+1})^2} K_h(U_{0,s+1} - U_{0,t+1}) \\
&= \frac{\partial}{\partial \gamma} \frac{K_h\left(U_{t+1}(\gamma) - U_{s+1}(\gamma)\right)}{g_\gamma(U_{t+1}(\gamma))}\Big|_{\gamma=\gamma_0}.
\end{aligned}$$

This allows computing the following for $1^m_{t+1} = 1$:

$$\begin{aligned}
d_T(R_{t+1}, q_t, R_{t+1}, q_t) &= -\frac{\left(\dot{g}_{\gamma_0}(U_{0,t+1}) + g'_{\gamma_0}(U_{0,t+1})\dot{U}_{0,t+1}\right)}{g_{\gamma_0}(U_{0,t+1})^2} h^{-1} K(0) = O_p\left(h^{-1}\right), \\
d_{T1}(R_{t+1}, q_t) &= \int \frac{\partial}{\partial \gamma} \frac{K_h\left(U_{t+1}(\gamma) - U(\gamma)\right)}{g_\gamma(U_{t+1}(\gamma))}\Big|_{\gamma=\gamma_0} \mathrm{d}F(r,q) \\
&= \frac{\partial}{\partial \gamma} \frac{\int K(z) g_\gamma(U_{t+1}(\gamma) + hz)\mathrm{d}z}{g_\gamma(U_{t+1}(\gamma))}\Big|_{\gamma=\gamma_0} = O_p\left(h\right), \\
d_{T2}(R_{t+1}, q_t) &= \int \frac{\partial}{\partial \gamma} \frac{K_h\left(U_{t+1}(\gamma) - U(\gamma)\right)}{g_\gamma(U(\gamma))}\Big|_{\gamma=\gamma_0} \mathrm{d}F(r,q) \\
&= \frac{\partial}{\partial \gamma} \int K_h\left(U_{t+1}(\gamma) - u\right)\mathrm{d}u\Big|_{\gamma=\gamma_0} = 0.
\end{aligned}$$

Therefore $\mathbb{E}\left(h\|1^m_{t+1}\tilde{d}_{T,t}\|\right) < \infty$, so by the ergodic theorem $\frac{1}{T\sqrt{T}}\sum_{t=1}^{T} 1^m_{t+1}\tilde{d}_{T,t} = O_p\left(\frac{1}{\sqrt{Th}}\right)$.

Focusing on the U-statistic, suppose that for some $\delta > 0$, $M < \infty$, positive series a_T, and all T

$$\iint \|u_T(R_{t+1},q_t,R_{s+1},q_s)\|^{2+\delta}\mathrm{d}F(R_{t+1},q_t)\mathrm{d}F(R_{s+1},q_s) \leq Ma_T, \tag{B.1}$$

and for all integers $t_1 < t_2$

$$\mathbb{E}\left(\|u_T(R_{t_2+1},q_{t_2},R_{t_1+1},q_{t_1})\|^{2+\delta}\right) \leq Ma_T. \tag{B.2}$$

By Assumptions A*(i)* and C*(iv)*, q_t is a measurable function of X_t so that measurable functions of (R_{t+1},q_t) are mixing at the rate in Assumption SE*(i)*. Yoshihara (1976, Lemma 2) then implies that (B.1) and (B.2) are sufficient for $\mathbb{E}\left(\|U_T\|^2\right) = O\left(T^{-1-\eta}a_T^{\frac{2}{2+\delta}}\right)$ with $\eta = 2\frac{\delta-\delta'}{\delta'(2+\delta)} > 0$. Writing $u_{T,ts} = \sum_{i,j\in\{l,m,u\}} 1_{t+1}^i 1_{s+1}^j u_{T,ts}$, we can establish (B.1) and (B.2) for the middle region ($i = j = m$), the tails ($i = j \neq m$), and the cross-terms ($i \neq j$) separately.

To establish (B.1) on the middle region, use $|\frac{1}{g_{\gamma_0}(U_{0,t+1})} - \frac{1}{g_{\gamma_0}(U_{0,s+1})}| \leq C\,|U_{0,t+1} - U_{0,s+1}|$ to write

$$\begin{aligned}
&1_{t+1}^m 1_{s+1}^m \|u_T(R_{t+1},q_t,R_{s+1},q_s)\| \\
&\leq C\left\|(U_{0,t+1}-U_{0,s+1})\left(\dot{U}_{0,s+1}-\dot{U}_{0,t+1}\right)K_h'\left(U_{0,s+1}-U_{0,t+1}\right)\right\| \\
&\quad + \left(C_0 + C_1\left(\|\dot{U}_{0,s+1}\| + \|\dot{U}_{0,t+1}\|\right)\right)\left|K_h\left(U_{0,s+1}-U_{0,t+1}\right)\right|,
\end{aligned} \tag{B.3}$$

where C, C_0 and C_1 are some large, positive constants. For the first term on the RHS of (B.3), note that

$$\begin{aligned}
&\iint 1_{t+1}^m 1_{s+1}^m \left\|(U_{0,t+1}-U_{0,s+1})\left(\dot{U}_{0,s+1}-\dot{U}_{0,t+1}\right)K_h'\left(U_{0,s+1}-U_{0,t+1}\right)\right\|^{2+\delta} \mathrm{d}F(R_{t+1},q_t)\mathrm{d}F(R_{s+1},q_s) \\
&\leq 2C\int 1_{s+1}^m \int_{v^*}^{1-v^*} \left\|(U_{0,t+1}-U_{0,s+1})\dot{U}_{0,s+1}K_h'\left(U_{0,s+1}-U_{0,t+1}\right)\right\|^{2+\delta} \mathrm{d}G_{\gamma_0}(U_{0,t+1})\mathrm{d}F(R_{s+1},q_s) \\
&\leq \frac{C}{h^{1+\delta}}\int |zK'(z)|^{2+\delta}\,\mathrm{d}z \max_{v^*\leq v\leq 1-v^*} \mathbb{E}\left(\|\dot{U}_0\|^{2+\delta} \mid U_0 = v\right)g_{\gamma_0}(v) = O\left(\frac{1}{h^{1+\delta}}\right),
\end{aligned}$$

using Minkowski's inequality and stationarity in the first step, and continuity of the functions inside the maximum in the second. To establish (B.2) in the middle region, for the first term on the RHS of (B.3) we derive for any $s > t$ that

$$\begin{aligned}
&\mathbb{E}\left(1_t^m 1_s^m \left\|(U_{0,t}-U_{0,s})\,(\dot{U}_{0,s}-\dot{U}_{0,t})K_h'\,(U_{0,s}-U_{0,t})\right\|^{2+\delta}\right)\\
&\leq \int_{v^*}^{1-v^*}\int_{v^*}^{1-v^*}\left|(U_{0,t}-U_{0,s})\,K_h'\,(U_{0,s}-U_{0,t})\right|^{2+\delta}\mathbb{E}\left(\|\dot{U}_{0,s}-\dot{U}_{0,t}\|^{2+\delta}\mid U_{0,t},U_{0,s}\right)\mathrm{d}G_{\gamma_0}(U_{0,t})\mathrm{d}G_{\gamma_0}(U_{0,s})\\
&\leq \frac{C}{h^{1+\delta}}\int\left|zK'\,(z)\right|^{2+\delta}\mathrm{d}z\max_{v^*\leq v\leq 1-v^*}\mathbb{E}\left(\|\dot{U}_{0,s}-\dot{U}_{0,t}\|^{2+\delta}\mid U_{0,s}=U_{0,t}=v\right)+O\left(h^{-\delta}\right),
\end{aligned}$$

which is $O\left(\frac{1}{h^{1+\delta}}\right)$ since the conditional moment inside the maximum is bounded on $[v^*, 1-v^*]$ by SE*(ii)*. The integrals in (B.1) and (B.2) for the second term on the RHS of (B.3) are also $O\left(\frac{1}{h^{1+\delta}}\right)$ by the moment condition on $\dot{U}_0$.

Meanwhile, when $1_{t+1}^l = 1$, with $\dot{U}_{0,t}^i := \dot{U}_t^i(\gamma_0)$ where $\dot{U}_t^i(\gamma) := \dot{U}_t(R_{t,i}^*;\gamma)$ for $i \in \{l, u\}$,

$$\begin{aligned}
d_T(R_{t+1},q_t,R_{s+1},q_s) &= \frac{1}{G_{0,\gamma_0}(U_{0,t}^l)}\left(\dot{U}_{0,t}^l-\dot{U}_{0,s+1}\right)K_h\left(U_{0,t}^l-U_{0,s+1}\right)\\
&\quad-\frac{\dot{G}_{0,\gamma_0}(U_{0,t}^l)+g_{0,\gamma_0}(U_{0,t}^l)\dot{U}_{0,t}^l}{G_{0,\gamma_0}(U_{0,t}^l)^2}F_K\left(\frac{U_{0,t}^l-U_{0,s+1}}{h}\right)\\
&=\frac{\partial}{\partial\gamma}\frac{F_K\left(\frac{1}{h}\left(U_t^l(\gamma)-U_{s+1}(\gamma)\right)\right)}{G_{0,\gamma}(U_t^l(\gamma))}\Bigg|_{\gamma=\gamma_0}.
\end{aligned}$$

Similar steps as for the middle region allow computing that $\mathbb{E}\left\|1_{t+1}^l d_{T,tt}\right\| = O(1)$, $\mathbb{E}\left\|1_{t+1}^l d_{T1,t}\right\| = O(h)$, and $\mathbb{E}\left\|1_{t+1}^l d_{T2,t}\right\| = O(1)$. The first of these is now $O(1)$ rather than $O(h^{-1})$ due to integrating the kernel function, while the second is now $O(h)$ rather than $O(h^2)$ as F_K is asymmetric. The term $d_{T2}(R_{t+1}, q_t)$ is bounded since $U_0^l(q) \in [v^*, 1-v^*]$ for all q. Combining implies $\mathbb{E}\left(\|1_{t+1}^l \tilde{d}_{T,t}\|\right) < \infty$, so by the ergodic theorem $\frac{1}{T\sqrt{T}}\sum_{t=1}^{T} 1_{t+1}^{l}\tilde{d}_T(R_{t+1},q_t) = O_p\left(\frac{1}{\sqrt{T}}\right)$. Furthermore, $U_{0,t}^{l} \in [v^*, 1-v^*]$ a.s. implies

$$\begin{aligned} 1_{t+1}^{l}\left\|d_T(R_{t+1},q_t,R_{s+1},q_s)\right\| \leq C\left\|\dot{U}_{0,t}^{l}-\dot{U}_{0,s+1}\right\|\left|K_h\left(U_{0,t}^{l}-U_{0,s+1}\right)\right| \\ +\left(C_0+C_1\|\dot{U}_{0,t}^{l}\|\right)\left|F_K\left(\frac{U_{0,t}^{l}-U_{0,s+1}}{h}\right)\right|. \end{aligned} \tag{B.4}$$

The first term on the RHS of (B.4) is similar to the second term in (B.3). Using that

$$\mathbb{E}(\|\dot{U}_{0,t}^{l}\|^{2+\delta}) = \mathbb{E}\left(\mathbb{E}(\|\dot{U}_{0,t+1}\|^{2+\delta} \mid U_0 = U_{0,t}^{l})\right) < \infty,$$

its integrals in (B.1) and (B.2) are $O\left(\frac{1}{h^{1+\delta}}\right)$. The integrals over the second term on the RHS of (B.4) are $O(1)$.

For cross-terms with one observation in the middle and the other in the lower tail, write

$$1_{t+1}^{m}1_{s+1}^{l}u_{T,ts}+1_{t+1}^{l}1_{s+1}^{m}u_{u,ts} = (1_{t+1}^{m}1_{s+1}^{l}d_{ts,T}+1_{t+1}^{l}1_{s+1}^{m}d_{st,T})+(1_{t+1}^{m}1_{s+1}^{l}d_{st,T}+1_{t+1}^{l}1_{s+1}^{m}d_{ts,T}).$$

The first bracketed term is ordered to satisfy the middle region bound (B.3), while the terms in the second bracket satisfy the tail bound (B.4). Therefore, their integrals in (B.1) and (B.2) are also $O\left(\frac{1}{h^{1+\delta}}\right)$. By symmetry, the same order is obtained when one or both observations are in the upper tail, i.e. when $1_{t+1}^{u} = 1$.

We conclude that (B.1) and (B.2) hold with $a_T = \frac{1}{h^{1+\delta}}$. As a result, we find $\sqrt{T}U_T = O_p\left(\left(T^{\frac{\delta-\delta'}{\delta'}}h^{1+\delta}\right)^{\frac{-1}{2+\delta}}\right)$, which is $o_p(1)$ under the bandwidth conditions in SE*(iv)*. □

The remainder of this subsection verifies the uniform convergence conditions imposed in Assumptions C, A, and N for the kernel estimators of g and G given by (3.4) and (3.11), respectively. Throughout, we make the following assumptions on the dynamic properties of (R,q) and the kernel K, where $\mu_k(K) = \int z^k K(z)\mathrm{d}z$:

Assumption K.

(i) $f_t(r) = f(r|X_t)$ *is a positive, continuous function of the stationary process* (R_{t+1}, X_t) *on* $\mathbb{R}_{++} \times \mathcal{X}$*;*

(ii) (R_{t+1}, X_t) *is strongly mixing with mixing coefficients* $\alpha(j) \leq Aj^{-\beta}$ *for some finite, positive* A *and* β*;*

(iii) $Z'(P)$, $u'(r;\gamma)$, *and* $c_t(\gamma)$ *are positive and continuous on* $(0,1)$, $\mathbb{R}_{++} \times \Theta$, *and* Θ *a.s., respectively;*

(iv) The kernel K *is Lipschitz continuous, has compact support, and* $\mu_0(K) = 1$.

Consider the family of time series processes $(U_{t+1}(\gamma))_{t=1}^T$ indexed by $\gamma \in \Theta \subset \mathbb{R}^k$. Let $g(u;\gamma)$ and $g_j(u,v;\gamma)$ denote the densities of $U_t(\gamma)$ and $(U_t(\gamma), U_{t+j}(\gamma))$, which do not depend on (t,T) due to the stationarity assumption. Assumptions K*(i)*-K*(iii)* imply that $g(u;\gamma)$ is positive and continuous by Lemma D. Furthermore, they imply that $q_t(r)$ is a measurable function of the mixing process X_t, and is therefore itself strongly mixing with polynomially decaying coefficients. Moreover, this property carries over to measurable functions of q_t and R_{t+1}, such as $U_{t+1}(\gamma)$.

First, we establish sufficient conditions for $\sup_{\gamma\in\Theta} \|\widehat{g}_\gamma - g_\gamma\|_{\infty,v^*} \xrightarrow{p} 0$ by adapting a result from Kristensen (2009).

Lemma K-1. *Suppose Assumption K and the following conditions hold:*

(i) $Z'(P)$, $u'(r;\gamma)$, *and* $c_t(\gamma)$ *are continuously differentiable on* $(0,1)$, $\mathbb{R}_{++} \times \Theta$, *and* Θ *a.s., respectively;*

(ii) $\mathbb{E}\left(\|\dot{U}_t(\gamma)\|^{2+\delta}|U_t(\gamma) = u\right) g(u;\gamma)$ *is continuous on* $(0,1) \times \Theta$ *for some* $\delta > 0$ *such that* $\beta > \frac{1+(1+\delta)(2+k)}{\delta}$*;*

(iii) For some $M \geq 0$, $\sup_{|t-s|\geq M} \mathbb{E}\left(|\dot{U}_t(\gamma)\dot{U}_s(\gamma)^\top|\middle|U_t(\gamma) = u, U_s(\gamma) = v\right) g_{t-s}(u,v;\gamma)$ *is continuous on* $(0,1)^2 \times \Theta$*;*

(iv) The bandwidth satisfies $h \to 0$ and $\log T/(T^{\theta}h) \to 0$ with

$$\theta = \frac{\beta - 2 - k - (1+\beta)/(1+\delta)}{\beta + 2 - k - (1+\beta)/(1+\delta)}.$$

Then $\sup_{\gamma \in \Theta} \|\widehat{g}_{\gamma} - g_{\gamma}\|_{\infty, v^*} \xrightarrow{p} 0$.

Proof. Assumptions *(i)*-*(iii)* are sufficient for A.1-A.4 and A.6.1 in Kristensen (2009, Theorem 1(i)) with $X_{t,T}(\gamma) = U_t(\gamma)$, $c_T = 1$, and $d_T = d$ for some large $d > 0$. In particular, *(i)* implies that g' is continuous and the derivative $\dot{U}_t(\gamma)$ exists. Condition *(ii)* implies that $\mathbb{E}(1^m_{t+1}\|\dot{U}_{t+1}(\gamma)\|^{2+\delta}) < \infty$ by the law of iterated expectations. Continuity of the function in *(iii)* guarantees its boundedness on $[v^*, 1-v^*]^2$. The theorem yields that $\sup_{\gamma \in \Theta} \|\widehat{g}_{\gamma} - \mathbb{E}\widehat{g}_{\gamma}\|_{\infty, v^*} = O_p(\sqrt{\log T/(Th)}) = o_p(1)$ under bandwidth condition *(iv)*. Furthermore, $\sup_{\gamma \in \Theta} \|\mathbb{E}\widehat{g}_{\gamma} - g_{\gamma}\|_{\infty, v^*} = O(h)$ by a Taylor expansion and the continuity of g'. The conclusion follows from the triangle inequality. □

Next we verify that $\sup_{\gamma \in \Theta} |\widehat{G}_{\gamma}(v^*) - G_{\gamma}(v^*)| \xrightarrow{p} 0$ for the kernel CDF estimator (3.11). Its uniform convergence would follow directly from Lemma K-1, provided Assumptions *(i)* and *(iii)* hold on the closed interval $[0, 1]$. However, this rules out asymptotes in g_{γ}, g'_{γ} and $\dot{G}_{\gamma}$ when v goes to 0 or 1 for any γ. Instead, the following Lemma requires some additional continuity and moment conditions:

Lemma K-1*. *The following conditions, in addition to those in Lemma K-1, imply that* $\sup_{\gamma \in \Theta} |\widehat{G}_{\gamma}(v^*) - G_{\gamma}(v^*)| \xrightarrow{p} 0$ *for the kernel CDF estimator:*

(i) $u'(r;\gamma)$ and $c_t(\gamma)$ are twice continuously differentiable in γ on $\mathbb{R}_{++} \times \Theta$ and Θ a.s., respectively;

(ii) For some $\delta > 0$ and all $\gamma \in \Theta$, $\mathbb{E}(\|\ddot{U}_t(\gamma)\|^{2+\delta}) < \infty$ and $\mathbb{E}(\|\dot{U}_t(\gamma)\|^{4+2\delta}) < \infty$;

(iii) The following quantities are finite and continuous in γ on Θ:

$$\tilde{B}_1^*(\gamma) = \mathbb{E}\left(\|W_t(\gamma)\| \mid U_t(\gamma) = v^*\right)$$
$$\tilde{B}_2^*(\gamma) = \sup_{|t-s|\geq M} \mathbb{E}\left(\|W_t(\gamma)W_s(\gamma)^\top\| \Big| U_t(\gamma) = v^*, U_s(\gamma) = v^*\right) g_{t-s}(v^*, v^*; \gamma),$$

for $W_t(\gamma) \in \{\frac{\partial^{1+|j|}}{\partial\gamma^{1+j}}U_t(\gamma), \dot{U}_t(\gamma)\frac{\partial^{|j|}}{\partial\gamma^j}U_t(\gamma)\}$ for multi-indices $|j| \leq 1$ and some $M > 0$.

Proof. The derivative of $\widehat{G}_\gamma(v^*)$ w.r.t. γ is the kernel-weighted average

$$\hat{\dot{G}}_\gamma(v^*) = -\frac{1}{T}\sum_{t=1}^{T} \dot{U}_{t+1}(\gamma)K_h(v^* - U_{t+1}(\gamma)).$$

Conditions *(i)*-*(iii)* imply those required in Kristensen (2009, Thm. 1(i)) for the uniform convergence $\sup_{\gamma\in\Theta} \|\hat{\dot{G}}_\gamma(v^*) - \mathbb{E}\hat{\dot{G}}_\gamma(v^*)\| \xrightarrow{p} 0$. In particular, *(i)* ensures that $U_t(\gamma)$ is a.s. twice differentiable. Meanwhile,

$$\sup_{\gamma\in\Theta} \|\mathbb{E}\hat{\dot{G}}_\gamma(v^*)\| = \sup_{\gamma\in\Theta} \left\|\mathbb{E}\left(\dot{U}_{t+1}(\gamma)K_h(v^* - U_{t+1}(\gamma))\right)\right\| = \sup_{\gamma\in\Theta} \|\dot{G}_\gamma(v^*)\| + O(h),$$

where the last equation uses that $\dot{g}(v;\gamma)$ is continuous in a neighborhood of v^* for all $\gamma \in \Theta$ under K-1*(i)*. The mean-value theorem thus establishes stochastic equicontinuity of $\widehat{G}(v^*;\gamma)$, whose pointwise convergence follows from the boundedness of F_K and the ergodic theorem. Therefore $\sup_{\gamma\in\Theta} |\widehat{G}_\gamma(v^*) - \mathbb{E}\widehat{G}_\gamma(v^*)| \xrightarrow{p} 0$ by a uniform law of large numbers such as (Andrews, 1992, Thm. 3). A Taylor expansion implies that $\sup_{\gamma\in\Theta} |\mathbb{E}\widehat{G}_\gamma(v^*) - G_\gamma(v^*)| = O(h)$. The conclusion follows from the triangle inequality. □

Next, we provide primitive conditions for the kernel estimator to satisfy the uniform convergence rate assumptions A*(ii)*. The rate conditions on the density estimator and its first derivatives follow from existing results for kernel estimators, such as Andrews (1995). In particular, the following lemma provides sufficient conditions.

Lemma K-2. *Suppose Assumption K and the following conditions hold, for some $\omega \geq 4$:*

(i) *$Z'(P)$ is ω times continuously differentiable on $(0,1)$, $u'(r;\gamma_0)$ and $\dot{u}'(r;\gamma_0)$ are ω and $\omega-1$ times continuously differentiable on $\mathbb{R}_{++}$, $f_t(r)$ is $\omega-1$ times continuously differentiable a.s. on $\mathbb{R}_{++}$, and $c_t(\gamma)$ is continuously differentiable a.s. at γ_0;*

(ii) *$\mathbb{E}\|\dot{U}_0\|^{2+\delta} < \infty$ for some $\delta > 0$ such that $\beta > \frac{2+\delta}{\delta}$;*

(iii) *$\mu_k(K) = 0$ for $k = 1, \dots, \omega-2$;*

(iv) *$h = O(T^{-\psi})$ and $h^{-1} = O(T^{\psi})$ for some $\psi > 0$ that satisfies $\frac{1}{4(\omega-1)} < \psi < \frac{1}{8}$.*

Then $\sqrt{T}\|\widehat{g} - g_0\|^2 \xrightarrow{p} 0$.

Proof. Condition *(i)* implies that g_{γ_0} and $\dot{g}_{\gamma_0}$ are continuously differentiable on $(0,1)$ up to orders $\omega \geq 4$ and $\omega - 1$, respectively, by Lemma D. Condition *(ii)* establishes that $(\dot{U}_0, U_0)$ are strongly mixing with coefficients such that $\sum_{l=1}^{\infty} \alpha(l)^{\frac{\delta}{2+\delta}} < \infty$. The result follows from Andrews (1995, Lemma A-1) by verifying its assumptions NP1-NP5 with $(Y_t, X_t) = (1, U_0)$ with $\lambda = 0, 1$, and $(Y_t, X_t) = (\dot{U}_0, U_0)$ with $\lambda = 1$, to establish the uniform convergence of $\widehat{g}_{\gamma_0}$, $\widehat{g}'_{\gamma_0}$, and $\widehat{\dot{g}}_{\gamma_0}$, respectively, on $[v^*, 1-v^*]$ for any $v^* > 0$. □

Lemma K-2*. *If the conditions in Lemma K-2 hold, and $\mathbb{E}\left(\|\dot{U}_0\|^2 \mid U_0 = v\right)$ exists and is continuously differentiable at v^*, then $\widehat{G}_{\gamma_0}(v^*) - G_{\gamma_0}(v^*) = o_p\left(T^{-1/4}\right)$ and $\widehat{\dot{G}}_{\gamma_0}(v^*) - \dot{G}_{\gamma_0}(v^*) = o_p\left(T^{-1/4}\right)$ for the kernel CDF estimator.*

Proof. For the result on $\widehat{G}_{\gamma_0}(v^*)$, suppose that $h \leq v^*$, which happens for T large enough. Using partial integration, the expectation of the kernel CDF estimator can then be written as

$$
\begin{aligned}
\mathbb{E}F_K\left(\frac{v^*-U_{0,t+1}}{h}\right) &= \int_0^1 F_K\left(\frac{v^*-v}{h}\right) g_{\gamma_0}(v)\mathrm{d}v \\
&= F_K\left(z\right) G_{\gamma_0}(v^*-hz)\Big|_{(v^*-1)/h}^{v^*/h} + \int_{-1}^1 K\left(z\right) G_{\gamma_0}(v^*-hz)\mathrm{d}z \\
&= G_{\gamma_0}(v^*) + O(h^{\omega-1}),
\end{aligned}
$$

using the higher order kernel Assumption K-2*(iii)*, and the $\omega+1$ times continuous differentiability of G_{γ_0} at v^*. Meanwhile, since the summands $F_K\left(\frac{v^*-U_{0,t+1}}{h}\right)$ inherit the i.i.d. property of $U_{0,t+1}$, the variance term equals $\mathrm{var}\left(\widehat{G}_{\gamma_0}(v^*)\right) = \frac{1}{T}\mathrm{var}\left(F_K\left(\frac{v^*-U_{0,t+1}}{h}\right)\right) \leq \frac{C}{T}$ since F_K is bounded. We conclude that $\widehat{G}_{\gamma_0}(v^*) - G_{\gamma_0}(v^*) = O(h^{\omega-1}) + O_p\left(T^{-1/2}\right) = o_p\left(T^{-1/4}\right)$ under the bandwidth condition K-2*(iv)*.

For the second result, differentiating G_γ w.r.t. γ yields the relation

$$
\dot{G}_{\gamma_0}(v) = -\mathbb{E}(\dot{U}_0 \mid U_0 = v) g_{\gamma_0}(v).
$$

With $\widehat{\dot{G}}_{\gamma_0}(v^*) = -\frac{1}{T}\sum_{t=1}^T \dot{U}_{0,t+1} K_h\left(v^* - U_{0,t+1}\right)$, we can decompose $\widehat{\dot{G}}_{\gamma_0}(v^*) - \dot{G}_{\gamma_0}(v^*) = B_T + V_T + D_T$ into bias, variance, and density estimation terms, where the bias term

$$
B_T = -\frac{1}{T}\sum_{t=1}^T \left(\mathbb{E}(\dot{U}_{0,t+1} \mid U_{0,t+1}) - \mathbb{E}(\dot{U}_{0,t+1} \mid U_{0,t+1} = v^*)\right) K_h\left(v^* - U_{0,t+1}\right)
$$

has i.i.d. summands with $\mathbb{E}B_T = O(h^{\omega-1})$ and $\mathrm{var}B_T = O\left(\frac{1}{Th}\right)$, the variance term equals

$$
V_T = -\frac{1}{T}\sum_{t=1}^T \left(\dot{U}_{0,t+1} - \mathbb{E}(\dot{U}_{0,t+1} \mid U_{0,t+1})\right) K_h\left(v^* - U_{0,t+1}\right),
$$

and the density estimation term, which is $o_p\left(T^{-1/4}\right)$ by Lemma K-2, equals

$$
D_T = -\mathbb{E}(\dot{U}_{0,t+1} \mid U_{0,t+1} = v^*)\left(\hat{g}_{\gamma_0}(v^*) - g_{\gamma_0}(v^*)\right).
$$

The variance term has mean $\mathbb{E}V_T = 0$ and long-run covariance matrix $\mathbb{E}V_T V_T^\top = V_{1T} + V_{2T}$, where the contemporaneous covariance is

$$
\begin{aligned}
V_{1T} &:= \frac{1}{T}\mathbb{E}\left(\dot{U}^c_{0,t+1}\dot{U}^{c\,\top}_{0,t+1}K_h^2\left(v^* - U_{0,t+1}\right)\right) \\
&= \frac{g_{\gamma_0}(v^*)}{Th}\mathbb{E}(\dot{U}_{0,t+1}\dot{U}^\top_{0,t+1} \mid U_{0,t+1} = v^*)\int K^2(z)dz + o\left(\frac{1}{Th}\right),
\end{aligned}
$$

with $\dot{U}^c_{0,t+1} := \dot{U}_{0,t+1} - \mathbb{E}(\dot{U}_{0,t+1} \mid U_{0,t+1})$, and the between-period covariances sum to

$$
\begin{aligned}
V_{2T} &:= \frac{1}{T^2}\sum_{l=1}^{T}\left(1 - \frac{l}{T}\right)\mathbb{E}\left(\dot{U}^c_{0,1}\dot{U}^{c\,\top}_{0,1+l}K_h\left(v^* - U_{0,1}\right)K_h\left(v^* - U_{0,1+l}\right)\right) \\
&\leq C\frac{1}{T^2h^2}\left(\mathbb{E}||\dot{U}_0\|^{2+\delta}\right)^{\frac{2}{2+\delta}}\sum_{l=1}^{T}\left(1 - \frac{l}{T}\right)\alpha(l)^{\frac{\delta}{2+\delta}},
\end{aligned}
$$

which is $O\left(\frac{1}{T^2h^2}\right)$ by K-2*(ii)*. We conclude that $\widehat{\dot{G}}_{\gamma_0}(v^*) - \dot{G}_{\gamma_0}(v^*) = O_p\left(h^{\omega-1} + \frac{1}{\sqrt{Th}}\right) + o_p\left(T^{-1/4}\right) = o_p\left(T^{-1/4}\right)$ by K-2*(iv)*. □

Lemma K-3. *Suppose Assumption K and the following conditions hold for some neighborhood $\mathcal{N}_0$ of γ_0:*

(i) For some $\delta > 0$, all $\gamma \in \mathcal{N}_0$, and multi-indices j,

$$
\begin{aligned}
&\mathbb{E}\left(\left|\frac{\partial^{|j|}}{\partial\gamma^j}U_t(\gamma)\right|^{2+\delta}\right) < \infty \text{ for } |j| \leq 3, \\
&\mathbb{E}\left(\left|\frac{\partial^{|j|}}{\partial\gamma^j}U_t(\gamma)\right|^{2+\delta}\|\dot{U}_t(\gamma)\|^{2+\delta}\right) < \infty \text{ for } |j| \leq 2;
\end{aligned}
$$

(ii) The mixing exponent β satisfies $\beta > \frac{1+(1+\delta)(2+k)}{\delta}$;

(iii) *The following quantities are continuous on* $(0,1)\times\mathcal{N}_0$*:*

$$\tilde{B}_1(\gamma,v)=\mathbb{E}\left(\|W_t(\gamma)\|\mid U_t(\gamma)=v\right),$$
$$\tilde{B}_2(\gamma,u,v)=\sup_{|t-s|\geq M}\mathbb{E}\left(\|W_t(\gamma)W_s(\gamma)^\top\|\Big|U_t(\gamma)=u,U_s(\gamma)=v\right)g_{t-s}(u,v;\gamma),$$

for $W_t(\gamma)\in\{\frac{\partial^{1+|j|}}{\partial\gamma^{1+j}}U_t(\gamma),\dot{U}_t(\gamma)\frac{\partial^{|j|}}{\partial\gamma^j}U_t(\gamma)\}$ *for multi-indices* $|j|\leq 2$ *and some* $M>0$*;*

(iv) K *is three times differentiable;*

(v) $\frac{\log T}{Th^5}\to 0$ *and* $\frac{\log T}{T^\theta h}\to 0$ *with* $\theta=\frac{\beta-2-k-(1+\beta)/(1+\delta)}{\beta+2-k-(1+\beta)/(1+\delta)}$*;*

(vi) $Z'(P)$*,* $u'(r;\gamma)$*, and* $c_t(\gamma)$ *are three times, and* $f_t(r)$ *two times, continuously differentiable on* $(0,1)$*,* $\mathbb{R}_{++}\times\mathcal{N}_0$*,* $\mathcal{N}_0$ *a.s., and* $\mathbb{R}_{++}$ *a.s., respectively.*

Then $\sup_{\gamma\in\mathcal{N}_0}\|\frac{\partial^{i+|j|}}{\partial v^i\partial\gamma^j}(\widehat{g}_\gamma-g_\gamma)\|_{\infty,v^*}\xrightarrow{p}0$ *for any non-negative integer* i *and multi-index* j *with* $i+|j|\leq 2$*.*

Proof. Condition *(vi)* implies that $g(v;\gamma)$ is three times continuously differentiable on $(0,1)\times\mathcal{N}_0$ by Lemma D, and that $U_t(\gamma)$ is a.s. three times differentiable on $\mathcal{N}_0$. The result follows from Kristensen (2009, Thm. 1(i)) by verifying its assumptions A.1-A.4 and A.6.1 with $(Y_t(\gamma),X_t(\gamma))=(\frac{\partial^{|j|}}{\partial\gamma^j}U_t(\gamma),U_t(\gamma))$ and $K_i(z)=\frac{\partial^i}{\partial^i z}K(z)$ for each (i,j) with $i+|j|\leq 2$, while setting $c_T=1$ and $d_T=\gamma_0+\epsilon$ for some $\epsilon>0$. It establishes that

$$\sup_{\gamma\in\mathcal{N}_0}\left\|\frac{\partial^{i+|j|}}{\partial v^i\partial\gamma^j}(\widehat{g}_\gamma-\mathbb{E}\widehat{g}_\gamma)\right\|_{\infty,v^*}=O_p\left(\sqrt{\frac{\log T}{Th^{1+2(i+|j|)}}}\right),$$

which is $o_p(1)$ by Assumption *(v)*. The conclusion then follows from $\sup_{\gamma\in\mathcal{N}_0}\|\frac{\partial^{i+|j|}}{\partial v^i\partial\gamma^j}(\mathbb{E}\widehat{g}_\gamma-g_\gamma)\|_{\infty,v^*}=O(h^{\omega-1-i})$ by a Taylor expansion of order $\omega-1$, and the triangle inequality. □

Lemma K-3*. *If the conditions in K-3 hold, then* $\sup_{\gamma\in\mathcal{N}_0}\left|\frac{\partial^{|j|}}{\partial\gamma^j}(\widehat{G}_\gamma(v^*)-G_\gamma(v^*))\right|\xrightarrow{p}0$ *for* $|j|\leq 2$*.*

Proof. For $|j| = 0, 1$ the result follows analogous to the proof of Lemma K-1*, whose conditions are satisfied by those in K-3 with $\mathcal{N}_0$ instead of Θ. For $|j| = 2$, write $\widehat{\ddot{G}}_\gamma(v^*) = \frac{1}{T}\sum_{t=1}^{T} \dot{U}_{t+1}(\gamma)\dot{U}_{t+1}^\top(\gamma)K_h'\left(v^* - U_{t+1}(\gamma)\right) - \frac{1}{T}\sum_{t=1}^{T} \ddot{U}_{t+1}(\gamma)K_h\left(v^* - U_{t+1}(\gamma)\right)$. Both components are kernel averages whose uniform convergence follows from Kristensen (2009, Thm. 1(i)) with $(Y_t(\gamma), X_t(\gamma)) = (\dot{U}_t(\gamma)\dot{U}_t^\top(\gamma), U_t(\gamma))$ and $(Y_t(\gamma), X_t(\gamma)) = (\ddot{U}_t(\gamma), U_t(\gamma))$ with kernels K' and K, respectively. The conclusion then follows from $\sup_{\gamma\in\mathcal{N}_0} |\frac{\partial^{|j|}}{\partial\gamma^j}(\mathbb{E}\widehat{G}_\gamma(v^*) - G_\gamma(v^*))| = O(h)$ by K-3*(vi)* and the triangle inequality. □

C Verification of conditions for common models

The assumptions made for the asymptotic theory in Section 3 impose certain regularity and smoothness conditions on the probability weighting function, such as positivity of the weighting density by Assumption I*(i)*, and continuity and differentiability by Assumptions C*(i)*, A*(i)* and N*(ii)*. Due to our trimming procedure, these conditions only need to hold on the open interval $(0, 1)$, which is easily verified for popular parametrizations of the probability weighting function.

For the probability weighting function (PWF) of Tversky and Kahneman (1992),

$$\overline{Z}'_{TK}(P) = \left[\frac{\delta - 1}{P^{1-\delta}} + \delta\frac{(1-P)^\delta}{P} + \frac{1}{(1-P)^{1-\delta}}\right]\frac{\overline{Z}_{TK}(P)}{P^\delta + (1-P)^\delta},$$

which diverges to infinity at the boundary points for $\delta < 1$ and to zero when $\delta > 1$. Meanwhile, for the PWF of Prelec (1998),

$$\overline{Z}'_P(P) = \frac{\alpha\beta}{P}(-\beta\log(P))^{\alpha-1}\overline{Z}_P(P).$$

When $\alpha < 1$, $\overline{Z}'_P(P)$ is $O\left(\frac{1}{P}\right)$ when $P \to 0$ and $O\left((\log(P))^{\alpha-1}\right)$ when $P \to 1$, and thus diverges to infinity at both boundary points. On the other hand, when $\alpha > 1$, $\overline{Z}'_P(P)$

vanishes to zero at the boundary points. The same is true for the associated cumulative probability weighting functions. As we trim at the boundary, our conditions allow for such divergence to infinity and/or zero at the boundary points $P = 0$ and $P = 1$. In particular, A*(i)* merely requires the probability weighting density $Z'(P)$ to be positive and continuously differentiable on the open interval $(0, 1)$, which both models satisfy. Moreover, our conditions allow for linear probability weighting $Z_0(P) = P$, which corresponds to the standard expected utility model.

Since $g_{\gamma_0}(v) = \frac{1}{Z'(Z^{-1}(v))}$ by the inverse function theorem, the density of $U_{t+1}(\gamma_0)$ goes to zero when v goes to zero and one when either PWF specification takes the inverse S-shape. Our Assumptions C*(iii)*, A*(ii)*, and N*(iii)* on the uniform convergence of $\hat{g}$ allow for this as $g_\gamma(v)$ remains bounded away from zero on the trimmed domain $[v^*, 1 - v^*]$, uniformly over γ, due to the positivity and continuity of $g(v; \gamma)$ on $(0, 1) \times \Theta$ established in Lemma D in the Appendix.

Similarly, our trimming procedure allows for asymptotes in the derivatives of $g_\gamma(v)$ near boundary points, which arise under common models. In particular, differentiating expression (A.2) for $G_\gamma(v)$ w.r.t. γ yields

$$\dot{G}_{\gamma_0}(v) = -\mathbb{E}\left(v \frac{c_t'(\gamma_0)}{c_t(\gamma_0)} - \int_0^v \frac{\dot{u}'(U_t^{-1}(u; \gamma_0); \gamma_0)}{u'(U_t^{-1}(u; \gamma_0); \gamma_0)} \mathrm{d}u \right) g_{\gamma_0}(v).$$

Under power utility, $\frac{\dot{u}'(r;\gamma)}{u'(r;\gamma)} = -\log r$. Since $\mathbb{E} \log U_t^{-1}(v; \gamma_0)$ diverges when $v \to 0$ or $v \to 1$ for any strictly increasing PWF, $\dot{g}_{\gamma_0}(u) = \frac{\partial}{\partial v} \dot{G}_{\gamma_0}(v)$ diverges at the boundaries. Still, $\dot{g}_\gamma$ remains positive and continuous on the trimmed domain, which enables stochastic equicontinuity for our kernel estimator by Lemma SE and the uniform convergence of its derivatives in Appendix B.

D Additional Simulation Results

D.1 Feasible inference performance

This appendix contains results regarding the performance of feasible inference techniques in the simulations.

D.1.1 Bootstrap confidence intervals

To compute confidence intervals for the PL estimation error $\hat{\gamma}-\gamma_0$, we consider the bootstrap described in Section 3.7. We consider confidence intervals based on the bootstrap quantiles as well as one exploiting the asymptotic normal distribution using the bootstrap standard deviation. Table 1 reports the associated coverage rates based on $K = 500$ bootstrap repetitions for each of the $N = 1000$ simulations. Both types of the bootstrapped intervals achieve coverage rates that are close to the nominal rates of 90% and 95%, with some slight under-coverage that can be explained by the parameter estimation biases. The similar performance of the quantile- and standard deviation-based intervals confirms the approximate normality of the PLE.

Table 1: Bootstrap results

		Z_{TK}		Z_P		Z_0	
	CI (%)	90	95	90	95	90	95
Quantile-based	Coverage (%)	88.3	92.7	88.8	93.7	85.7	92.4
	Width	3.23	3.84	4.39	5.22	5.59	6.65
St.d.-based	Coverage (%)	89.8	93.9	89.6	94.8	89.0	94.9
	Width	3.25	3.87	4.41	5.26	5.62	6.70

Note: This table displays the result of the nonparametric bootstrap, see Section 3.7. The columns are based on different DGPs, in the form of a different true probability weighting function. The bandwidth is set at $h = 0.2$, while the trimming level is set at $v^* = 0.01$ for Z_0 and $v^* = 0$ for Z_{TK} and Z_P. We consider $K = 500$ bootstrap replications for each of the $N = 1000$ simulations.

D.1.2 Testing for the presence of nonlinear probability weighting

If the probability weighting function is truly linear, then the PITs $U_{t+1}(\gamma_0)$ have a standard uniform distribution. Moreover, the MLE is correctly specified and has good finite-sample properties (see Table 1). Therefore, in this subsection, we consider testing the null of linear probability weighting, leveraging the method of Bai (2003). This test uses a Kolmogorov-Smirnov test for the uniformity of the PITs, while using a martingale transformation to correct for the estimation error of the unknown parameter. The correction relies on the score function, which is straightforward to compute under the expected utility model.[1]

As we consider the correct utility function in the Monte Carlo simulations, the only asymptotic cause of rejection is the presence of nonlinear probability weighting. The rejection rates of the Bai (2003) test are displayed in Table 2 below, based on the MLE using various levels of trimming. The empirical size of the test corresponds to the rejection rates under Z_0. They are close to, and generally below, their nominal value for all choices of α and v^*. Meanwhile, the power of the test against the inverse S-shaped alternative Z_{TK} is excellent, but less so for the nearly globally concave alternative Z_P, which is harder to distinguish from the marginal utility function, as discussed above. For a confidence level of $\alpha = 0.05$, the power of the test is slightly increasing with v^*, whereas its size is slightly decreasing.

[1] Estimating the correction term requires reducing the domain in the Kolmogorov-Smirnov statistic from $[0, 1]$ to $[0, 1 - \varepsilon]$ for some $\varepsilon > 0$. Though Bai (2003) suggests taking ε small, we find that this leads to near-singularity of a matrix that needs to be inverted to compute the correction term. However, the simple choice of $\varepsilon = \frac{1}{2}$ results in nearly nominal size, whereas lower values lead to poor size, and higher values to low power. The resulting test focuses on the left half of the distribution, so we apply it to both the cumulative, or primal, and decumulative, or dual, probability weighting functions separately.

Table 2: Testing the probability weighting function

	α (%)	Z_{TK} 10	5	1	Z_P 10	5	1	Z_0 10	5	1
Rejection rate (%)	$v^* = 0$	100	100	100	23.9	14.6	4.1	9.8	5.4	0.8
	$v^* = 0.001$	100	100	100	24.1	14.9	4.1	9.9	5.2	0.8
	$v^* = 0.01$	100	100	100	26.0	15.5	4.2	9.5	4.4	0.7
	$v^* = 0.05$	100	100	99.9	24.5	15.6	3.9	9.0	4.1	0.7

Note: This table displays the rejection rate of the Bai (2003) test at level α. The first two sets of three columns consider the power under two different alternatives, the last set considers a DGP where the null is true and display the size of the test. The rejection rates are based on $N = 1000$ simulations. Critical values are constructed by simulation.

D.2 Smaller sample size

This appendix repeats the analysis of Section 4, but halves the sample size and uses only $T = 150$ observations. The bias is of a relatively similar magnitude as the baseline case of $T = 300$, and the variance is roughly doubled, as should be expected given the asymptotic rate.

Table 3: Profile likelihood performance, smaller sample size

h	$v^*=0$				$v^*=0.001$				$v^*=0.01$				$v^*=0.05$			
	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-0.05	1.77	3.13	1.74	0.20	1.87	3.53	1.90	0.41	2.10	4.58	2.37	0.94	2.01	4.93	2.68
0.20	0.06	1.40	1.95	1.10	0.29	1.48	2.26	1.24	0.54	1.70	3.19	1.65	1.06	1.81	4.39	2.30
0.25	0.22	1.23	1.57	0.85	0.44	1.31	1.90	0.99	0.75	1.51	2.83	1.43	1.32	1.69	4.62	2.30
0.30	0.43	1.18	1.57	0.80	0.64	1.24	1.94	0.96	1.01	1.40	2.97	1.46	1.66	1.55	5.15	2.47
MLE	2.02	1.10	5.29	2.18	2.02	1.10	5.30	2.18	2.01	1.12	5.28	2.16	1.91	1.17	5.03	1.99

(a) Z_{TK}

h	$v^*=0$				$v^*=0.001$				$v^*=0.01$				$v^*=0.05$			
	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	0.08	2.13	4.53	1.41	0.71	2.22	5.42	1.59	1.05	2.49	7.30	2.13	1.32	2.13	6.27	2.24
0.20	0.21	1.90	3.66	1.08	0.77	1.98	4.50	1.25	1.19	2.18	6.17	1.75	1.45	2.13	6.65	2.07
0.25	0.32	1.75	3.17	0.85	0.84	1.82	4.02	1.04	1.30	1.98	5.62	1.54	1.64	2.19	7.50	2.11
0.30	0.43	1.71	3.13	0.76	0.93	1.77	4.00	0.96	1.42	1.91	5.66	1.50	1.87	2.19	8.27	2.21
MLE	2.19	1.61	7.38	1.60	2.19	1.61	7.40	1.60	2.21	1.63	7.53	1.60	2.22	1.68	7.75	1.59

(b) Z_P

h	$v^*=0$				$v^*=0.001$				$v^*=0.01$				$v^*=0.05$			
	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-1.70	2.23	7.85	2.11	-0.77	2.32	5.96	1.68	-0.20	2.67	7.17	1.97	0.87	2.11	5.22	1.85
0.20	-1.72	2.01	6.99	1.82	-0.87	2.08	5.10	1.40	-0.16	2.37	5.66	1.56	0.64	2.20	5.25	1.62
0.25	-1.71	1.85	6.35	1.58	-0.93	1.91	4.53	1.18	-0.14	2.15	4.65	1.28	0.60	2.31	5.68	1.54
0.30	-1.69	1.79	6.06	1.45	-0.94	1.85	4.32	1.07	-0.08	2.07	4.29	1.16	0.73	2.35	6.08	1.53
MLE	0.22	1.71	2.98	0.59	0.23	1.72	2.99	0.58	0.22	1.73	3.06	0.57	0.19	1.77	3.18	0.54

(c) Z_0

Note: Subtables display the bias, standard deviation, and mean squared error (MSE) of the profile likelihood estimator $\hat{\gamma}$ over $N = 1000$ replications of samples of length $T = 150$ for various levels of trimming v^*, with bandwidth $h^* = h/2$ for the CDF estimator. Each subpanel represents a different true probability weighting function. Columns labeled IMSE display the integrated mean squared error of the nonparametric estimator $\widehat{Z}$, multiplied by 1,000. The bottom rows represent the maximum likelihood estimator of the expected utility model, which fixes the weighting function as the identity map.

Table 4: Testing the probability weighting function, smaller sample size

		Z_{TK}			Z_P			Z_0		
	α (%)	10	5	1	10	5	1	10	5	1
Rejection rate (%)	$v^*=0$	98.6	96.8	87.3	15.7	8.2	1.4	8.9	5.1	1.0
	$v^*=0.001$	98.7	96.9	87.3	15.6	8.3	1.3	8.8	4.8	1.0
	$v^*=0.01$	98.7	97.2	87.7	15.8	8.7	1.4	7.9	4.7	0.9
	$v^*=0.05$	98.8	97.3	88.0	16.7	9.1	1.7	8.3	3.5	0.8

Note: This table displays the rejection rate of the Bai (2003) test at level α. The first two sets of three columns consider the power under two different alternatives, the last set considers a DGP where the null is true and display the size of the test. The rejection rates are based on $N = 1000$ simulations. Critical values are constructed by simulation.

D.3 Weekly observations

This appendix considers the same DGP as Section 4, but changes the observation scheme to match weekly data instead of monthly. Historic weekly data does not extend as far back as monthly data, so we consider a sample of 12 years, or $T = 624$. This is not a *ceteris paribus* sample size increase, as the shortened observation span affects the shape of the distributions.

Table 5: Profile likelihood performance, weekly observations

	$v^* = 0$				$v^* = 0.001$				$v^* = 0.01$				$v^* = 0.05$			
h	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-0.05	1.76	3.09	0.41	0.38	1.82	3.47	0.48	0.74	1.97	4.44	0.60	1.00	2.09	5.40	0.72
0.20	0.24	1.44	2.14	0.31	0.60	1.48	2.55	0.38	0.98	1.62	3.58	0.50	1.22	1.88	5.01	0.67
0.25	0.55	1.27	1.92	0.29	0.86	1.30	2.42	0.36	1.26	1.44	3.65	0.51	1.55	1.71	5.32	0.71
0.30	0.89	1.20	2.22	0.33	1.17	1.21	2.84	0.41	1.58	1.35	4.33	0.58	1.96	1.55	6.24	0.81
MLE	3.29	1.09	12.04	1.34	3.30	1.09	12.08	1.34	3.30	1.10	12.07	1.34	3.22	1.14	11.69	1.29

(a) Z_{TK}

	$v^* = 0$				$v^* = 0.001$				$v^* = 0.01$				$v^* = 0.05$			
h	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	0.55	2.24	5.32	0.39	1.52	2.24	7.32	0.52	1.98	2.44	9.84	0.66	1.89	2.53	9.98	0.71
0.20	0.91	2.02	4.93	0.34	1.77	2.01	7.18	0.49	2.24	2.18	9.79	0.63	2.18	2.47	10.89	0.73
0.25	1.19	1.85	4.83	0.32	1.97	1.83	7.22	0.47	2.44	2.00	9.95	0.63	2.41	2.36	11.41	0.74
0.30	1.43	1.78	5.22	0.33	2.15	1.75	7.72	0.49	2.62	1.93	10.58	0.65	2.64	2.28	12.13	0.78
MLE	3.92	1.64	18.03	1.00	3.94	1.61	18.08	1.00	4.00	1.62	18.64	1.03	4.09	1.67	19.48	1.06

(b) Z_P

	$v^* = 0$				$v^* = 0.001$				$v^* = 0.01$				$v^* = 0.05$			
h	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-2.82	2.32	13.38	0.71	-1.49	2.33	7.68	0.46	-0.68	2.58	7.13	0.45	0.05	2.46	6.07	0.44
0.20	-2.76	2.11	12.07	0.62	-1.58	2.11	6.96	0.40	-0.70	2.31	5.85	0.36	-0.16	2.50	6.26	0.41
0.25	-2.71	1.91	10.98	0.55	-1.63	1.90	6.26	0.35	-0.72	2.11	4.97	0.31	-0.13	2.43	5.93	0.38
0.30	-2.64	1.82	10.30	0.51	-1.63	1.80	5.92	0.32	-0.71	2.03	4.63	0.28	0.05	2.39	5.73	0.36
MLE	0.01	1.66	2.77	0.14	0.04	1.65	2.73	0.14	0.06	1.68	2.83	0.14	0.04	1.72	2.96	0.13

(c) Z_0

Note: Subtables display the bias, standard deviation, and mean squared error (MSE) of the profile likelihood estimator $\hat{\gamma}$ over $N = 1000$ replications of samples of length $T = 624$ for various levels of trimming v^*, with bandwidth $h^* = h/2$ for the CDF estimator. The observation window mimics weekly data, as opposed to the monthly frequency used in the baseline case. Each subpanel represents a different true probability weighting function. Columns labeled IMSE display the integrated mean squared error of the nonparametric estimator $\widehat{Z}$, multiplied by 1,000. The bottom rows represent the maximum likelihood estimator of the expected utility model, which fixes the weighting function as the identity map.

Table 6: Testing the probability weighting function, weekly observations

	α (%)	Z_{TK} 10	5	1	Z_P 10	5	1	Z_0 10	5	1
Rejection rate (%)	$v^* = 0$	100	100	100	49.4	37.9	18.3	9.9	5.6	1.3
	$v^* = 0.001$	100	100	100	49.9	38.3	17.9	9.8	5.4	1.2
	$v^* = 0.01$	100	100	100	51.9	39.6	19.5	8.8	4.8	1.3
	$v^* = 0.05$	100	100	100	48.5	37.2	17.6	8.3	4.0	1.1

Note: This table displays the rejection rate of the Bai (2003) test at level α. The first two sets of three columns consider the power under two different alternatives, the last set considers a DGP where the null is true and display the size of the test. The rejection rates are based on $N = 1000$ simulations. Critical values are constructed by simulation.

D.4 Different CRRA parameter: Higher risk aversion

This appendix considers the same DGP as Section 4, but increases the CRRA parameter to $\gamma_0 = 10$ from its baseline value of $\gamma_0 = 2$. This is perhaps not an empirically realistic case; we study it primarily to investigate the robustness of our estimator to relatively extreme values of γ_0. The interval for the parameter under trimming is set to $\Theta = [0, 15]$, to maintain the same width relative to the baseline case.

Table 7: Profile likelihood performance, higher risk aversion

h	$v^*=0$ Bias	St.d.	MSE	IMSE	$v^*=0.001$ Bias	St.d.	MSE	IMSE	$v^*=0.01$ Bias	St.d.	MSE	IMSE	$v^*=0.05$ Bias	St.d.	MSE	IMSE
0.15	-0.27	1.30	1.76	1.02	-0.18	1.27	1.64	0.95	0.01	1.45	2.10	1.15	-0.36	1.38	2.03	1.22
0.20	-0.16	1.23	1.54	0.94	-0.03	1.01	1.02	0.60	0.11	1.14	1.32	0.74	-0.02	1.23	1.51	0.90
0.25	0.03	1.12	1.25	0.76	0.14	0.88	0.80	0.45	0.24	1.00	1.07	0.59	0.27	1.14	1.37	0.80
0.30	0.28	0.95	0.98	0.54	0.36	0.84	0.83	0.44	0.42	0.94	1.05	0.57	0.49	1.03	1.31	0.76
MLE	1.92	0.76	4.26	1.88	1.92	0.76	4.26	1.88	1.93	0.76	4.30	1.89	2.00	0.80	4.65	2.02

(a) Z_{TK}

h	$v^*=0$ Bias	St.d.	MSE	IMSE	$v^*=0.001$ Bias	St.d.	MSE	IMSE	$v^*=0.01$ Bias	St.d.	MSE	IMSE	$v^*=0.05$ Bias	St.d.	MSE	IMSE
0.15	-0.13	1.47	2.18	0.72	0.05	1.50	2.25	0.78	0.08	1.73	2.98	0.97	-0.87	1.38	2.66	1.08
0.20	0.02	1.33	1.76	0.54	0.17	1.34	1.82	0.60	0.13	1.49	2.25	0.72	-0.32	1.39	2.04	0.80
0.25	0.14	1.22	1.51	0.42	0.28	1.22	1.57	0.48	0.18	1.36	1.88	0.58	0.05	1.44	2.09	0.74
0.30	0.26	1.19	1.48	0.37	0.39	1.18	1.55	0.43	0.25	1.31	1.79	0.53	0.26	1.44	2.15	0.72
MLE	2.04	1.12	5.41	1.26	2.05	1.11	5.42	1.26	2.08	1.11	5.55	1.29	2.14	1.15	5.93	1.34

(b) Z_P

h	$v^*=0$ Bias	St.d.	MSE	IMSE	$v^*=0.001$ Bias	St.d.	MSE	IMSE	$v^*=0.01$ Bias	St.d.	MSE	IMSE	$v^*=0.05$ Bias	St.d.	MSE	IMSE
0.15	-1.91	1.54	6.02	1.52	-1.54	1.59	4.91	1.31	-1.29	1.84	5.04	1.35	-1.33	1.37	3.64	1.20
0.20	-1.90	1.40	5.58	1.36	-1.57	1.44	4.53	1.17	-1.33	1.61	4.38	1.15	-1.08	1.45	3.26	1.00
0.25	-1.89	1.29	5.22	1.22	-1.58	1.31	4.22	1.04	-1.36	1.46	3.97	1.02	-0.89	1.55	3.19	0.91
0.30	-1.85	1.24	4.99	1.13	-1.57	1.26	4.05	0.96	-1.35	1.40	3.77	0.95	-0.75	1.59	3.09	0.86
MLE	0.07	1.18	1.40	0.28	0.08	1.18	1.39	0.28	0.08	1.18	1.40	0.28	0.06	1.21	1.47	0.27

(c) Z_0

Note: Subtables display the bias, standard deviation, and mean squared error (MSE) of the profile likelihood estimator $\hat{\gamma}$ over $N = 1000$ replications of samples of length $T = 300$ for various levels of trimming v^*, with bandwidth $h^* = h/2$ for the CDF estimator. Each subpanel represents a different true probability weighting function. Columns labeled IMSE display the integrated mean squared error of the nonparametric estimator $\widehat{Z}$, multiplied by 1,000. The bottom rows represent the maximum likelihood estimator of the expected utility model, which fixes the weighting function as the identity map.

Table 8: Testing the probability weighting function, higher risk aversion

		Z_{TK}			Z_P			Z_0		
	α (%)	10	5	1	10	5	1	10	5	1
Rejection rate (%)	$v^*=0$	100	100	100	23.9	14.6	4.1	9.9	5.4	0.8
	$v^*=0.001$	100	100	100	24.2	15.1	4.1	9.8	5.1	0.7
	$v^*=0.01$	100	100	100	25.8	15.8	4.3	9.2	4.2	0.8
	$v^*=0.05$	100	100	99.8	23.1	13.8	3.7	8.2	4.2	0.7

Note: This table displays the rejection rate of the Bai (2003) test at level α. The first two sets of three columns consider the power under two different alternatives, the last set considers a DGP where the null is true and display the size of the test. The rejection rates are based on $N = 1000$ simulations. Critical values are constructed by simulation.

D.5 Different $\mathbb{P}$-dynamics

This appendix considers a different DGP as Section 4, using instead the following $\mathbb{P}$-dynamics

$$\begin{aligned} \mathrm{d}\log F_t &= \left(-\tfrac{1}{2}V_t - \mu_J\lambda_t\right)\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_{1,t} + J_t\mathrm{d}N_t, \\ \mathrm{d}V_t &= 8(0.015 - V_t)\mathrm{d}t + 0.3\sqrt{V_t}\mathrm{d}W_{2,t} + J_t^V 1_{\{J_t<0\}}\mathrm{d}N_t, \end{aligned} \tag{D.1}$$

where $W_{1,t}$ and $W_{2,t}$ are correlated with coefficient $\rho = -0.6$, and $\lambda_t = 30V_t$. Jump parameters, and utility and probability weighting functions are the same as in Section 4. This model does not satisfy Assumption I, as there is only one state variable. Nonetheless, the utility parameter does seem to be identified.

Table 9: Profile likelihood performance, different $\mathbb{P}$-dynamics

	$v^*=0$				$v^*=0.001$				$v^*=0.01$				$v^*=0.05$			
h	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-0.15	1.64	2.70	0.83	0.12	1.69	2.86	0.89	0.30	1.87	3.58	1.09	0.66	1.85	3.87	1.25
0.20	0.04	1.30	1.70	0.53	0.27	1.34	1.86	0.58	0.51	1.48	2.46	0.76	0.80	1.68	3.46	1.09
0.25	0.26	1.13	1.35	0.42	0.47	1.16	1.57	0.49	0.76	1.29	2.25	0.69	1.11	1.54	3.61	1.11
0.30	0.50	1.07	1.39	0.42	0.71	1.09	1.69	0.52	1.03	1.21	2.53	0.76	1.50	1.37	4.15	1.25
MLE	2.27	0.94	6.01	1.58	2.27	0.94	6.03	1.58	2.26	0.94	6.00	1.57	2.20	0.99	5.80	1.50

(a) Z_{TK}

	$v^*=0$				$v^*=0.001$				$v^*=0.01$				$v^*=0.05$			
h	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	0.20	2.00	4.02	0.68	0.79	1.99	4.58	0.79	1.24	2.27	6.67	1.08	1.24	2.00	5.53	1.08
0.20	0.39	1.81	3.41	0.54	0.92	1.79	4.06	0.66	1.40	1.99	5.92	0.94	1.38	2.08	6.21	1.08
0.25	0.54	1.64	2.98	0.44	1.03	1.62	3.70	0.57	1.52	1.80	5.54	0.86	1.56	2.09	6.80	1.12
0.30	0.69	1.58	2.96	0.42	1.15	1.56	3.75	0.56	1.65	1.72	5.68	0.86	1.80	2.04	7.39	1.19
MLE	2.57	1.42	8.62	1.20	2.59	1.42	8.70	1.21	2.61	1.43	8.84	1.22	2.66	1.48	9.25	1.25

(b) Z_P

	$v^*=0$				$v^*=0.001$				$v^*=0.01$				$v^*=0.05$			
h	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE	Bias	St.d.	MSE	IMSE
0.15	-2.14	2.12	9.06	1.23	-1.21	2.11	5.91	0.87	-0.49	2.42	6.11	0.94	0.29	1.97	3.98	0.82
0.20	-2.14	1.93	8.31	1.09	-1.29	1.93	5.40	0.76	-0.48	2.14	4.79	0.73	-0.03	2.12	4.51	0.77
0.25	-2.12	1.75	7.58	0.97	-1.34	1.75	4.88	0.66	-0.48	1.92	3.92	0.60	-0.06	2.20	4.83	0.76
0.30	-2.08	1.67	7.13	0.89	-1.35	1.68	4.64	0.60	-0.45	1.83	3.56	0.54	0.12	2.21	4.91	0.76
MLE	-0.01	1.55	2.39	0.28	0.01	1.54	2.38	0.28	0.02	1.56	2.44	0.27	0.01	1.60	2.56	0.27

(c) Z_0

Note: Subtables display the bias, standard deviation, and mean squared error (MSE) of the profile likelihood estimator $\hat{\gamma}$ over $N = 1000$ replications of samples of length $T = 300$ for various levels of trimming v^*, with bandwidth $h^* = h/2$ for the CDF estimator. Each subpanel represents a different true probability weighting function. Columns labeled IMSE display the integrated mean squared error of the nonparametric estimator $\widehat{Z}$, multiplied by 1,000. The bottom rows represent the maximum likelihood estimator of the expected utility model, which fixes the weighting function as the identity map.

Table 10: Testing the probability weighting function, different $\mathbb{P}$-dynamics

		Z_{TK}			Z_P			Z_0		
	α (%)	10	5	1	10	5	1	10	5	1
Rejection rate (%)	$v^*=0$	100	100	100	27.4	19.6	6.3	9.9	4.9	0.6
	$v^*=0.001$	100	100	100	27.5	19.8	6.3	9.6	4.8	0.6
	$v^*=0.01$	100	100	100	28.4	20.3	6.9	9.0	4.3	0.4
	$v^*=0.05$	100	100	100	28.2	20.1	6.7	9.1	4.1	0.7

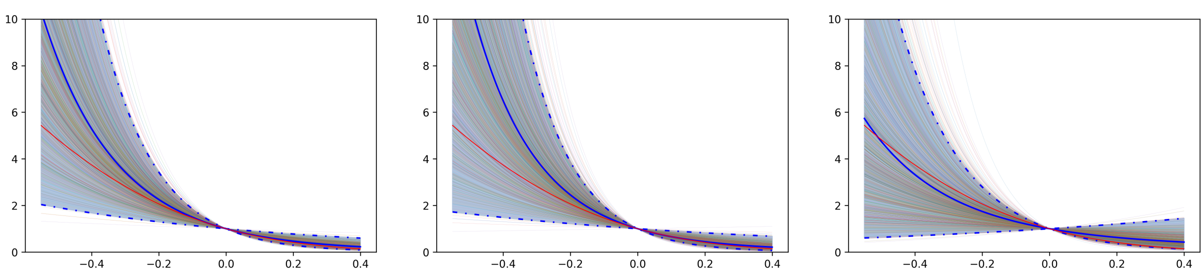
Note: This table displays the rejection rate of the Bai (2003) test at level α. The first two sets of three columns consider the power under two different alternatives, the last set considers a DGP where the null is true and display the size of the test. The rejection rates are based on $N = 1000$ simulations. Critical values are constructed by simulation.

D.6 Misspecified marginal utility

This appendix investigates the effect of potential misspecification in the parametric utility function. In particular, our estimation considers CRRA utility $u'(R;\gamma) = R^{-\gamma}$, while the true utility is of the CARA type $u'(R;\theta_0) = e^{-\theta_0 R}$, with $\theta_0 = 4$. The rest of the simulation set-up is the same as in the baseline case. The tuning parameters $h = 0.2$ and $v^* = 0.001$ are the same as in the empirical analysis.

Figure 1 plots the resulting estimates of the misspecified parametric marginal utility function, and Figure 2 the corresponding nonparametric weighting functions. Misspecification clearly leads to difficulties in fitting the parametric component, as the differences in curvature of the two utility specifications effectively leads to incorrect fitting for either the left or right side of the distribution. For Z_{TK} and Z_P, the misspecification is mainly visible in the left tail, and *vice versa* under Z_0, but the 95% intervals contain the true values over the considered domain in all settings. Importantly, this asymmetry in fit does not lead to a comparable, opposite bias in probabilistic risk aversion: even under misspecification of the parametric component, the estimated probability weighting functions trace the true probability weighting function similarly close as in the correctly specified case of Section 4.

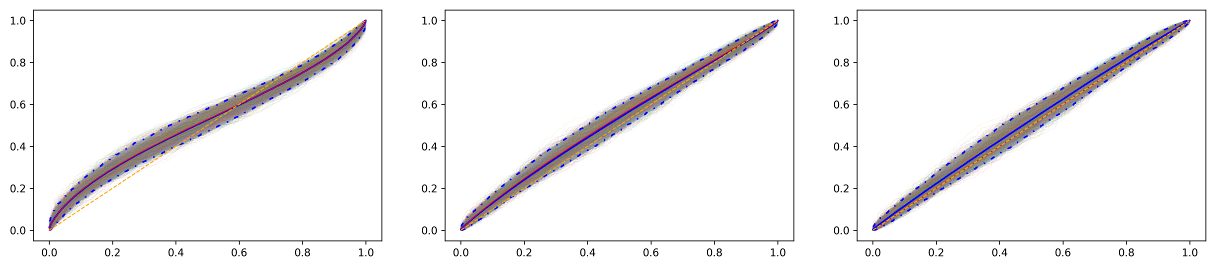
Figure 1: Simulation estimates of the misspecified marginal utility function.



(a) Z_{TK} (b) Z_P (c) Z_0

Note: Subpanels display the parametric estimates of the marginal utility function from a rank-dependent utility model with CRRA utility, when the true utility function is CARA. Marginal utilities are scaled such that $u'(0) = 1$. Tuning parameters are set as $v^* = 0.001$ and $h = 0.2$. Each of the $N = 1000$ lines represents a single simulation. The mean (blue, solid), lower and upper 2.5% percentiles (blue, dash-dotted), and the true marginal utility function (red, solid), are also displayed, along with the 45-degree line (orange, dashed).

Figure 2: Simulation estimates of the probability weighting function, misspecified utility.



(a) Z_{TK} (b) Z_P (c) Z_0

Note: The subpanels display the nonparametric estimates $\widehat{Z}$ of the probability weighting function for $v^* = 0.001$ and $h = 0.2$, for three different true probability weighting functions, when the parametric component of the estimation is misspecified. Each of the $N = 1000$ lines represents a single simulation. The mean (blue, solid), lower and upper 2.5% percentiles (blue, dash-dotted), and the true PW function (red, solid), are also displayed, along with the 45-degree line (orange, dashed).

References

Andrews, D. W. (1992): "Generic uniform convergence," *Econometric Theory*, 8, 241–257.

——— (1994): "Asymptotics for semiparametric econometric models via stochastic equicontinuity," *Econometrica*, 62, 43–72.

——— (1995): "Nonparametric kernel estimation for semiparametric models," *Econometric Theory*, 11, 560–586.

Bai, J. (2003): "Testing parametric conditional distributions of dynamic models," *Review of Economics and Statistics*, 85, 531–549.

Davidson, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricians*, OUP Oxford.

Kristensen, D. (2009): "Uniform convergence rates of kernel estimators with heterogeneous dependent data," *Econometric Theory*, 25, 1433–1445.

Prelec, D. (1998): "The probability weighting function," *Econometrica*, 66, 497–527.

Tversky, A. and D. Kahneman (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Yoshihara, K.-I. (1976): "Limiting behavior of U-statistics for stationary, absolutely regular processes," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 35, 237–252.