

NOTE ON MEASURABILITY PROBLEM IN BAYESIAN INFERENCE ON PARTIALLY IDENTIFIED MODELS

KENGO KATO

ABSTRACT. This note is concerned with a measurability problem that appears in Bayesian inference on partially identified models developed recently by Kitagawa (2012) and Kline and Tamer (2015), which boils down to proving certain measurability of the composition map of the singleton-valued and set-valued functions. We provide fairly general conditions that guarantee the measurability of the composition map.

1. INTRODUCTION

There has been growing interest in Bayesian inference on partially identified models. Recent contributions include Kitagawa (2012) and Kline and Tamer (2015)¹. Their approach is based on drawing posterior probability statements on the identified set for a function of the partially identified parameter from a posterior distribution for the point-identified reduced form parameter by inverting the composition map of the singleton-valued and set-valued functions. Importantly, this approach enables us to construct Bayesian credible regions for the identified set and these credible regions enjoy frequentist validity under suitable regularity conditions.

However, as pointed out in Kline and Tamer (2015), it is not clear under what conditions the set-valued composition map is suitably measurable in the sense that posterior probabilities of events about the identified set make sense. The purpose of this note is to argue that these posterior probabilities make sense under fairly general conditions, and in most cases we do not need to worry about this measurability problem. The following results might be known (or obvious) to the experts in the area of random set theory, but the author could not find a right reference.

2. RESULTS

We partly follow the setup in Kline and Tamer (2015). Suppose that there is a partially identified parameter $\theta \in \Theta$ together with a reduced form parameter $\mu \in M$ that can be point identified, and they satisfy the structural

Date: This version: September 2, 2016.

K. Kato is supported by Grant-in-Aid for Scientific Research (C) (15K03392) from the JSPS.

¹See Kitagawa (2012) and Kline and Tamer (2015) for the literature review on this topic.

restriction $\theta \in \Gamma(\mu)$ for some (known) correspondence $\Gamma : M \rightrightarrows \Theta$, that is, Γ is a set-valued map from M into the collection of subsets of Θ . We refer the reader to Arintis and Border (2006, Chapters 17 and 18) for details about correspondences. The identified set for θ at given $\mu \in M$ is thus $\Gamma(\mu)$. The goal is to make inference on $\Delta(\theta)$ for some (known) map $\Delta : \Theta \rightarrow D$, and the identified set for $\Delta(\theta)$ at given μ is

$$(\Delta \circ \Gamma)(\mu) := \Delta(\Gamma(\mu)) := \{\Delta(\theta) : \theta \in \Gamma(\mu)\}.$$

Suppose that a posterior distribution $\Pi(\cdot \mid X)$ for μ conditionally on data X is given.

Kitagawa (2012) and Kline and Tamer (2015) study Bayesian inference on the identified set for $\Delta(\theta)$ based on posterior probabilities such as

$$\Pi\{(\Delta \circ \Gamma)^u(B) \mid X\} \quad \text{and} \quad \Pi\{(\Delta \circ \Gamma)^\ell(B) \mid X\}$$

for $B \subset D$, where $(\Delta \circ \Gamma)^u(B)$ and $(\Delta \circ \Gamma)^\ell(B)$ are the upper and lower inverses of B by $\Delta \circ \Gamma$ (Arintis and Border, 2006, p.557)², respectively, i.e.,

$$(\Delta \circ \Gamma)^u(B) := \{\mu \in M : (\Delta \circ \Gamma)(\mu) \subset B\},$$

$$(\Delta \circ \Gamma)^\ell(B) := \{\mu \in M : (\Delta \circ \Gamma)(\mu) \cap B \neq \emptyset\}.$$

See Kitagawa (2012) for the interpretation of these posterior probabilities. Kline and Tamer (2015) point out that verifying measurability of these sets is a non-trivial technical problem.

Putting the measurability problem aside, the posterior probability of the form $\Pi\{(\Delta \circ \Gamma)^u(B) \mid X\}$ is of particular importance, since for a given $\alpha \in (0, 1)$, a set $B_{1-\alpha} = B_{1-\alpha}(X) \subset D$ (that may depend on X) verifying $\Pi\{(\Delta \circ \Gamma)^u(B_{1-\alpha}) \mid X\} \geq 1 - \alpha$ can be interpreted as a Bayesian credible region for the identified set for $\Delta(\theta)$, and enjoys frequentist validity under suitable regularity conditions (see Kitagawa, 2012; Kline and Tamer, 2015).

We consider the setting more general than Kitagawa (2012) and Kline and Tamer (2015). We begin with introduction of some definitions. A *Polish space* is a separable, completely metrizable topological space. A separable metrizable topological space S is called a *Suslin space* if there exist a Polish space T and a Borel measurable map from T onto S . We refer the reader to Dudley (2002, Section 13.1) for the equivalent definitions³. The Borel isomorphism theorem (Dudley, 2002, Theorem 13.1) yields that any Borel set in a Polish space equipped with the relative topology is Suslin. A Suslin space equipped with its Borel σ -field will be called a *Suslin measurable space*.

We will make the following assumptions:

- (i) $(\Theta, \mathcal{B}(\Theta))$ is a Suslin measurable space.
- (ii) $(M, \mathcal{M})$ is a measurable space.

²For a correspondence $\varphi : S \rightrightarrows T$, the upper and lower inverses of a subset $A \subset T$ are defined by $\varphi^u(A) := \{x \in S : \varphi(x) \subset A\}$ and $\varphi^\ell(A) := \{x \in S : \varphi(x) \cap A \neq \emptyset\}$, respectively. If φ is singleton-valued, then these inverses coincide with the usual inverse.

³For applications of Suslin (or analytic) sets in econometrics and mathematical economics, we refer to the illuminating paper by Stinchcomb and White (1992).

- (iii) $(D, \mathcal{D})$ is a measurable space.
- (iv) $\Gamma : M \twoheadrightarrow \Theta$ is a correspondence such that $\text{graph}(\Gamma) := \{(\theta, \mu) \in \Theta \times M : \theta \in \Gamma(\mu)\} \in \mathcal{B}(\Theta) \otimes \mathcal{M}$.
- (v) $\Delta : \Theta \rightarrow D$ is a measurable map.

These assumptions are fairly general. Comparisons with conditions in Kitagawa (2012) and Kline and Tamer (2015) will be discussed after the proof of Proposition 2 below.

Proving measurability of sets $(\Delta \circ \Gamma)^u(B)$ and $(\Delta \circ \Gamma)^\ell(B)$ boils down to proving measurability of the correspondence $\Delta \circ \Gamma : M \twoheadrightarrow D$. For a correspondence $\varphi : \Omega \twoheadrightarrow S$ from a measurable space $(\Omega, \mathcal{F})$ into another measurable space $(S, \mathcal{S})$, φ is said to be $\mathcal{F}/\mathcal{S}$ -measurable if $\varphi^\ell(B) = \{\omega \in \Omega : \varphi(\omega) \cap B \neq \emptyset\} \in \mathcal{F}$ for all $B \in \mathcal{S}$ (see Aprintis and Border, 2006, p.592, for various notions of measurability for correspondences). Because $\varphi^u(B) = \{\omega \in \Omega : \varphi(\omega) \subset B\} = (\varphi^\ell(B^c))^c$, φ is $\mathcal{F}/\mathcal{S}$ -measurable if and only if $\varphi^u(B) \in \mathcal{F}$ for all $B \in \mathcal{S}$. For a measurable space $(\Omega, \mathcal{F})$, a subset $A \subset \Omega$ is said to be *universally measurable* if for every probability measure P on $\mathcal{F}$, A is measurable for the completion of P . The collection of all universally measurable sets in Ω is a σ -field containing $\mathcal{F}$ and called the *universal completion* of $\mathcal{F}$.

The following proposition is the main result of this note.

Proposition 1. *The correspondence $\Delta \circ \Gamma : M \twoheadrightarrow D$ is $\overline{\mathcal{M}}/\mathcal{D}$ -measurable, where $\overline{\mathcal{M}}$ is the universal completion of $\mathcal{M}$.*

Before proving Proposition 1, we shall verify that a posterior distribution can be extended to the universal completion.

Proposition 2. *Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$ be measurable spaces, and let $P_x, x \in S$ be a class of probability measures on $\mathcal{F}$ dominated by a probability measure λ and such that $S \ni x \mapsto P_x(A)$ is measurable for each $A \in \mathcal{F}$. Let $\overline{\mathcal{F}}$ denote the universal completion of $\mathcal{F}$. Then there exist probability measures $\hat{P}_x, x \in S$ on $\overline{\mathcal{F}}$ such that the restriction of $\hat{P}_x$ to $\mathcal{F}$ coincides with P_x for each $x \in S$, and such that $S \ni x \mapsto \hat{P}_x(A)$ is measurable for each $A \in \overline{\mathcal{F}}$.*

Remark 1. If a set A in Ω is universally measurable, then for each $x \in S$ there exists a set $\hat{A} \in \mathcal{F}$ such that $(A \setminus \hat{A}) \cup (\hat{A} \setminus A)$ is a P_x -null set, and thus we may extend P_x by $P_x(A) = P_x(\hat{A})$. However, the set $\hat{A}$ in this discussion may depend on x , and so measurability of $x \mapsto P_x(A)$ does not directly follow.

Proof of Proposition 2. For each $A \in \overline{\mathcal{F}}$, there exists a set $\hat{A} \in \mathcal{F}$ such that $(A \setminus \hat{A}) \cup (\hat{A} \setminus A)$ is a λ -null set, and define $\hat{P}_x(A) = P_x(\hat{A})$ for each $x \in S$. Since $P_x \ll \lambda$ for each $x \in S$, $\hat{P}_x$ is a well-defined probability measure on $\overline{\mathcal{F}}$. Further, it is obvious from its construction that $x \mapsto \hat{P}_x(A)$ is measurable for each $A \in \overline{\mathcal{F}}$. $\square$

Suppose now that X takes values in a measurable space $(S, \mathcal{S})$. By definition, a posterior distribution for μ conditionally on X is a regular conditional

probability $(S, \mathcal{M}) \ni (x, A) \mapsto \Pi(A \mid x) = \Pi_x(A)$, that is, $A \mapsto \Pi_x(A)$ is a probability measure on $\mathcal{M}$ for each fixed x and $x \mapsto \Pi_x(A)$ is measurable for each fixed A . Suppose that there is a probability measure that dominates Π_x for all $x \in S$ (for example, take the prior distribution on μ). Proposition 2 then guarantees that $\Pi_x, x \in \mathcal{X}$ can be extended to a regular conditional probability on $(S, \overline{\mathcal{M}})$.

Remark 2 (Comparisons with Kitagawa (2012) and Kline and Tamer (2015)). Lemma 3 in the supplemental appendix of Kline and Tamer (2015) provides sufficient conditions that guarantee measurability of the correspondence $\Delta \circ \Gamma$. They assume that Θ, M , and D are (Borel subsets of) finite dimensional Euclidean spaces, Θ is closed, $\Gamma(\mu)$ is of the form $\Gamma(\mu) = \{\theta : Q(\theta, \mu) = 0\}$ for some continuous function $Q : \Theta \times M \rightarrow [0, \infty)$, and Δ is continuous. These conditions are completely covered by ours. It is worthwhile to point out that the proof of their Lemma 5 extends to the case where the parameter spaces are infinite dimensional under the additional assumption that Θ is σ -compact, but a σ -compact subset in an infinite dimensional norm space (say) is rather small. Kitagawa (2012) considers this measurability problem in Lemma A.1. He (implicitly) assumes that the parameters spaces are Polish, and the correspondence $\Delta \circ \Gamma$ maps M into the collection of closed sets in D . He then applies Theorem 2.6 in Molchanov (2005) to prove measurability of $\Delta \circ \Gamma$ (precisely speaking he should assume that $(\Delta \circ \Gamma)(\mu)$ is regular closed in order to apply Theorem 2.6 in Molchanov (2005); recall that a set in a topological space is said to be regular closed if it coincides with the closure of its interior. For example, the boundary of a non-empty open set is closed but not regular closed). However, the assumption that $\Delta \circ \Gamma$ takes values in the collection of closed sets in D would not be straightforward to verify.

We turn to proving Proposition 1. The proof relies on the following projection theorem of Sainte-Beuve (1974). See Dudley (2002, Theorem 5.3.2) for a textbook exposition.

Theorem 1 (Sainte-Beuve). *Let $(\Omega, \mathcal{F})$ be a measurable space and let $(S, \mathcal{S})$ be a Suslin measurable space. Then for every $A \in \mathcal{S} \otimes \mathcal{F}$,*

$$\{\omega \in \Omega : (x, \omega) \in A \text{ for some } x \in S\}$$

is universally measurable in Ω .

Proof of Proposition 1. Because $(\Delta \circ \Gamma)^u(B) = \Gamma^u(\Delta^{-1}(B))$, it suffices to prove $\overline{\mathcal{M}}/\mathcal{B}(\Theta)$ -measurability of the correspondence Γ , but this follows from the fact that for each $A \in \mathcal{B}(\Theta)$, $\Gamma^\ell(A)$ is expressed as $\Gamma^\ell(A) = \{\mu : (\theta, \mu) \in \text{graph}(\Gamma) \cap (A \times M) \text{ for some } \theta \in \Theta\}$, and Theorem 1 yields that $\Gamma^\ell(A)$ is universally measurable in M . $\square$

Remark 3. Theorem 1 provides conditions under which the supremum of a stochastic process is measurable. Let $(\Omega, \mathcal{F})$ be a measurable space and let $Y(t, \omega), t \in T, \omega \in \Omega$ be a stochastic process indexed by a Suslin measurable

space T . If $Y(t, \omega)$ is in addition jointly measurable in (t, ω) , then for every $r \in \mathbb{R}$, $\{\omega : \sup_{t \in T} Y(t, \omega) > r\} = \{\omega : Y(t, \omega) > r \text{ for some } t \in T\}$ is universally measurable in Ω by Theorem 1, and so $\sup_t Y(t)$ is measurable with respect to the universal completion of $\mathcal{F}$. See Dudley (1999, Section 5.3) for further details and applications to empirical process theory. Projection theorems such as Theorem 1 are closely connected with measurable selection theorems (see Dudley, 1999, Theorem 5.3.2), which provide conditions for existence of measurable extrema (cf. Brown and Purves, 1973; Stinchcomb and White, 1992).

REFERENCES

- Aprintis, C.D. and Border, K.C. (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide (3rd Edition)*. Springer.
- Brown, L.D. and Purves, R. (1973). Measurable selections of extrema. *Ann. Statist.* **5** 902-912.
- Dudley, R.M. (1999). *Uniform Central Limit Theorems*. Cambridge University Press.
- Dudley, R.M. (2002). *Real Analysis and Probability*. Cambridge University Press.
- Kitagawa, T. (2012). Estimation and inference for set-identified parameters using posterior lower probability. Working paper.
- Kline, B. and Tamer, E. (2015). Bayesian inference in a class of partially identified models. *Quantitative Economics*, to appear.
- Molchanov, I. (2005). *Theory of Random Sets*. Springer.
- Sainte-Beuve, M.-F. (1974). On the extension of the von Neumann-Aumann's theorem. *J. Functional Anal.* **17** 112-129.
- Stinchcomb, M.B. and White, H. (1992). Some measurability results for extrema of random functions over random sets. *Rev. Econ. Stud.* **59** 495-512.

(K. Kato) GRADUATE SCHOOL OF ECONOMICS, UNIVERSITY OF TOKYO 7-3-1 HONGO, BUNKYO-KU, TOKYO 113-0033, JAPAN.

E-mail address: `kkato@e.u-tokyo.ac.jp`