Bayesian and Frequentist Inference in Partially Identified Models
Supplementary Material: Proofs and Derivations

Hyungsik Roger Moon
University of Southern California

Frank Schorfheide[1]
University of Pennsylvania
CEPR and NBER

This supplement contains proofs and derivations for results presented in the paper "Bayesian and Frequentist Inference in Partially Identified Models" by H. Moon and F. Schorfheide.

[1]Correspondence: H.R. Moon: Department of Economics, KAP 300, University Park Campus, University of Southern California, Los Angeles, CA 90089. Email: hyungsikmoon@gmail.com. F. Schorfheide: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104. Email: schorf@ssc.upenn.edu.

A Proofs of Theorems and Corollary in Main Text

This section contains proofs for Theorems 1 and 2 as well as Corollary 1. The proof of Theorem 1 requires Lemma A.1 stated below.

Proof of Theorem 1(ii): Since the L_1 distance satisfies the triangle inequality

$$\|P^{\theta}_{Y^n} - P^{\theta}_{\hat{\phi}_n}\| \leq \|P^{\theta}_{Y^n} - P^{\theta}_{N,Y^n}\| + \|P^{\theta}_{N,Y^n} - P^{\theta}_{\hat{\phi}_n}\|$$

it suffices to show that $\|P^{\theta}_{N,Y^n} - P^{\theta}_{\hat{\phi}_n}\| \xrightarrow{\mathbb{P}} 0$.

$$\begin{aligned}
&\left\|P^{\theta}_{N,Y^n} - P^{\theta}_{\hat{\phi}_n}\right\| \\
&\quad\leq \int_{\mathbb{R}^m} \left\|P^{\theta}_{\hat{\phi}_n+\hat{J}_n^{-1/2}D_n^{-1}s} - P^{\theta}_{\hat{\phi}_n}\right\| dN\big(0,I\big)(s) \\
&\quad\leq \int_{\mathbb{R}^m} I\{\|\hat{\phi}_n - \phi_0\| < \delta\} I\{\|\hat{\phi}_n - \phi_0 + \hat{J}_n^{-1/2}D_n^{-1}s\| < \delta\} \left\|P^{\theta}_{\hat{\phi}_n+\hat{J}_n^{-1/2}D_n^{-1}s} - P^{\theta}_{\hat{\phi}_n}\right\| dN\big(0,I\big)(s) \\
&\qquad +2I\{\|\hat{\phi}_n - \phi_0\| \geq \delta\} + 2\int_{\mathbb{R}^m} I\{\|\hat{\phi}_n - \phi_0 + \hat{J}_n^{-1/2}D_n^{-1}s\| \geq \delta\} dN\big(0,I\big)(s) \\
&\quad\leq \int_{\mathbb{R}^m} M(\phi_0,\delta)\|\hat{J}_n^{-1/2}D_n^{-1}s\| dN\big(0,I\big)(s) + 2I\{\|\hat{\phi}_n - \phi_0\| \geq \delta\} \\
&\qquad +2I\{\|\hat{\phi}_n - \phi_0\| \geq \delta/2\} + 2\int_{\mathbb{R}^m} I\{\|\hat{J}_n^{-1/2}D_n^{-1}s\| \geq \delta/2\} dN\big(0,I\big)(s) \\
&\quad\leq M(\phi_0,\delta)\|\hat{J}_n^{-1/2}\|\|D_n^{-1}\| \int_{\mathbb{R}^m} \|s\| dN\big(0,I\big)(s) + o_p(1) \xrightarrow{\mathbb{P}} 0.
\end{aligned}$$

For the second inequality we bound the L_1 distance $\|P^{\theta}_{\hat{\phi}_n+\hat{J}_n^{-1/2}D_n^{-1}s} - P^{\theta}_{\hat{\phi}_n}\|$ by 2 if either $\hat{\phi}_n$ or $\hat{\phi}_n + \hat{J}_n^{-1/2}D_n^{-1}s$ lie outside of the $N_\delta(\phi_0)$ neighborhood. For the third inequality we use the Lipschitz bound of Assumption 2 and the inequality $I\{\|x+y\| \geq \delta\} \leq I\{\|x\| \geq \delta/2\} + I\{\|y\| \geq \delta/2\}$. The last line follows from Assumption 1 that $\hat{\phi}_n$ converges in probability to ϕ_0 , $\|D_n\| \uparrow \infty$, and $\hat{J}_n^{-1/2} = O_p(1)$. A similar argument can be used to establish the convergence of $P^{\theta}_{Y^n}$ to $P^{\theta}_{\phi_0}$. ■

The following Lemma is needed for the subsequent proof of Theorem 2. To simplify the notation let $p_Y(\theta) = p(\theta|Y^n)$, and $p_0(\theta) = p(\theta|\phi_0)$. Similarly, we abbreviate the thresholds κ_{Y^n} and κ_{ϕ_0} by κ_Y and κ_0. A proof is provided in the Online Appendix.

Lemma A.1 *Suppose that* $\int |p_Y(\theta) - p_0(\theta)|d\theta = o_p(1)$ *and* $\int I\{p_0(\theta) = \kappa_0\}p_0(\theta)d\theta = 0$, *where* $\kappa_0 < \infty$. *Then*

$$\int \Big| I\{p_Y(\theta) \geq \kappa_0\} - I\{p_0(\theta) \geq \kappa_0\} \Big| p_Y(\theta)d\theta = o_p(1).$$

Proof of Lemma A.1: (This Lemma is used to prove Theorem 2. Write

$$\begin{aligned}
&\int \Big| I\{p_Y(\theta) \geq \kappa_0\} - I\{p_0(\theta) \geq \kappa_0\} \Big| p_Y(\theta)d\theta \\
&= \int I\{\theta \mid p_Y(\theta) \geq \kappa_0,\ p_0(\theta) < \kappa_0\}p_Y(\theta)d\theta + \int I\{\theta \mid p_Y(\theta) < \kappa_0,\ p_0(\theta) \geq \kappa_0\}p_Y(\theta)d\theta \\
&= \int_{\theta\in A_n} p_Y(\theta)d\theta + \int_{\theta\in B_n} p_Y(\theta)d\theta = (I) + (II),
\end{aligned}$$

say. We will subsequently construct $o_p(1)$ bounds for terms (I) and (II).

Bound for (I)**:** We deduce from the L_1 convergence assumption of $p_Y(\theta)$ to $p_0(\theta)$ that

$$(I) = \int_{\theta\in A_n} p_Y(\theta)d\theta = \int_{\theta\in A_n} p_0(\theta)d\theta + o_p(1) = (Ia) + o_p(1).$$

Thus, it suffices construct an $o_p(1)$ bound for Ia. Define the function

$$f_n(\theta) = p_Y(\theta) - p_0(\theta)$$

and notice that $f_n(\theta) > 0$ for $\theta \in A_n$. With this definition,

$$\begin{aligned}
\int_{A_n} f_n(\theta)p_0(\theta)d\theta &= \int_{A_n} |p_Y(\theta) - p_0(\theta)|p_0(\theta)d\theta \\
&\leq \kappa_0 \int_{A_n} |p_Y(\theta) - p_0(\theta)|d\theta = o_p(1).
\end{aligned} \tag{A.1}$$

The inequality follows from $p_0(\theta) < \kappa_0$ on the set A_n. The $o_p(1)$ statement is a consequence of the assumptions that $p_Y(\theta)$ converges to $p_0(\theta)$ in L_1 and that κ_0 is finite.

Now notice that

$$I\{\theta \in A_n\} = I\Big\{ I\{\theta \in A_n\}f_n(\theta) > 0 \Big\}. \tag{A.2}$$

If $\theta \in A_n$ then $f_n(\theta) > 0$, which means that $I\{\theta \in A_n\}f_n(\theta) > 0$. Moreover, for any $\eta > 0$ we obtain the inequality

$$I\Big\{ I\{\theta \in A_n\}f_n(\theta) > \eta \Big\} \leq \frac{1}{\eta} I\{\theta \in A_n\}f_n(\theta). \tag{A.3}$$

Thus,

$$\begin{aligned}(Ia) &= \int I\Big\{I\{\theta \in A_n\}f_n(\theta) > 0\Big\}p_0(\theta)d\theta \\ &\leq \int I\Big\{I\{\theta \in A_n\}f_n(\theta) > 0\Big\}p_0(\theta)d\theta - \int I\Big\{I\{\theta \in A_n\}f_n(\theta) > \eta\Big\}p_0(\theta)d\theta \\ &\quad +\frac{1}{\eta}\int_{A_n} f_n(\theta)p_0(\theta)d\theta \\ &= \int I\Big\{0 < I\{\theta \in A_n\}f_n(\theta) \leq \eta\Big\}p_0(\theta)d\theta + \frac{1}{\eta}\int_{A_n} f_n(\theta)p_0(\theta)d\theta \\ &= (Ib) + (Ic),\end{aligned}$$

say. The first equality follows from (A.2). The inequality is a consequence of (A.3).

To bound (Ib) notice that

$$I\Big\{0 < I\{\theta \in A_n\}f_n(\theta) \leq \eta\Big\} \leq I\Big\{\kappa_0 - \eta \leq p_0(\theta) \leq \kappa_0 + \eta\Big\}.$$

For the indicator function on the left-hand-side to be one, it has to be the case that $\theta \in A_n$ and $f_n(\theta) \leq \eta$. On the set A_n $p_Y(\theta) \geq \kappa_0$ which leads to

$$\kappa_0 \leq p_Y(\theta) = p_0(\theta) + f_n(\theta) \leq p_0(\theta) + \eta,$$

that is,

$$\kappa_0 - \eta \leq p_0(\theta).$$

Moreover, $p_0(\theta) < \kappa_0 \leq \kappa_0 + \eta$ and therefore the following inequality is satisfied:

$$\kappa_0 - \eta \leq p_0(\theta) \leq \kappa_0 + \eta.$$

Thus,

$$(Ib) \leq \int I\Big\{\kappa_0 - \eta \leq p_0(\theta) \leq \kappa_0 + \eta\Big\}p_0(\theta)d\theta.$$

Based on the Dominated Convergence Theorem and the assumption $\int I\{p_0(\theta) = \kappa_0\}p_0(\theta) = 0$ we deduce that

$$\lim_{\eta \longrightarrow 0} \int I\Big\{\kappa_0 - \eta \leq p_0(\theta) \leq \kappa_0 + \eta\Big\}p_0(\theta)d\theta = \int I\Big\{p_0(\theta) = \kappa_0\Big\}p_0(\theta) = 0. \qquad \text{(A.4)}$$

Notice that our bound for (Ib) is deterministic.

To establish that $(Ia) \xrightarrow{\mathbb{P}} 0$ it suffices to show that for every $\epsilon > 0$ and $\delta > 0$ there exists an $N(\epsilon, \delta)$ such that for $n \geq N(\epsilon, \delta)$

$$\mathbb{P}\{Ia > \epsilon\} \leq \mathbb{P}\{Ib > \epsilon/2\} + \mathbb{P}\{Ic > \epsilon/2\} < \delta.$$

Based on (A.4) we can find an $\eta(\epsilon) > 0$ such that $\mathbb{P}\{Ib > \epsilon/2\} = 0$. To obtain a bound for (Ic) define $Z_n = \int_{A_n} f_n(\theta)p_0(\theta)d\theta$ such that $(Ic) = Z_n/\eta$. According to (A.1), $Z_n = o_p(1)$. Thus, we can find an $N(\epsilon, \delta)$ such that

$$\mathbb{P}\left\{|Z_n| > \eta(\epsilon)\frac{\epsilon}{2}\right\} < \delta$$

whenever $n \geq N(\epsilon, \delta)$, which shows that $(Ia) = o_p(1)$.

Bound for (II): This bound can be obtained following the same steps. Change the definition of $f_n(\theta)$ to

$$f_n(\theta) = p_0(\theta) - p_Y(\theta).$$

Using this definition we obtain that

$$\begin{aligned}
\int_{\theta \in B_n} f_n(\theta)p_Y(\theta)d\theta &= \int_{\theta \in B_n} (p_0(\theta) - p_Y(\theta))p_Y(\theta)d\theta \\
&\leq \kappa_0 \int_{\theta \in B_n} \big|p_0(\theta) - p_Y(\theta)\big|d\theta = o_p(1)
\end{aligned}$$

because on the set B_n the density $p_Y(\theta)$ is bounded by κ_0. Now consider

$$\begin{aligned}
(II) &= \int_{B_n} p_Y(\theta)d\theta = \int I\Big\{I\{\theta \in B_n\}f_n(\theta) > 0\Big\}p_Y(\theta)d\theta \\
&\leq \int I\Big\{I\{\theta \in B_n\}f_n(\theta) > 0\Big\}p_Y(\theta)d\theta - \int I\Big\{I\{\theta \in B_n\}f_n(\theta) > \eta\Big\}p_Y(\theta)d\theta \\
&\quad + \frac{1}{\eta}\int_{B_n} f_n(\theta)p_Y(\theta)d\theta \\
&= \int I\Big\{0 < I\{\theta \in B_n\}f_n(\theta) \leq \eta\Big\}p_0(\theta)d\theta + \frac{1}{\eta}\int_{B_n} f_n(\theta)p_Y(\theta)d\theta + o_p(1) \\
&= (IIb) + (IIc) + o_p(1).
\end{aligned}$$

In the last line we used the L_1 convergence to replace $p_Y(\theta)$ by $p_0(\theta)$ in the definition of term (IIb) which introduces an additional $o_p(1)$ term.

To bound (IIb) notice that

$$I\Big\{0 < I\{\theta \in B_n\} f_n(\theta) \leq \eta\Big\} \leq I\Big\{\kappa_0 - \eta \leq p_n(\theta) \leq \kappa_0 + \eta\Big\}.$$

For the indicator function on the left-hand-side to be one, it has to be the case that $\theta \in B_n$ and $f_n(\theta) \leq \eta$. On the set B_n $p_Y(\theta) < \kappa_0$ which leads to

$$\kappa_0 > p_Y(\theta) = p_0(\theta) - f_n(\theta) \geq p_0(\theta) - \eta.$$

that is,

$$\kappa_0 + \eta \geq p_0(\theta).$$

Moreover, $p_0(\theta) \geq \kappa_0 \geq \kappa_0 - \eta$ and therefore the following inequality is satisfied:

$$\kappa_0 - \eta \leq p_0(\theta) \leq \kappa_0 + \eta.$$

Thus,

$$(IIb) \leq \int I\Big\{\kappa_0 \leq p_0(\theta) < \kappa_0 + \eta\Big\} p_0(\theta) d\theta.$$

Dominated convergence implies that the bound converges to zero as $\eta \longrightarrow 0$. The remaining steps needed to establish that $(II) = o_p(1)$ are identical to the steps followed for term (I). ■

Proof of Theorem 2: Throughout the proof we express the symmetric difference between two sets in terms of indicator functions: $A \ominus B = |I\{x \in A\} - I\{x \in B\}|$. Part (i): To simplify the notation let $p_Y(\theta) = p(\theta|Y^n)$ and $p_0(\theta) = p(\theta|\phi_0)$. Similarly, we abbreviate the thresholds κ_{Y^n} and κ_{ϕ_0} by κ_Y and κ_0. Write

$$\begin{aligned}
&\int \Big| I\{p_Y(\theta) \geq \kappa_Y\} - I\{p_0(\theta) \geq \kappa_0\} \Big| p_Y(\theta) d\theta \\
&= \int \Big| I\{p_Y(\theta) \geq \kappa_Y\} - I\{p_Y(\theta) \geq \kappa_0\} \Big| p_Y(\theta) d\theta \\
&\quad + \int \Big| I\{p_Y(\theta) \geq \kappa_0\} - I\{p_0(\theta) \geq \kappa_0\} \Big| p_Y(\theta) d\theta \\
&= (I) + (II),
\end{aligned}$$

say. In view of our assumptions Lemma A.1 provides an $o_p(1)$ bound for term (II). Now consider term (I). Since by construction

$$\int I\{p_Y(\theta) \geq \kappa_Y\} p_Y(\theta) d\theta = 1 - \tau,$$

we can write term I as

$$\begin{aligned}
(I) &= \int I\Big\{p_Y(\theta) \geq \min\{\kappa_0, \kappa_Y\}\Big\} p_Y(\theta) d\theta - \int I\Big\{p_Y(\theta) \geq \max\{\kappa_0, \kappa_Y\}\Big\} p_Y(\theta) d\theta \\
&= I\{\kappa_0 \geq \kappa_Y\} \left[(1-\tau) - \int I\{p_Y(\theta) \geq \kappa_0\} p_Y(\theta) d\theta\right] \\
&\quad + I\{\kappa_0 < \kappa_Y\} \left[\int I\{p_Y(\theta) \geq \kappa_0\} p_Y(\theta) d\theta - (1-\tau)\right] \\
&= \left| \int I\{p_Y(\theta) \geq \kappa_0\} p_Y(\theta) d\theta - (1-\tau) \right|.
\end{aligned}$$

In order to show that $I = o_p(1)$ we add and subtract $\int I\{p_0(\theta) \geq \kappa_0\} p_Y(\theta) d\theta$ and using the triangle inequality:

$$\begin{aligned}
(I) &\leq \left| \int I\{p_Y(\theta) \geq \kappa_0\} p_Y(\theta) d\theta - \int I\{p_0(\theta) \geq \kappa_0\} p_Y(\theta) d\theta \right| \\
&\quad + \left| \int I\{p_0(\theta) \geq \kappa_0\} p_Y(\theta) d\theta - (1-\tau) \right| \\
&= \left| \int I\{p_Y(\theta) \geq \kappa_0\} p_Y(\theta) d\theta - \int I\{p_0(\theta) \geq \kappa_0\} p_Y(\theta) d\theta \right| \\
&\quad + \left| \int I\{p_0(\theta) \geq \kappa_0\} p_Y(\theta) d\theta - \int I\{p_0(\theta) \geq \kappa_0\} p_0(\theta) d\theta \right| \\
&\leq \int \Big| I\{p_Y(\theta) \geq \kappa_0\} - I\{p_0(\theta) \geq \kappa_0\} \Big| p_Y(\theta) d\theta \\
&\quad + \int I\{p_0(\theta) \geq \kappa_0\} \big| p_Y(\theta) - p_0(\theta) \big| d\theta = o_p(1).
\end{aligned}$$

The first equality holds because $\int I\{p_0(\theta) \geq \kappa_0\} p_0(\theta) d\theta = 1 - \tau$. The final $o_p(1)$ result follows from Lemma A.1 and the L_1 convergence of the posterior densities established in Theorem 1.

Part (ii): The triangle inequality implies that

$$\|P^\theta_{\hat{\phi}_n} - P^\theta_{\phi_0}\| \leq \|P^\theta_{Y^n} - P^\theta_{\hat{\phi}_n}\| + \|P^\theta_{Y^n} - P^\theta_{\phi_0}\| \stackrel{\mathbb{P}}{\longrightarrow} 0$$

by Theorem 1(ii). Let $p_n(\theta) = p(\theta|\hat{\phi}_n)$ and $\kappa_n = \kappa_{\hat{\phi}_n}$. Then using the same argument as for Part (i), replacing $p_Y(\theta)$ by $p_n(\theta)$ and κ_Y by κ_n we can easily establish that

$$\int \left| I\{\theta \in CS^{\theta}_{HPD}(\hat{\phi}_n)\} - I\{\theta \in CS^{\theta}_{HPD}(\phi_0)\} \right| dP^{\theta}_{Y^n} \xrightarrow{\mathbb{P}} 0. \tag{A.5}$$

Now consider the following inequality

$$\begin{aligned} \big|I\{\theta \in A\} - I\{\theta \in B\}\big| &\leq \big|I\{\theta \in A\} - I\{\theta \in C\}\big| + \big|I\{\theta \in B\} - I\{\theta \in C\}\big| \quad \text{(A.6)} \\ &= (I) + (II). \end{aligned}$$

If the left-hand side of (A.6) is zero, then the inequality is trivially satisfied. The left-hand side of (A.6) is one if $\theta \in A$ and $\theta \notin B$ or if $\theta \notin A$ and $\theta \in B$. Since the statement of the inequality is symmetric in A and B we focus on the first case. If $\theta \in A$, $\theta \notin B$, and $\theta \in C$, then $(I) = |1-1| = 0$ and $(II) = |0-1| = 1$. If $\theta \in A$, $\theta \notin B$, and $\theta \notin C$, then $(I) = |1-0| = 1$ and $(II) = |0+0| = 0$. We deduce that whenever the left-hand side of (A.6) is equal to one, the right-hand side is equal to one as well, which confirms the inequality. Now let

$$A = CS^{\theta}_{HPD}(Y^n), \quad B = CS^{\theta}_{HPD}(\hat{\phi}_n), \quad \text{and} \quad C = CS^{\theta}_{HPD}(\phi_0).$$

Integrating both sides of (A.6) yields

$$\begin{aligned} &\int \big|I\{\theta \in A\} - I\{\theta \in B\}\big| p_Y(\theta) d\theta \\ &\leq \int \big|I\{\theta \in A\} - I\{\theta \in C\}\big| p_Y(\theta) d\theta + \int \big|I\{\theta \in B\} - I\{\theta \in C\}\big| p_Y(\theta) d\theta = o_p(1). \end{aligned}$$

The $o_p(1)$ statement follows from Part (i) and (A.5). ■

Proof of Corollary 1: Recall that $\Theta(\hat{\phi}_n) \subset CS^{\theta}_F(Y^n)$ and $CS^{\theta}_{HPD}(Y^n) \subset \Theta$. Part (i) follows from the inequalities

$$\begin{aligned} &P^{\theta}_{Y^n}(CS^{\theta}_{HPD}(Y^n) \backslash CS^{\theta}_F(Y^n)) \\ &\leq P^{\theta}_{Y^n}(\Theta \backslash \Theta(\hat{\phi}_n)) \\ &= 1 - P^{\theta}_{Y^n}(\Theta(\hat{\phi}_n)) \\ &\leq 1 - P^{\theta}_{\hat{\phi}_n}(\Theta(\hat{\phi}_n)) + \left| P^{\theta}_{\hat{\phi}_n}(\Theta(\hat{\phi}_n)) - P^{\theta}_{Y^n}(\Theta(\hat{\phi}_n)) \right| \\ &\xrightarrow{\mathbb{P}} 0. \end{aligned}$$

The probability limit is obtained from $P^{\theta}_{\hat{\phi}_n}(\Theta(\hat{\phi}_n)) = 1$ and Theorem 1(ii).

Part (ii) can be deduced from the following inequalities:

$$\begin{aligned}
& P^{\theta}_{Y^n}(CS^{\theta}_F(Y^n)\backslash CS^{\theta}_{HPD}(Y^n)) \\
& \quad \geq \quad P^{\theta}_{Y^n}(\Theta(\hat{\phi}_n)\backslash CS^{\theta}_{HPD}(Y^n)) \\
& \quad \geq \quad P^{\theta}_{Y^n}(\Theta(\hat{\phi}_n)) - P^{\theta}_{Y^n}(CS^{\theta}_{HPD}(Y^n)) \\
& \quad \geq \quad P^{\theta}_{\hat{\phi}_n}(\Theta(\hat{\phi}_n)) - P^{\theta}_{Y^n}(CS^{\theta}_{HPD}(Y^n)) - \left| P^{\theta}_{Y^n}(\Theta(\hat{\phi}_n)) - P^{\theta}_{\hat{\phi}_n}(\Theta(\hat{\phi}_n)) \right| \\
& \quad \xrightarrow{\mathbb{P}} \quad 1 - (1 - \tau) = \tau.
\end{aligned}$$

The probability limit is obtained from $P^{\theta}_{\hat{\phi}_n}(\Theta(\hat{\phi}_n)) = 1$, $P^{\theta}_{Y^n}(CS^{\theta}_{HPD}(Y^n)) = 1 - \tau$, and Theorem 1(ii). ■

B Derivations of Results Presented in Main Text

This section contains derivations for Section 2, derivations for Remark 2 in Section 3, as well as detailed derivations for the entry game illustration in Section 4.

Derivations for Section 2

Direct calculation of the posterior density of θ:

$$\begin{aligned} p(\theta|Y^n) &= \frac{1}{\sqrt{2\pi/n}} \int_{-\infty}^{\infty} \frac{1}{\lambda} I\{\phi \leq \theta \leq \phi + \lambda\} \exp\left\{-\frac{n}{2}(\phi - \hat{\phi}_n)^2\right\} d\phi \\ &= \frac{1}{\lambda} \frac{1}{\sqrt{2\pi}} \int_{\sqrt{n}(\theta - \hat{\phi}_n - \lambda)}^{\sqrt{n}(\theta - \hat{\phi}_n)} \exp\left\{-\frac{s^2}{2}\right\} ds \\ &= \frac{1}{\lambda}\left[\Phi_N(\sqrt{n}(\theta - \hat{\phi}_n)) - \Phi_N(\sqrt{n}(\theta - \hat{\phi}_n - \lambda))\right]. \end{aligned}$$

The second equality follows from re-arranging the inequalities in the indicator function and the change of variables $s = \sqrt{n}(\phi - \hat{\phi}_n)$. It is straightforward to verify that $p(\theta|Y^n)$ has a single mode at $\theta = \hat{\phi}_n + \lambda/2$ and is symmetric around the mode. ■

Derivations for Section 3

Direct Calculations to Verify Equation (18): We begin with the change of variable $s = \hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n + \tilde{s})$, which leads to

$$\begin{aligned} p(\theta|Y^n) &= p_N(\theta|Y^n) \\ &= \frac{1}{\lambda_n} \int f\left(\frac{\theta - \hat{\phi}_n - \hat{J}_n^{-1/2} D_n^{-1} s}{\lambda_n}\right) \varphi_N(s) ds \\ &= \frac{1}{\lambda_n} |\hat{J}_n^{1/2} D_n| \int_{\tilde{s}=-\lambda_n}^{0} f(-\lambda_n^{-1}\tilde{s}) \varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n + \tilde{s})) d\tilde{s}. \end{aligned}$$

The second equality makes use of the assumption that $f(x) = 0$ outside of the unit interval. The L_1 distance can be bounded as follows:

$$\begin{aligned}
&\int_\theta \left| p_N(\theta|Y^n) - |\hat{J}_n^{1/2} D_n| \varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n)) \right| d\theta \\
&= |\hat{J}_n^{1/2} D_n| \int_\theta \left| \int_{\tilde{s}=-\lambda_n}^{0} \frac{1}{\lambda_n} f(-\lambda_n^{-1}\tilde{s}) \left[\varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n + \tilde{s})) - \varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n)) \right] d\tilde{s} \right| d\theta \\
&\leq |\hat{J}_n^{1/2} D_n| \int_{\tilde{s}=-\lambda_n}^{0} \int_\theta \frac{1}{\lambda_n} f(-\lambda_n^{-1}\tilde{s}) \left| \varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n + \tilde{s})) - \varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n)) \right| d\theta d\tilde{s} \\
&\leq \int_{\tilde{s}=-\lambda_n}^{0} \frac{1}{\lambda_n} f(-\lambda_n^{-1}\tilde{s}) \int_{\tilde{\theta}} \left| \varphi_N(\tilde{\theta} + \hat{J}_n^{1/2} D_n \tilde{s}) - \varphi_N(\tilde{\theta}) \right| d\tilde{\theta} d\tilde{s}. \qquad \text{(B.1)}
\end{aligned}$$

The first equality follows because $\int_0^1 f(x)dx = 1$ and $\varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n))$ does not depend on $\tilde{s}$. The last inequality is based on the change of variables $\tilde{\theta} = \hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n)$.

Now consider the difference $\varphi_N(\tilde{\theta} + h) - \varphi_N(\tilde{\theta})$ for $-\bar{h} \leq h \leq 0$. By direct calculation we obtain

$$\begin{aligned}
|\varphi_N(\tilde{\theta} + h) - \varphi_N(\tilde{\theta})| &= \left| (2\pi)^{-1/2} \exp\left\{ -\frac{1}{2}(\tilde{\theta} + h)^2 \right\} - \varphi_N(\tilde{\theta}) \right| \\
&= \left| \exp\left\{ -\frac{1}{2}(2\tilde{\theta}h + h^2) \right\} - 1 \right| \varphi_N(\tilde{\theta}).
\end{aligned}$$

A first-order Taylor series expansion around $h = 0$ yields

$$\exp\left\{ -\frac{1}{2}(2\tilde{\theta}h + h^2) \right\} - 1 = -(\tilde{\theta} + h_*(\tilde{\theta})) \exp\{-\tilde{\theta} h_*(\tilde{\theta})\} \exp\{-h_*^2(\tilde{\theta})/2\} h,$$

where $-\bar{h} \leq h_*(\tilde{\theta}) \leq 0$. Thus, on the interval $-\bar{h} \leq h \leq 0$ we obtain the bound

$$\left| \exp\left\{ -\frac{1}{2}(2\tilde{\theta}h + h^2) \right\} - 1 \right| \varphi_N(\tilde{\theta}) \leq (|\tilde{\theta}| + \bar{h}) \exp\{-\tilde{\theta}\bar{h} I\{\tilde{\theta} \leq 0\}\} \bar{h} \varphi_N(\tilde{\theta}). \qquad \text{(B.2)}$$

Replacing $\bar{h}$ by $\hat{J}_n^{1/2} D_n \lambda_n$ in (B.2) and combining (B.1) with (B.2) leads to

$$\begin{aligned}
&\int_\theta \left| p_N(\theta|Y^n) - |\hat{J}_n^{1/2} D_n| \varphi_N(\hat{J}_n^{1/2} D_n(\theta - \hat{\phi}_n)) \right| d\theta \\
&\leq \hat{J}_n^{1/2} D_n \lambda_n \int_{\tilde{\theta}} (|\tilde{\theta}| + \hat{J}_n^{1/2} D_n \lambda_n) \exp\{-\tilde{\theta} \hat{J}_n^{1/2} D_n \lambda_n I\{\tilde{\theta} \leq 0\}\} \varphi_N(\tilde{\theta}) d\tilde{\theta} = o_p(1).
\end{aligned}$$

The $o_p(1)$ statement follows because $D_n\lambda_n \longrightarrow 0$ and we can find a finite constant M and an N_M such that for $n > N_M$

$$\int_{\tilde{\theta}} (|\tilde{\theta}| + \hat{J}_n^{1/2} D_n \lambda_n) \exp\{-\tilde{\theta} \hat{J}_n^{1/2} D_n \lambda_n I\{\tilde{\theta} \leq 0\}\} \varphi_N(\tilde{\theta}) d\tilde{\theta} \leq M$$

with probability approaching one. ■

Derivations for Section 4

The probabilities that firm i is profitable as monopolist and duopolist are

$$m_i = \Phi_N(\beta_i) \quad \text{and} \quad d_i = \Phi_N(\beta_i - \gamma_i). \tag{B.3}$$

The relationship between the reduced-form entry probabilities and m_i and d_i, $i = 1, 2$ is given by

$$\begin{aligned} \phi_{11} &= d_1 d_2 && \text{(B.4)} \\ \phi_{00} &= (1 - m_1)(1 - m_2) && \text{(B.5)} \\ \phi_{10} &= m_1(1 - m_2) + d_1(m_2 - d_2) + \psi(m_1 - d_1)(m_2 - d_2) && \text{(B.6)} \\ &= m_1(1 - d_2) - (1 - \psi)(m_1 - d_1)(m_2 - d_2), \end{aligned}$$

where $\psi \in [0, 1]$. The vector of non-redundant reduced form parameters is given by $\phi = [\phi_{11}, \phi_{00}, \phi_{10}]'$ and the structural parameters are $\theta = [\beta_1, \gamma_1, \beta_2, \gamma_2]'$. In addition, there is an auxiliary parameter ψ.

Identified Set

We will now provide a characterization of the identified set $\Theta(\phi)$. Define

$$G(\theta, \alpha) = \begin{bmatrix} G_1(\theta) \\ G_2(\theta) \end{bmatrix} - \begin{bmatrix} 0_{2\times 1} \\ \alpha \end{bmatrix}, \tag{B.7}$$

where

$$G_1(\theta) = \begin{bmatrix} d_1 d_2 \\ (1-m_1)(1-m_2) \end{bmatrix}, \quad G_2(\theta) = m_1(1-d_2).$$

and

$$\alpha = (1-\psi)(m_1-d_1)(m_2-d_2).$$

Moreover, let

$$\bar{\alpha}(\theta) = (m_1-d_1)(m_2-d_2) \tag{B.8}$$

and

$$Q(\theta;\phi) = \min_{0\le\alpha\le\bar{\alpha}(\theta)} \Big\| \phi - G(\theta,\alpha) \Big\|. \tag{B.9}$$

Notice that by construction $Q(\theta;\phi) \geq 0$. In view of (B.4) to (B.6) and (B.7) it is straightforward to verify that the identified set can be characterized as follows:

$$\theta \in \Theta(\phi) \quad \text{iff} \quad Q(\theta;\phi) = 0.$$

Suppose we partition θ into $\theta = [\theta_1', \theta_2']'$. (B.4) and (B.5) imply that conditional on ϕ and θ_1 the subvector θ_2 is uniquely determined. Thus, the dimension of the identified set $\Theta(\phi)$ is 2. Since the entry game is symmetric with respect to Firm 1 and Firm 2, our illustration focuses on inference for θ_1. We denote the identified set for this subvector by $\Theta_1(\phi)$ and it can be characterized by the projection

$$\Theta_1(\phi) = \left\{ \theta_1 \;\middle|\; \exists\, \theta_2 \text{ s.t. } Q([\theta_1', \theta_2']';\phi) = 0 \right\}.$$

Frequentist Inference

Starting point of the frequentist inference is a large sample approximation of the sampling distribution of $\hat{\phi}_n$, defined as

$$\hat{\phi}_n = \left[\frac{n_{11}}{n}, \frac{n_{00}}{n}, \frac{n_{10}}{n} \right]', \tag{B.10}$$

where n_{11} is the number of markets with a duopoly, n_{00} is the number of markets without entry, and n_{10} is the number of markets with a Firm 1 monopoly. We assume that

$$\sqrt{n}(\hat{\phi}_n - \phi) \Longrightarrow N(0, \Lambda(\phi)) \tag{B.11}$$

uniformly in ϕ, where $\Lambda(\phi)$ can be consistently estimated by $\hat{\Lambda}$. Now define

$$Q_n(\theta; \hat{\phi}_n) = \min_{0 \le \alpha \le \bar{\alpha}(\theta)} n \left\| \hat{\phi}_n - G(\theta, \alpha) \right\|_{\hat{\Lambda}^{-1}}. \tag{B.12}$$

We shall construct a confidence set for θ as a level set of $Q_n(\theta; \hat{\phi}_n)$. To do so, we examine the sampling distribution of $Q_n(\theta; \hat{\phi}_n)$ for $\theta \in \Theta(\phi)$.

We partition $\hat{\phi}_n$ into $\hat{\phi}_{1,n}$ and $\hat{\phi}_{2,n}$ where the partitions conform with $G_1(\theta)$ and $G_2(\theta)$. Moreover, define

$$\hat{H}_1(\theta) = \hat{\phi}_{1,n} - G_1(\theta), \quad \hat{H}_2(\theta) = \hat{\phi}_{2,n} - G_2(\theta),$$

and partition $\hat{\Lambda}$ accordingly. In addition, let

$$\hat{H}_{2.11}(\theta) = \hat{H}_2(\theta) - \hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1}\hat{H}_1(\theta), \quad \hat{\Lambda}_{2.11} = \hat{\Lambda}_{22} - \hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1}\hat{\Lambda}_{12}.$$

Using the formula for factorizing a joint normal density into a marginal and a conditional density we can re-write the objective function as

$$Q_n(\theta; \hat{\phi}_n) = \min_{0 \le \alpha \le \bar{\alpha}(\theta)} n \left(\left\| \hat{H}_1(\theta) \right\|_{\hat{\Lambda}_{11}^{-1}} + \left\| \hat{H}_{2.11}(\theta) + \alpha \right\|_{\hat{\Lambda}_{2.11}^{-1}} \right). \tag{B.13}$$

The minimizing value of α which we denote by $\hat{\alpha}(\theta)$ is given by

$$\hat{\alpha}(\theta) = \begin{cases} 0 & \text{if } 0 \le \hat{H}_{2.11}(\theta) \\ -\hat{H}_{2.11}(\theta) & \text{if } -\bar{\alpha}(\theta) \le \hat{H}_{2.11}(\theta) < 0 \\ \bar{\alpha}(\theta) & \text{otherwise} \end{cases}. \tag{B.14}$$

In turn, the objective function becomes

$$Q_n(\theta; \hat{\phi}_n) = \begin{cases} n\left\| \hat{H}_1(\theta) \right\|_{\hat{\Lambda}_{11}^{-1}} + n\left\| \hat{H}_{2.11}(\theta) \right\|_{\hat{\Lambda}_{2.11}^{-1}} & \text{if } 0 \le \hat{H}_{2.11}(\theta) \\ n\left\| \hat{H}_1(\theta) \right\|_{\hat{\Lambda}_{11}^{-1}} & \text{if } -\bar{\alpha}(\theta) \le \hat{H}_{2.11}(\theta) < 0 \\ n\left\| \hat{H}_1(\theta) \right\|_{\hat{\Lambda}_{11}^{-1}} + n\left\| \hat{H}_{2.11}(\theta) + \bar{\alpha}(\theta) \right\|_{\hat{\Lambda}_{2.11}^{-1}} & \text{otherwise} \end{cases}. \tag{B.15}$$

As shown in Andrews and Guggenberger (2009), critical values for the construction of uniformly valid confidence sets can be obtained by considering the behavior of the objective function $Q_n(\cdot)$ under sequences of parameters. To do so, suppose data are generated based on $\phi_n = G(\theta_n, \alpha_n)$. To approximate the distribution of $Q_n(\theta_n; \hat{\phi}_n)$, notice that

$$\begin{aligned}
\hat{H}_1(\theta_n) &= \hat{\phi}_{1,n} - G_1(\theta_n) \\
&= \hat{\phi}_{1,n} - \phi_{1,n} \\
\hat{H}_{2.11}(\theta_n) &= \hat{\phi}_{2,n} - G_2(\theta_n) - \hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1}[\hat{\phi}_{1,n} - G_1(\theta_n)] \\
&= \hat{\phi}_{2,n} - \phi_{2,n} - \alpha_n - \hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1}(\hat{\phi}_{1,n} - \phi_{1,n}).
\end{aligned}$$

Let

$$Z_{1,n} = \sqrt{n}\hat{\Lambda}_{11}^{-1/2}(\hat{\phi}_{1,n} - \phi_{1,n}), \quad Z_{2.11,n} = \sqrt{n}\hat{\Lambda}_{2.11}^{-1/2}[\hat{\phi}_{2,n} - \phi_{2,n} - \hat{\Lambda}_{21}\hat{\Lambda}_{11}^{-1}(\hat{\phi}_{1,n} - \phi_{1,n})].$$

Using this notation, we can rewrite the objective function as

$$Q_n(\theta_n; \hat{\phi}_n) = \begin{cases} \|Z_{1,n}\| + \|Z_{2.11,n} - \sqrt{n}\hat{\Lambda}_{2.11}^{-1/2}\alpha_n\| & \text{if } \sqrt{n}\hat{\Lambda}_{2.11}^{-1/2}\alpha_n \leq Z_{2.11,n} \\ \|Z_{1,n}\| + \|Z_{2.11,n} + \sqrt{n}\hat{\Lambda}_{2.11}^{-1/2}(\bar{\alpha}(\theta_n) - \alpha_n)\| & \text{if } Z_{2.11,n} < -\sqrt{n}\hat{\Lambda}_{2.11}^{-1/2}(\bar{\alpha}(\theta_n) - \alpha_n) \\ \|Z_{1,n}\| & \text{otherwise} \end{cases}. \tag{B.16}$$

Now suppose that $\sqrt{n}\Lambda_{2.11}^{-1/2}\alpha_n \longrightarrow a$, $\sqrt{n}\Lambda_{2.11}^{-1/2}(\bar{\alpha}(\theta_n) - \alpha_n) \longrightarrow \bar{a}$, where $a \in \mathbb{R}^+ \cup \infty$ and $\bar{a} \in \mathbb{R}^+ \cup \infty$. Thus,

$$Q_n(\theta_n; \hat{\phi}_n) \Longrightarrow \begin{cases} \|Z_1\| + \|Z_{2.11} - a\| & \text{if } a \leq Z_{2.11} \\ \|Z_1\| + \|Z_{2.11} + \bar{a}\| & \text{if } Z_{2.11} < -\bar{a} \\ \|Z_1\| & \text{otherwise} \end{cases}, \tag{B.17}$$

where $Z_1 \sim N(0, I_2)$ and $Z_{2.11} \sim N(0,1)$ and Z_1 and $Z_{2.11}$ are independent. We have to distinguish three cases. First,

$$Q_n(\theta_n; \hat{\phi}_n) \Longrightarrow \|Z_1\| \leq \|Z_1\| + \|Z_{2.11}\|I\{Z_{2.11} \geq 0\} \quad \text{if} \quad a = \infty,\ \bar{a} = \infty.$$

Second,

$$Q_n(\theta_n; \hat{\phi}_n) \Longrightarrow \|Z_1\| + \|Z_{2.11} - a\|I\{Z_{2.11} \geq a\} \leq \|Z_1\| + \|Z_{2.11}\|I\{Z_{2.11} \geq 0\} \quad \text{if} \quad a < \infty,\ \bar{a} = \infty.$$

Third,

$$\begin{aligned} Q_n(\theta_n;\hat{\phi}_n) &\Longrightarrow \|Z_1\| + \|Z_{2.11} - a\| I\{Z_{2.11} \geq a\} + \|Z_{2.11} + \bar{a}\| I\{Z_{2.11} < -\bar{a}\} \quad \text{if} \quad a < \infty,\ \bar{a} < \infty \\ &\leq \|Z_1\| + \|Z_{2.11}\|. \end{aligned}$$

The bound for this last case is weaker than the bounds for the first two cases. The case $\bar{a} < 0$ arises only if $\bar{\alpha}(\theta_n) \longrightarrow 0$ sufficiently fast, meaning that θ_n approaches an area of the parameter space in which the model is point identified. From the definition of $\bar{\alpha}(\theta)$ in (B.8) it follows that the third case arises if one of the interaction parameters is close to zero. In our numerical illustration we use a conservative fixed critical value obtained from the $1-\tau$ quantile of a $\chi^2(df = 3)$.

A frequentist confidence set for the 4-dimensional parameter vector θ can then be defined as the level set

$$CS_F^{\theta}(Y^n) = \{\theta \mid Q_n(\theta;\hat{\phi}_n) \leq c_\tau^2\}. \tag{B.18}$$

We are restricting our attention to confidence sets constructed from fixed (rather than sample-size and θ dependent) critical values. In principle, one can construct the set $CS_F^{\theta}(Y^n)$ by evaluating the objective function $Q_n(\theta\ \hat{\phi}_n)$ on a 4-dimensional grid. However, since the identified set $\Theta(\phi)$ lies in a 2-dimensional subspace the specification of a suitable grid is difficult. Moreover, our goal is to construct a confidence set for the subvector θ_1. Thus, we let

$$\underline{Q}_n(\theta_1;\hat{\phi}_n) = \min_{\theta_2}\ Q_n([\theta_1', \theta_2]';\hat{\phi}_n)$$

and define

$$CS_F^{\theta_1}(Y^n) = \{\theta \mid \underline{Q}_n(\theta_1;\hat{\phi}_n) \leq c_\tau^2\}. \tag{B.19}$$

The confidence set $CS_F^{\theta_1}(Y^n)$ is the projection of $CS_F^{\theta}(Y^n)$ onto the domain of θ_1. To compute the projection-based confidence set we specify a 2-dimensional grid for θ_1 and evaluate the objective function $\underline{Q}_n(\theta_1;\hat{\phi}_n)$ for each grid point. A parameter value is included in the confidence set if $\underline{Q}_n(\theta_1;\hat{\phi}_n) \leq c_\tau^2$.

Bayesian Inference – Draws from Conditional Prior

Prior 1 and Prior 2 are specified on the θ - ψ space through densities $p(\theta, \psi)$. These priors induce a prior distribution on the reduced form parameters ϕ. As explained in the main text, the conditional prior of θ given ϕ will not get updated through the likelihood function and the posterior will converge to $p(\theta|\hat{\phi}_n)$. In order to characterize the conditional prior $p(\theta_1|\phi)$ we conduct the following change of variables. Let

$$Z = [\beta_1, \gamma_1, \beta_2, \gamma_2, \psi]' \tag{B.20}$$

and

$$X = [\beta_1, \gamma_1, \phi_{11}, \phi_{00}, \phi_{10}]'. \tag{B.21}$$

To convert a prior density for $Z = f(X)$ into a prior for X, we can use

$$p_X(X) = p_Z(f(X))|f'(X)|. \tag{B.22}$$

Once we have derived $p_X(X)$ we can proceed as follows. Notice that

$$p(\theta_1|\phi) \propto p(\theta_1, \phi). \tag{B.23}$$

We use a Random-Walk Metropolis Algorithm to generate draws from $p(\theta_1|\phi)$. For this algorithm it is sufficient to be able evaluate the joint density $p(\theta_1, \phi)$ numerically. Descriptions of the algorithm can be found in many textbooks, e.g., Geweke (2005). Our proposal density is multivariate Gaussian with a covariance matrix that equals a suitably scaled identity matrix.

We shall proceed by characterizing the function $f(X)$, component by component and then derive the Jacobian $f'(X)$. The following functional relationships will be useful:

$$m_1 = \Phi_N(\beta_1), \quad m_2 = \Phi_N(\beta_2), \quad d_1 = \Phi_N(\beta_1 - \gamma_1), \quad d_2 = \Phi_N(\beta_2 - \gamma_2).$$

Since we will have to solve for β_2 and γ_2, notice that

$$\beta_2 = \Phi_N^{-1}(m_2), \quad \gamma_2 = \Phi_N^{-1}(m_2) - \Phi_N^{-1}(d_2).$$

The Nash equilibrium conditions imply that

$$\begin{aligned}\phi_{00} &= (1-m_1)(1-m_2)\\ \phi_{11} &= d_1 d_2\\ \phi_{10} &= m_1(1-m_2)+d_1(m_2-d_2)+\psi(m_1-d_1)(m_2-d_2).\end{aligned}$$

We can use these conditions to solve for m_2, d_2, and ψ:

$$\begin{aligned}m_2 &= 1-\frac{\phi_{00}}{1-m_1}\\ d_2 &= \frac{\phi_{11}}{d_1}\\ \psi &= \frac{\phi_{10}-m_1(1-m_2)-d_1(m_2-d_2)}{(m_1-d_1)(m_2-d_2)}.\end{aligned}$$

The expression for ψ can be simplified by replacing m_2 and d_2:

$$\begin{aligned}\psi &= \frac{\phi_{10}-m_1(1-m_2)-d_1(m_2-d_2)}{(m_1-d_1)(m_2-d_2)}\\ &= \frac{\phi_{10}-\phi_{00}\frac{m_1}{1-m_1}-d_1\left(1-\frac{\phi_{00}}{1-m_1}-\frac{\phi_{11}}{d_1}\right)}{(m_1-d_1)\left(1-\frac{\phi_{00}}{1-m_1}-\frac{\phi_{11}}{d_1}\right)}\\ &= \frac{\phi_{10}(1-m_1)-\phi_{00}m_1-d_1\left(1-m_1-\phi_{00}-\frac{\phi_{11}(1-m_1)}{d_1}\right)}{(m_1-d_1)\left(1-m_1-\phi_{00}-\frac{\phi_{11}(1-m_1)}{d_1}\right)}\\ &= \frac{\phi_{10}(1-m_1)-\phi_{00}m_1-d_1 g(X)}{(m_1-d_1)g(X)},\end{aligned}$$

where

$$g(X)=\left(1-m_1-\phi_{00}-\frac{\phi_{11}(1-m_1)}{d_1}\right).$$

Combining terms, we obtain the following expressions for the components of $f(X)$:

$$\begin{aligned}f_1(X) &= \beta_1\\ f_2(X) &= \gamma_1\\ f_3(X) &= \Phi_N^{-1}\left(1-\frac{\phi_{00}}{1-\Phi_N(\beta_1)}\right)\\ f_4(X) &= f_3(X)-\Phi_N^{-1}\left(\frac{\phi_{11}}{\Phi_N(\beta_1-\gamma_1)}\right)\\ f_5(X) &= \frac{A_5(X)}{B_5(X)}=\frac{\phi_{10}(1-\Phi_N(\beta_1))-\phi_{00}\Phi_N(\beta_1)-\Phi_N(\beta_1-\gamma_1)g(X)}{(\Phi_N(\beta_1)-\Phi_N(\beta_1-\gamma_1))g(X)}\end{aligned}$$

where

$$g(X) = \left(1 - \Phi_N(\beta_1) - \phi_{00} - \frac{\phi_{11}(1 - \Phi_N(\beta_1))}{\Phi_N(\beta_1 - \gamma_1)}\right).$$

Now we can calculate the derivatives for the jacobian matrix. For this define

$$\psi(z) = \frac{\partial \Phi_N^{-1}(z)}{\partial z} = \frac{1}{\phi_N(\Phi_N^{-1}(z))}.$$

Term $f_1(X)$:

$$\frac{\partial f_1(X)}{\partial \beta_1} = 1.$$

Term $f_2(X)$:

$$\frac{\partial f_2(X)}{\partial \gamma_1} = 1.$$

Term $f_3(X)$:

$$\begin{aligned}
\frac{\partial f_3(X)}{\partial \beta_1} &= -\psi\left(1 - \frac{\phi_{00}}{1 - \Phi_N(\beta_1)}\right) \frac{\phi_{00}}{[1 - \Phi_N(\beta_1)]^2} \phi_N(\beta_1) \\
\frac{\partial f_3(X)}{\partial \phi_{00}} &= -\psi\left(1 - \frac{\phi_{00}}{1 - \Phi_N(\beta_1)}\right) \frac{1}{1 - \Phi_N(\beta_1)}.
\end{aligned}$$

Term $f_4(X)$:

$$\begin{aligned}
\frac{\partial f_4(X)}{\partial \beta_1} &= \frac{\partial f_3(X)}{\partial \beta_1} + \psi\left(\frac{\phi_{11}}{\Phi_N(\beta_1 - \gamma_1)}\right) \frac{\phi_{11}\phi_N(\beta_1 - \gamma_1)}{\Phi_N^2(\beta_1 - \gamma_1)} \\
\frac{\partial f_4(X)}{\partial \gamma_1} &= -\psi\left(\frac{\phi_{11}}{\Phi_N(\beta_1 - \gamma_1)}\right) \frac{\phi_{11}\phi_N(\beta_1 - \gamma_1)}{\Phi_N^2(\beta_1 - \gamma_1)} \\
\frac{\partial f_4(X)}{\partial \phi_{11}} &= -\psi\left(\frac{\phi_{11}}{\Phi_N(\beta_1 - \gamma_1)}\right) \frac{1}{\Phi_N(\beta_1 - \gamma_1)} \\
\frac{\partial f_4(X)}{\partial \phi_{00}} &= \frac{\partial f_3(X)}{\partial \phi_{00}}.
\end{aligned}$$

Term $f_5(X)$:

$$\frac{\partial f_5(X)}{\partial x} = \frac{\frac{\partial A(X)}{\partial x} B(X) - A(X) \frac{\partial B(X)}{\partial x}}{B(X)^2}.$$

Term $A(X)$:

$$\begin{aligned}
\frac{\partial A(X)}{\partial \beta_1} &= -(\phi_{10}+\phi_{00})\phi_N(\beta_1) - \phi_N(\beta_1-\gamma_1)g(X) - \Phi_N(\beta_1-\gamma_1)\frac{\partial g(X)}{\partial \beta_1} \\
\frac{\partial A(X)}{\partial \gamma_1} &= \phi_N(\beta_1-\gamma_1)g(X) - \Phi_N(\beta_1-\gamma_1)\frac{\partial g(X)}{\partial \gamma_1} \\
\frac{\partial A(X)}{\partial \phi_{11}} &= -\Phi_N(\beta_1-\gamma_1)\frac{\partial g(X)}{\partial \phi_{11}} \\
\frac{\partial A(X)}{\partial \phi_{00}} &= -\Phi_N(\beta_1) - \Phi_N(\beta_1-\gamma_1)\frac{\partial g(X)}{\partial \phi_{00}} \\
\frac{\partial A(X)}{\partial \phi_{10}} &= (1-\Phi_N(\beta_1)) - \Phi_N(\beta_1-\gamma_1)\frac{\partial g(X)}{\partial \phi_{10}}.
\end{aligned}$$

Term $B(X)$:

$$\begin{aligned}
\frac{\partial B(X)}{\partial \beta_1} &= (\phi_N(\beta_1) - \phi_N(\beta_1-\gamma_1))g(X) + (\Phi_N(\beta_1) - \Phi_N(\beta_1-\gamma_1))\frac{\partial g(X)}{\partial \beta_1} \\
\frac{\partial B(X)}{\partial \gamma_1} &= \phi_N(\beta_1-\gamma_1)g(X) + (\Phi_N(\beta_1) - \Phi_N(\beta_1-\gamma_1))\frac{\partial g(X)}{\partial \gamma_1} \\
\frac{\partial B(X)}{\partial \phi_{11}} &= (\Phi_N(\beta_1) - \Phi_N(\beta_1-\gamma_1))\frac{\partial g(X)}{\partial \phi_{11}} \\
\frac{\partial B(X)}{\partial \phi_{00}} &= (\Phi_N(\beta_1) - \Phi_N(\beta_1-\gamma_1))\frac{\partial g(X)}{\partial \phi_{00}}.
\end{aligned}$$

Term $g(X)$:

$$\begin{aligned}
\frac{\partial g(X)}{\partial \beta_1} &= -\phi_N(\beta_1) + \frac{\phi_{11}\phi_N(\beta_1)}{\Phi_N(\beta_1-\gamma_1)} + \frac{\phi_{11}(1-\Phi_N(\beta_1))\phi_N(\beta_1-\gamma_1)}{\Phi_N^2(\beta_1-\gamma_1)} \\
\frac{\partial g(X)}{\partial \gamma_1} &= -\frac{\phi_{11}(1-\Phi_N(\beta_1))\phi_N(\beta_1-\gamma_1)}{\Phi_N^2(\beta_1-\gamma_1)} \\
\frac{\partial g(X)}{\partial \phi_{11}} &= -\frac{1-\Phi_N(\beta_1)}{\Phi_N(\beta_1-\gamma_1)} \\
\frac{\partial g(X)}{\partial \phi_{00}} &= -1.
\end{aligned}$$

Bayesian Inference – Draws from Posterior

According to Equations (B.3) to (B.6) we can express the reduced form probabilities as functions of θ and ψ. Thus, the likelihood function is given by

$$p(Y^n|\theta,\psi) = \phi_{11}^{n_{11}}(\theta,\psi)\phi_{00}^{n_{00}}(\theta,\psi)\phi_{10}^{n_{10}}(\theta,\psi)\phi_{01}^{n_{01}}(\theta,\psi). \tag{B.24}$$

If this prior distribution is combined with a prior specified on the θ - ψ space, then the posterior is given by

$$p(\theta, \psi|Y^n) \propto p(Y^n|\theta, \psi)p(\theta, \psi) \tag{B.25}$$

and draws can be generated with a Random Walk Metropolis Algorithm.

In addition to Priors 1 and 2 we consider a prior that is flat with respect to the reduced form parameters. Conditional on ϕ, the prior for θ_1 is uniform on the identified set $\Theta_1(\phi)$. In order to obtain draws from the posterior distribution of θ_1 we sample from (i) $p(\phi|Y^n)$ and (ii) from $p(\theta_1|\phi)$. For Step (i) notice that under the flat prior for ϕ, the posterior distribution $P^{\phi}_{Y^n}$ takes the form of a Dirichlet distribution

$$[\phi_{11}, \phi_{00}, \phi_{10}, \phi_{01}]' \sim \text{Dirichlet } (n_{11}+1, n_{00}+1, n_{10}+1, n_{01}).$$

A draw from this Dirichlet distribution can be generated as follows: Let $a_j \sim \mathcal{G}(n_j + 1, 1)$, where $j \in \{11, 00, 10, 01\}$ and $\mathcal{G}(\alpha, 1)$ denotes a Gamma distribution with shape parameter α and scale parameter 1. Then set

$$\phi = [a_{11}, a_{00}, a_{10}, a_{01}]'/(a_{11} + a_{00} + a_{10} + a_{01}).$$

For Step (ii) we specify a two-dimensional grid for θ_1 in order construct projections of the identified set $\Theta_1(\phi)$ onto the β_1 and γ_1 ordinates. Let these projections be delimited by $\underline{\beta}_1$, $\bar{\beta}_1$, $\underline{\gamma}_1$, and $\bar{\gamma}_1$. We then use an acceptance sampler with a proposal density that is uniform on $[\underline{\beta}_1, \bar{\beta}_1] \otimes [\underline{\gamma}_1, \bar{\gamma}_1]$ to obtain a draw of θ_1 conditional on ϕ.

Bayesian Inference – Credible Sets

Credible sets are computed according to the following steps:

1. Construct two independent sequences $\{\theta^{(1)}_{1,s}\}^S_{s=1}$ and $\{\theta^{(2)}_{1,s}\}^S_{s=1}$ of draws from the distribution of θ_1.

2. Use the $\{\theta^{(1)}_{1,s}\}^S_{s=1}$ draws to construct Kernel density estimates $\hat{p}(\theta^{(2)}_{1,s})$ for each $\theta^{(2)}_{1,s}$, $s = 1, \ldots, S$.

3. Find a cutoff κ such that $(1-\tau)S$ of the density estimates $\hat{p}(\theta_{1,s}^{(2)})$ are greater or equal than κ.

4. Use the $\{\theta_{1,s}^{(1)}\}_{s=1}^{S}$ draws to construct Kernel density estimates $\hat{p}(\theta_1)$ for values of θ_1 on a 2-dimensional grid. Include a particular grid point into the credible set if $\hat{p}(\theta_1) \geq \kappa$.