

ARTICLE

# On Accuracy and Coherence with Infinite Opinion Sets

Mikayla Kelley

Department of Philosophy, Stanford University, Stanford, CA, US  
Email: [mikelley@stanford.edu](mailto:mikelley@stanford.edu)

(Received 09 September 2020; revised 20 August 2021; accepted 23 November 2021; first published online 12 January 2022)

## Abstract

There is a well-known equivalence between avoiding accuracy dominance and having probabilistically coherent credences (see, e.g., de Finetti 1974; Joyce 2009; Predd et al. 2009; Pettigrew 2016). However, this equivalence has been established only when the set of propositions on which credence functions are defined is finite. In this article, I establish connections between accuracy dominance and coherence when credence functions are defined on an infinite set of propositions. In particular, I establish the necessary results to extend the classic accuracy argument for probabilism to certain classes of infinite sets of propositions, including countably infinite partitions.

## 1. Introduction

A central norm in the epistemology of partial belief is probabilism: a person's degrees of belief—or *credences*—should satisfy the laws of probability.<sup>1</sup> There is a long tradition in the spirit of Savage (1971) and de Finetti (1974) of appealing to the epistemic virtue of accuracy to justify probabilism (also see Rosenkrantz 1981). One particular form of argument is the accuracy dominance argument for probabilism introduced by Joyce (1998). Let a set  $\mathcal{F}$  of propositions be an *opinion set* and a function  $c : \mathcal{F} \rightarrow [0, 1]$  be a *credence function* on  $\mathcal{F}$ . Let a credence function be *coherent* if it satisfies the axioms of probability. A credence function  $c'$  on  $\mathcal{F}$  *accuracy dominates* a credence function  $c$  on  $\mathcal{F}$  if  $c$  is more inaccurate than  $c'$ , no matter how the world turns out to be (where inaccuracy is precisified as in Section 2). Then the existing accuracy dominance arguments purport to vindicate probabilism by showing that a credence function is not accuracy dominated if and only if it is coherent.

---

<sup>1</sup> This article is based on work done by Kelley (2019).

However, there is a limitation to almost all of the literature on accuracy dominance arguments for probabilism: the opinion set is assumed to be finite.<sup>2</sup> Indeed, de Finetti (1974), Lindley (1987), Joyce (1998, 2009), Predd et al. (2009), Leitgeb and Pettigrew (2010a, b), and Pettigrew (2016) all establish their dominance results only for finite opinion sets.<sup>3</sup> In this paper, I remove this assumption and prove dominance results that I hope to be useful in evaluating the extent to which accuracy dominance arguments for probabilism succeed when the opinion set is infinite.

I begin in Section 2 by reviewing the mathematical framework and the standard dominance result for finite opinion sets. Sections 3–5 are concerned with accuracy and coherence in the infinite setting. In Sections 3 and 4, I make headway on characterizing the opinion sets and accuracy measures for which there is an equivalence between coherence and avoiding dominance as in the finite case. In Section 5, I extend the accuracy framework to the uncountable setting and prove that coherence is necessary to avoid dominance on uncountable opinion sets. I conclude in Section 6 with a discussion of the results established in Sections 3–5.

## 2. The finite case

I first set up the framework that will be used throughout the article. Fix a set  $W$  (not necessarily finite) that represents the set of *possible worlds* and, for now, a finite set  $\mathcal{F} \subseteq \mathcal{P}(W)$  of *propositions* that represents an *opinion set*—the set of propositions that an agent has beliefs about.

**Definition 2.1.** An *algebra* over  $W$  is a subset  $\mathcal{F}^* \subseteq \mathcal{P}(W)$  such that:

1.  $W \in \mathcal{F}^*$ ;
2. if  $p, p' \in \mathcal{F}^*$ , then  $p \cup p' \in \mathcal{F}^*$ ; and
3. if  $p \in \mathcal{F}^*$ , then  $W \setminus p \in \mathcal{F}^*$ .

**Definition 2.2.**

- i. A *credence function* on an opinion set  $\mathcal{F}$  is a function from  $\mathcal{F}$  to  $[0, 1]$ .
- ii. A credence function  $c$  is *coherent* if it can be extended to a finitely additive probability function on an algebra  $\mathcal{F}^*$  over  $W$  containing  $\mathcal{F}$ . That is, there is an algebra  $\mathcal{F}^* \supseteq \mathcal{F}$  over  $W$  and a function  $c^* : \mathcal{F}^* \rightarrow [0, 1]$  such that:

<sup>2</sup> In an unpublished manuscript, Walsh (ms) proves an accuracy dominance result in the countably infinite context, to which we return in Section 3. In a related but distinct area, Huttegger (2013) and Easwaran (2013) extend to the infinite setting part of the literature on using the minimization of expected inaccuracy to vindicate epistemic principles. See, for example, Greaves and Wallace (2005). Schervish et al. (2014) prove that in certain countably infinite cases, coherence is sufficient to avoid *strong dominance*. Schervish et al. (2009) and Steeger (2019) explore a different way to weaken the assumption that the opinion set is finite. We will return to their work in Section 4.

<sup>3</sup> The same holds for accuracy dominance results pertaining to approximating coherence (De Bona and Staffel 2018; Staffel 2019, chap. 5) and accuracy dominance results that significantly weaken the additivity assumption on the measure of inaccuracy (Pettigrew 2022; Nielsen 2022), although neither will be my focus here.

- (a)  $c^*(p) = c(p)$  for all  $p \in \mathcal{F}$ ;
- (b)  $c^*(p \cup p') = c^*(p) + c^*(p')$  for  $p, p' \in \mathcal{F}^*$  with  $p \cap p' = \emptyset$ ; and
- (c)  $c^*(W) = 1$ .

iii. A credence function that is not coherent is *incoherent*.

**Remark 2.3.** If  $\mathcal{F} = \{p_1, \dots, p_n\}$ , a credence function  $c$  over  $\mathcal{F}$  can be identified with the vector  $(c(p_1), \dots, c(p_n)) \in [0, 1]^n$ . Thus, the space of all credence functions over  $\mathcal{F}$  can be identified with  $[0, 1]^n \subseteq \mathbb{R}^n$ . We often simplify notation by setting  $c_i := c(p_i)$ .

We introduce an important subclass of the class of all credence functions, namely, the (coherent) credence functions that match the truth values of  $\mathcal{F}$  at a world  $w$  exactly.

**Definition 2.4.** Fix an opinion set  $\mathcal{F}$ . For each  $w \in W$ , let  $v_w : \mathcal{F} \rightarrow \{0, 1\}$  be defined by  $v_w(p) = 1$  if and only if  $w \in p$ . We call  $v_w$  the *omniscient credence function at world  $w$* . We let  $\mathcal{V}_{\mathcal{F}}$  denote the set of all omniscient credence functions on  $\mathcal{F}$ . Note that  $|\mathcal{V}_{\mathcal{F}}| \leq 2^{|\mathcal{F}|}$ .

Next, we specify the inaccuracy measures that we will be concerned with in this section. Fix a finite opinion set  $\mathcal{F}$ , and let  $\mathcal{C}$  denote the set of credence functions on  $\mathcal{F}$ . We define an *inaccuracy measure* to be a function of the form

$$\mathcal{I} : \mathcal{C} \times W \rightarrow [0, \infty].$$

The class of inaccuracy measures we consider is a generalization of the class defended by Pettigrew (2016): the inaccuracy measures defined in terms of what we call a *quasi-additive Bregman divergence*. It is a subclass of the inaccuracy measures assumed by Predd et al. (2009).<sup>4</sup>

**Definition 2.5.** Suppose  $\mathcal{D} : [0, 1]^n \times [0, 1]^n \rightarrow [0, \infty]$ .

1.  $\mathcal{D}$  is a *divergence* if  $\mathcal{D}(x, y) \geq 0$  for all  $x, y \in [0, 1]^n$  with equality if and only if  $x = y$ .
2.  $\mathcal{D}$  is *quasi-additive* if there exists a function  $\mathfrak{d} : [0, 1]^2 \rightarrow [0, \infty]$  and a sequence of elements  $\{a_i\}_{i=1}^n$  from  $(0, \infty)$  such that

$$\mathcal{D}(x, y) = \sum_{i=1}^n a_i \mathfrak{d}(x_i, y_i),$$

in which case we say  $\mathcal{D}$  is *generated* by  $\mathfrak{d}$  and  $\{a_i\}_{i=1}^n$ .

3.  $\mathcal{D}$  is a *quasi-additive Bregman divergence* if  $\mathcal{D}$  is a quasi-additive divergence generated by  $\mathfrak{d}$  and  $\{a_i\}_{i=1}^n$ , and in addition, there is a function  $\varphi : [0, 1] \rightarrow \mathbb{R}$  such that:

---

<sup>4</sup> Using terminology from Definition 2.5, Predd et al. (2009) consider a more general class in allowing different one-dimensional Bregman divergences for different propositions.

- (a)  $\varphi$  is continuous and strictly convex on  $[0, 1]$ ;<sup>5</sup>
- (b)  $\varphi$  is continuously differentiable on  $(0, 1)$  with the formal definition

$$\varphi'(i) := \lim_{x \rightarrow i} \varphi'(x)$$

for  $i \in \{0, 1\}$ ;<sup>6</sup> and

- (c) for all  $x, y \in [0, 1]$ , we have

$$\mathfrak{D}(x, y) = \varphi(x) - \varphi(y) - \varphi'(y)(x - y).$$

We call such a  $\mathfrak{D}$  a *one-dimensional Bregman divergence*.

We take the inaccuracy of a credence function  $c$  at a world  $w$  to be the distance between  $c$  and the omniscient credence function  $v_w$ , where distance is measured with a quasi-additive Bregman divergence.

**Definition 2.6.** Let a *legitimate inaccuracy measure* be an inaccuracy measure given by

$$\mathcal{F}(c, w) = \mathfrak{D}(v_w, c),$$

where  $\mathfrak{D}$  is a quasi-additive Bregman divergence.

By allowing different weights depending on the proposition, we can accommodate the intuition that some propositions are more important to know than others.<sup>7</sup> Even if one thinks that inaccuracy measures should be additive, as Pettigrew (2016) does, relaxing this restriction makes our results more widely relevant. A popular example of an additive legitimate inaccuracy measure is the Brier score (see Section 12, “Homage to the Brier Score,” in Joyce 2009):

$$\mathcal{F}(c, w) = \sum_{i=1}^n (v_w(p_i) - c(p_i))^2.$$

**Remark 2.7.** The class of additive Bregman divergences is the class of additive and continuous *strictly proper scoring rules* (See Pettigrew [2016, 66]; also see, for example, Banerjee et al. [2005] and Gneiting and Raftery [2007] for more details on Bregman divergences, as well as their connection to strictly proper scoring rules).

We now recall the dominance result connecting coherence to accuracy dominance when the opinion set is finite. It was first proved for the Brier score by de Finetti (1974, 87–90) and extended to any legitimate inaccuracy measure by Predd et al.

<sup>5</sup> Therefore,  $\varphi$  is bounded because it is a continuous function on a compact interval.

<sup>6</sup> We do not require  $\varphi'(i) < \infty$  for  $i \in \{0, 1\}$ .

<sup>7</sup> Although see Levinstein 2019 for an argument that one should expect weights to vary with respect to worlds as well as propositions.

(2009).<sup>8</sup> See Schervish et al. (2009) and Pettigrew (2022) (see also Nielsen 2022) for further generalizations of the finite result.

**Definition 2.8.** For each pair of credence functions  $c, c^*$  over  $\mathcal{F}$ :

1.  $c^*$  *weakly dominates*  $c$  relative to an inaccuracy measure  $\mathcal{I}$  if  $\mathcal{I}(c, w) \geq \mathcal{I}(c^*, w)$  for all  $w \in W$  and  $\mathcal{I}(c, w) > \mathcal{I}(c^*, w)$  for some  $w \in W$ ;
2.  $c^*$  *strongly dominates*  $c$  relative to  $\mathcal{I}$  if  $\mathcal{I}(c, w) > \mathcal{I}(c^*, w)$  for all  $w \in W$ .

**Theorem 2.9.** (de Finetti 1974; Predd et al. 2009) Let  $\mathcal{F}$  be a finite opinion set,  $\mathcal{I}$  be a legitimate inaccuracy measure, and  $c$  be a credence function on  $\mathcal{F}$ . Then the following are equivalent:

1.  $c$  is not strongly dominated;
2.  $c$  is not weakly dominated;
3.  $c$  is coherent.

Further, if  $c$  is incoherent, then  $c$  is strongly dominated by a coherent credence function.

On the basis of Theorem 2.9, authors in the accuracy literature conclude that an incoherent credence function is objectionable because there is an undominated coherent credence function that does strictly better in terms of accuracy, no matter how the world turns out to be, whereas coherent credence functions are not accuracy dominated in this way. Because it is the basis of the accuracy dominance argument for probabilism in the finite case, Theorem 2.9 is the result we would like to extend to infinite opinion sets. We now make progress toward this goal when  $\mathcal{F}$  is countably infinite.

*Note that any missing proofs in Sections 3–5 can be found in the appendix.*

### 3. The countably infinite case: Coherence is necessary

#### 3.1 Generalized legitimate inaccuracy measures

We begin with a discussion of how to measure inaccuracy in the countably infinite setting. Fix a countably infinite opinion set  $\mathcal{F}$  over a set  $W$  of worlds (of arbitrary cardinality). Let  $\mathcal{C}$  be the set of credence functions over  $\mathcal{F}$ , which can be identified with  $[0, 1]^\infty$  (see Remark 2.3). An *inaccuracy measure* remains a map from  $\mathcal{C} \times W$  to  $[0, \infty]$ .

The class of inaccuracy measures that we use is defined in terms of generalizations of quasi-additive Bregman divergences.

**Definition 3.1.** Suppose  $\mathcal{D} : [0, 1]^\infty \times [0, 1]^\infty \rightarrow [0, \infty]$ . Then we call  $\mathcal{D}$  a *generalized quasi-additive Bregman divergence* if

---

<sup>8</sup> See Section 7 of Predd et al. (2009).

$$\mathfrak{D}(x, y) = \sum_{i=1}^{\infty} a_i \mathfrak{d}(x_i, y_i),$$

where  $\mathfrak{d}$  is a bounded<sup>9</sup> one-dimensional Bregman divergence as in Definition 2.5.3 and  $\{a_i\}_{i=1}^{\infty}$  is a sequence of elements from  $(0, \infty)$  with  $\sup_i a_i < \infty$ .<sup>10</sup>

**Remark 3.2.** Note that  $\mathfrak{d}$ —defined in terms of  $\varphi$ —being bounded is equivalent to  $\varphi'$  being bounded on  $[0, 1]$ . Further, we may assume that  $\varphi(0) = \varphi'(0) = 0$  because  $\mathfrak{d}_{\varphi} = \mathfrak{d}_{\bar{\varphi}}$  if  $\varphi$  and  $\bar{\varphi}$  differ by a linear function.<sup>11</sup>

In the appendix, I show that generalized quasi-additive Bregman divergences are examples of what Csiszár (1995) calls *Bregman distances*, which are generalizations of quasi-additive Bregman divergences defined on spaces of nonnegative functions.

We consider the suggestively named inaccuracy measures:

**Definition 3.3.** Given an enumeration of  $\mathcal{F}$ ,<sup>12</sup> let a *generalized legitimate inaccuracy measure* be an inaccuracy measure  $\mathcal{F} : \mathcal{C} \times W \rightarrow [0, \infty]$  given by

$$\mathcal{F}(c, w) = \mathfrak{D}(v_w, c) \tag{1}$$

for  $\mathfrak{D}$  a generalized quasi-additive Bregman divergence.<sup>13</sup>

Notice that the Brier score extends to a generalized legitimate inaccuracy measure, namely, the squared  $\ell^2(\mathcal{F})$  norm

$$\mathcal{F}(c, w) = \|v_w - c\|_{\ell^2(\mathcal{F})}^2 = \sum_{i=1}^{\infty} (v_w(p_i) - c(p_i))^2. \tag{2}$$

We call (2) the *generalized Brier score*.

The name “generalized legitimate inaccuracy measure” is motivated by the observation that a generalized legitimate inaccuracy measure naturally restricted to the finite opinion sets is a legitimate inaccuracy measure. This is because 1) for both the generalized and finite legitimate inaccuracy measures, the score of an individual proposition is defined by a one-dimensional Bregman divergence, and 2) for both the generalized and finite legitimate inaccuracy measures, the scores of individual propositions are combined in a weighted, additive way to give a score for the entire credence function. To use the terminology of Leitgeb and Pettigrew (2010a), in the finite and countably infinite setting, the local scores are the same, and the global scores relate to the local scores in the same way. These observations support the view that insofar as quasi-additive Bregman divergences are the appropriate functions to

<sup>9</sup> Boundedness is assumed for technical reasons.

<sup>10</sup> Recall that  $\sup_i a_i = a \in \mathbb{R} \cup \{+\infty, -\infty\}$  such that  $a_i \leq a$  for all  $i \in \mathbb{N}$ , and for any  $b < a$ , there is some  $a_i$  such that  $b < a_i \leq a$ .

<sup>11</sup> **Proof.** Let  $\bar{\varphi}(x) = \varphi(x) + ax + b$ . Then  $\mathfrak{d}_{\bar{\varphi}}(x, y) = \varphi(x + ax + b - \varphi(y) - ay - b - (\varphi'(y) + a)(x - y)) = \varphi(x) + ax + b - \varphi(y) - ay - b - \varphi'(y)(x - y) - ax + ay = \varphi(x) - \varphi(y) - \varphi'(y)(x - y) = \mathfrak{d}_{\varphi}(x, y)$ . Further, if  $\varphi$  satisfies the conditions in Definition 2.5.3, then  $\bar{\varphi}$  does as well.

<sup>12</sup> The choice of enumeration does not matter because the terms in the infinite sum defining inaccuracy are nonnegative. Thus, convergence is absolute and independent of order.

<sup>13</sup> Note that in the infinite setting, we may have  $v_{w_1} = v_{w_2}$  with  $w_1 \neq w_2$ .

use for measuring inaccuracy in the finite setting, generalized quasi-additive Bregman divergences are the appropriate functions to use for measuring inaccuracy in the countably infinite setting (see Section 6 for further discussion of generalized legitimate inaccuracy measures).<sup>14</sup>

### 3.2 Coherence is necessary

I now state one of the main results: coherence is necessary to avoid accuracy dominance in the countably infinite case.

**Theorem 3.4.** Let  $\mathcal{F}$  be a countably infinite opinion set,  $\mathcal{J}$  a generalized legitimate inaccuracy measure, and  $c$  an incoherent credence function. Then:

1.  $c$  is weakly dominated relative to  $\mathcal{J}$  by a coherent credence function; and
2. if  $\mathcal{J}(c, w) < \infty$  for each  $w \in W$ , then  $c$  is strongly dominated relative to  $\mathcal{J}$  by a coherent credence function.

**Remark 3.5.** By analyzing the proof of Theorem 3.4, one can see that the most general way to state the theorem is as follows: assume  $c$  is incoherent; if  $\mathcal{J}(c, w) < \infty$  for some  $w$ , then there is a coherent credence function  $d$  such that  $\mathcal{J}(d, w) < \mathcal{J}(c, w)$  for all  $w$  such that  $\mathcal{J}(c, w) < \infty$ ; if  $\mathcal{J}(c, w) = \infty$  for all  $w \in W$ , then any omniscient credence function weakly dominates  $c$ .

**Remark 3.6.** The following is easy to prove from the results of Schervish et al. (2009): any incoherent credence function  $c$  over a countably infinite opinion set is weakly dominated but not necessarily by a *coherent* credence function; and if  $\mathcal{J}(c, w) < \infty$  for each  $w \in W$ , then  $c$  is strongly dominated but not necessarily by a *coherent* credence function.<sup>15</sup> Thus, the value in the proof strategy to come is that the dominating credence function is proven to be coherent, which is analogous to the finite case.<sup>16,17</sup>

---

<sup>14</sup> The class of generalized legitimate inaccuracy measures could also be justified by defending the following axiom, which picks out the generalized legitimate inaccuracy measures: let  $\mathcal{F}$  be a countable opinion set, and let  $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$  be a sequence of increasing subsets of  $\mathcal{F}$  whose union is  $\mathcal{F}$ . Then there are legitimate inaccuracy measures  $\mathcal{J}_1, \mathcal{J}_2, \dots$  generated by the same one-dimensional Bregman divergence and compatible weights such that the inaccuracy of a credence function  $c$  on  $\mathcal{F}$  at world  $w$  is given by  $\mathcal{J}(c, w) = \lim_{n \rightarrow \infty} \mathcal{J}_n(c|_{\mathcal{F}_n}, w)$ , where  $c|_{\mathcal{F}_n}$  is the restriction of  $c$  to  $\mathcal{F}_n$ . “Compatible weights” means that if  $\mathcal{J}_n$  and  $\mathcal{J}_{n+1}$  are generated by weights  $\{a_i\}_{i=1}^{|\mathcal{F}_n|}$  and  $\{b_i\}_{i=1}^{|\mathcal{F}_{n+1}|}$ , respectively, then  $a_i = b_i$  for  $1 \leq i \leq |\mathcal{F}_n|$ .

<sup>15</sup> **Proof sketch.** If  $c$  is incoherent, then there is some finite  $\overline{\mathcal{F}} \subseteq \mathcal{F}$  on which  $c$  is incoherent. Restrict  $c$  to  $c|_{\overline{\mathcal{F}}}$  on  $\overline{\mathcal{F}}$ . Then, by Theorem 2.9, there is some  $\overline{d}$  that strongly dominates  $c|_{\overline{\mathcal{F}}}$ . Extend  $\overline{d}$  to a credence function  $d$  on  $\mathcal{F}$  by copying  $c$  off of  $\overline{\mathcal{F}}$ . Then, so long as  $c$  has finite inaccuracy at some world,  $d$  will weakly dominate  $c$ .

<sup>16</sup> Thanks to Teddy Seidenfeld for suggesting this connection to the finite case.

<sup>17</sup> Further, it is often argued that not all dominated credence functions are irrational—only those that are dominated by a credence function that is itself not dominated (see discussion of various dominance principles in Pettigrew 2016, 22). For the opinion sets and inaccuracy measures discussed in Section 4, the undominated credence functions will be precisely the coherent credence functions, and so the added strength of Theorem 3.4 is normatively important, as well.

Note that one direction of Walsh’s (ms) accuracy dominance result follows immediately from Theorem 3.4. We first recall his result.

**Theorem 3.7.** (Walsh ms) Let  $\mathcal{F}$  be a countably infinite opinion set. Let

$$\mathcal{F}(c, w) = \sum_{i=1}^{\infty} 2^{-i}(v_w(p_i) - c(p_i))^2. \tag{3}$$

Then:

1. if  $c$  is incoherent, then  $c$  is strongly dominated relative to  $\mathcal{F}$  by a coherent credence function; and
2. if  $c$  is coherent, then  $c$  is not weakly dominated relative to  $\mathcal{F}$  by any credence function  $d \neq c$ .

Part 1 of this result follows from Theorem 3.4 by defining  $\mathcal{F}$  in terms of the generalized quasi-additive Bregman divergence generated by  $\{2^{-i}\}_{i=1}^{\infty}$  and

$$\mathfrak{d}(x, y) = x^2 - y^2 - 2y(x - y) = \varphi(x) - \varphi(y) - \varphi'(y)(x - y),$$

where  $\varphi(x) = x^2$ . Note that  $\mathcal{F}(c, w) < \infty$  for all  $c \in \mathcal{C}$  and  $w \in W$  as  $\sum_{i=1}^{\infty} 2^{-i} < \infty$ .

#### 4. The countably infinite case: The sufficiency of coherence

Unlike coherent credence functions on finite opinion sets, coherent credence functions on countably infinite opinion sets can be strongly dominated.

**Example 4.1.** Let  $\mathcal{F} = \{\{n \geq N : n \in \mathbb{N}\} : N \in \mathbb{N}\}$  be an opinion set over  $\mathbb{N}$  (including zero). Let

$$c(\{n \geq N\}) = \frac{1}{\sqrt{N + 1}}.$$

Then  $c$  is coherent—in fact, countably coherent (see Definition 4.6)—but  $\mathcal{F}(c, w) = \infty$  for all  $w \in W$  when  $\mathcal{F}$  is the generalized Brier score. Thus, any omniscient credence function strongly dominates  $c$ .

In fact, the classic example of a merely finitely additive probability function—the 0-1 function defined on the finite-cofinite algebra over  $\mathbb{N}$  taking value zero on finite sets—restricts to a coherent dominated credence function.

**Example 4.2.** Let  $\mathcal{F} = \{\{n \leq N : n \in \mathbb{N}\} : N \in \mathbb{N}\}$  be an opinion set over  $\mathbb{N}$  (including zero). Let

$$c(\{n \leq N\}) = 0.$$

Then  $c$  is coherent—as well as finitely supported and not countably coherent—but  $\mathcal{F}(c, w) = \infty$  for all  $w \in W$  when  $\mathcal{F}$  is the generalized Brier score. Thus, any omniscient credence function strongly dominates  $c$ .

The goal of this section is to characterize the opinion sets and inaccuracy measures for which some variant of Theorem 2.9 holds. We can extend Theorem 2.9 by proving dominance results for *countably coherent* credence functions and using an opinion set compactification construction to transfer these results to merely coherent credence functions. At points, the results will only apply to the generalized Brier score. I conjecture that any such result extends to any generalized legitimate inaccuracy measure. In any case, this is a well-motivated restriction because the Brier score has been defended by many—including Horwich (1982), Maher (2002), Joyce (2009), and Leitgeb and Pettigrew (2010a)—as being a particularly appropriate way to measure inaccuracy.

A summary of the main results from Sections 3-4 can be found in Figure 4.1.

*Throughout the rest of Section 4, we assume that the opinion set  $\mathcal{F}$  is countably infinite, unless otherwise stated.*

### 4.1 Countable coherence

I begin by introducing the notion of a *countably coherent* credence function and establishing a characterization theorem regarding countable coherence on *countably discriminating* opinion sets that extends a result of de Finetti (1974).

**Definition 4.3.** For an opinion set  $\mathcal{F} \subseteq \mathcal{P}(W)$ , we define an equivalence relation  $\sim$  on  $W$  such that  $w \sim w'$  if and only if  $\{p \in \mathcal{F} : w \in p\} = \{p \in \mathcal{F} : w' \in p\}$ . We call the set of equivalence classes of  $W$  the *quotient of  $W$  relative to  $\mathcal{F}$* . If the quotient of  $W$  relative to  $\mathcal{F}$  is countable, then we call  $\mathcal{F}$  *countably discriminating*.

Clearly, any countable opinion set over a countable set of worlds is countably discriminating.

The following characterization of the coherent credence functions on finite opinion sets is due to de Finetti (1974). Recall that  $\mathcal{V}_{\mathcal{F}}$  is the set of omniscient credence functions on  $\mathcal{F}$ , which is finite when  $\mathcal{F}$  is finite.

**Theorem 4.4.** (de Finetti 1974)  $c$  is a coherent credence function on a finite opinion set  $\mathcal{F}$  if and only if there are  $\lambda_w \in [0, 1]$  with  $\sum_{v_w \in \mathcal{V}_{\mathcal{F}}} \lambda_w = 1$  such that

$$c(p) = \sum_{v_w \in \mathcal{V}_{\mathcal{F}}} \lambda_w v_w(p)$$

for all  $p \in \mathcal{F}$ .

Theorem 4.4 is integral to Predd et al.’s (2009) proof that coherence is sufficient to avoid dominance in Theorem 2.9. I now show de Finetti’s (1974) characterization of the coherent credence functions on finite opinion sets extends to countably coherent credence functions on countably infinite opinion sets that are countably discriminating.

**Definition 4.5.** A  $\sigma$ -algebra over  $W$  is a subset  $\mathcal{F}^* \subseteq \mathcal{P}(W)$  such that:

1.  $W \in \mathcal{F}^*$ ;
2. if  $\{p_i\}_{i=1}^{\infty} \subseteq \mathcal{F}^*$ , then  $\bigcup_{i=1}^{\infty} p_i \in \mathcal{F}^*$ ; and
3. if  $p \in \mathcal{F}^*$ , then  $W \setminus p \in \mathcal{F}^*$ .

**Definition 4.6.** Let a credence function  $c$  be *countably coherent* if  $c$  extends to a countably additive probability function on a  $\sigma$ -algebra  $\mathcal{F}^*$  containing  $\mathcal{F}$ .<sup>18</sup> That is, there is a  $c^* : \mathcal{F}^* \rightarrow [0, 1]$  such that:

1.  $c^*(p) = c(p)$  for all  $p \in \mathcal{F}$ ;
2.  $c^*(\bigcup_{i=1}^{\infty} p_i) = \sum_{i=1}^{\infty} c^*(p_i)$  for  $\{p_i\}_{i=1}^{\infty} \subseteq \mathcal{F}^*$  with  $p_i \cap p_j = \emptyset$  for  $i \neq j$ ; and
3.  $c^*(W) = 1$ .

Otherwise, a credence function is *countably incoherent*.

**Proposition 4.7.** Let  $\mathcal{F}$  be a countably infinite opinion set that is countably discriminating (thus,  $V_{\mathcal{F}}$  is countable). Then a credence function  $c$  is countably coherent if and only if there are  $\lambda_{v_w} \in [0, 1]$  with  $\sum_{v_w \in V_{\mathcal{F}}} \lambda_{v_w} = 1$  such that

$$c(p) = \sum_{v_w \in V_{\mathcal{F}}} \lambda_{v_w} v_w(p)$$

for all  $p \in \mathcal{F}$ .

### 4.2 Compactification of an opinion space

In this section, I introduce the compactification construction of what is called an *opinion space*. The construction will be relevant to transferring dominance results for countably coherent credence functions to merely coherent credence functions, the reason being that merely coherent credence functions become countably coherent if the underlying set of worlds is “compactified.”

**Definition 4.8.** An *opinion space* is a pair  $(W, \mathcal{F})$ , where  $W$  is a nonempty set and  $\mathcal{F} \subseteq \mathcal{P}(W)$ .

From here on out, I will speak in terms of opinion spaces as opposed to opinion sets in order to keep track of the underlying set of worlds. We continue to assume that  $\mathcal{F}$  is countably infinite.

Borkar et al. (2003) proved that the opinion spaces that satisfy a certain compactness property are precisely those where the set of coherent credence functions and the set of countably coherent credence functions coincide.<sup>19</sup>

**Definition 4.9.** Let  $(W, \mathcal{F})$  be an opinion space. Let  $f(n) \in \{0, 1\}$  and set  $p_n^{f(n)} = p_n$  if  $f(n) = 0$  and  $p_n^{f(n)} = p_n^c$  if  $f(n) = 1$ . Then  $(W, \mathcal{F})$  is *compact* if for any choice of  $\{p_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$  and  $f : \mathbb{N} \rightarrow \{0, 1\}$ , if  $\bigcap_{n=1}^N p_n^{f(n)}$  is nonempty for every  $N$ , then  $\bigcap_{n=1}^{\infty} p_n^{f(n)}$  is nonempty.

<sup>18</sup> Note that if  $c$  is countably coherent on  $\mathcal{F}$ , then  $c$  extends to a countably additive probability function on  $\sigma(\mathcal{F})$ , the  $\sigma$ -algebra generated by  $\mathcal{F}$ .

<sup>19</sup> Borkar et al., (2003) do not restrict attention to opinion spaces where  $\mathcal{F}$  is countably infinite, and it is easy to see that there are analogues of the following compactification results in the uncountable setting. However, we continue to restrict our attention to the countably infinite setting because we have not yet extended the accuracy framework beyond that.

As an example, note that the opinion spaces from Examples 4.1 and 4.2 are not compact. Indeed, for the first example,  $\bigcap_{n=1}^{\infty} p_n = \emptyset$ , and yet every finite subset of  $\mathcal{F}$  has a nonempty intersection; for the second example,  $\bigcap_{n=1}^{\infty} p_n^c = \emptyset$ , whereas  $\bigcap_{n=1}^N p_n^c \neq \emptyset$  for every  $N$ .

**Theorem 4.10.** (Borkar et al. 2003) The following are equivalent:

1.  $(W, \mathcal{F})$  is compact;
2. for every credence function  $c$  on  $(W, \mathcal{F})$ ,  $c$  is coherent if and only if  $c$  is countably coherent.

I now show how to turn any space into a compact space and, in light of Theorem 4.10, any coherent credence function into a countably coherent credence function. Let  $(W, \mathcal{F})$  be an opinion space. Let  $S$  denote the set of sequences of the form  $\{p_n^{f(n)}\}$  (as in Definition 4.9) such that  $\bigcap_{n=1}^N p_n^{f(n)} \neq \emptyset$  for every  $N$  but  $\bigcap_{n=1}^{\infty} p_n^{f(n)} = \emptyset$ . Define  $W^* = W \cup \{x_s : s \in S\}$ , where each  $x_s$  is a formal point corresponding to the element  $s \in S$ . Define  $\mathcal{F}^* \subseteq \mathcal{P}(W^*)$  as follows: for each  $p \in \mathcal{F}$ , let  $S_p$  denote the set of sequences  $s$  of the form  $\{p_n^{f(n)}\}$  (as in Definition 4.9) such that  $s \in S$ ,  $p_n = p$  for some  $n$ , and  $f(n) = 0$ . Then define

$$p^* = p \cup \{x_s : s \in S_p\}.$$

Finally, let  $\mathcal{F}^* = \{p^* : p \in \mathcal{F}\}$ . We call  $(W^*, \mathcal{F}^*)$  the compactification of  $(W, \mathcal{F})$ . We always denote the compactification of  $(W, \mathcal{F})$  by  $(W^*, \mathcal{F}^*)$ . Further, we let  $\Psi$  denote the natural bijection from  $\mathcal{F}$  to  $\mathcal{F}^*$  given by  $\Psi(p) = p^*$ .

We first note that  $(W^*, \mathcal{F}^*)$  is in fact compact.

**Lemma 4.11.** For  $(W, \mathcal{F})$  an opinion space,  $(W^*, \mathcal{F}^*)$  is compact.

Next, we note that, as suggested, we can naturally turn a coherent credence function into a countably coherent credence function by compactifying the underlying opinion space.

**Lemma 4.12.** Let  $(W, \mathcal{F})$  be an opinion space and  $c$  be a coherent credence function on  $(W, \mathcal{F})$ . Let  $(W^*, \mathcal{F}^*)$  be the compactification of  $(W, \mathcal{F})$ , and define  $c^*(\Psi(p)) := c(p)$  for each  $p \in \mathcal{F}$ . Then  $c^*$  is a countably coherent credence function on  $(W^*, \mathcal{F}^*)$ , and  $\mathcal{I}(c, w) = \mathcal{I}(c^*, w)$  for  $w \in W$ .

For a coherent credence function  $c$  defined on an opinion space  $(W, \mathcal{F})$ , we let  $c^*$  denote the countably coherent credence function on  $(W^*, \mathcal{F}^*)$  given as in Lemma 4.12.

**Example 4.13.** As an example, let us compute the compactification of the opinion space from Example 4.2 and show how to identify a coherent credence function on the space with a countably coherent credence function on its compactification. We note that only for  $f(n) = 1$  for all  $n \in \mathbb{N}$  is  $\bigcap_{n=1}^N p_n^{f(n)}$  nonempty for every  $N$  while  $\bigcap_{n=1}^{\infty} p_n^{f(n)} = \emptyset$ . Indeed, assume  $f(m) = 0$  for some  $m$ . If  $f(i) = 1$  for some  $i \geq m + 1$ ,

then because  $p_i^c \cap p_m = \emptyset$ , we have  $\bigcap_{n=1}^i p_n^{f(n)} = \emptyset$  which contradicts our assumption. So  $f(i) = 0$  for all  $i \geq m + 1$ . But then because  $\bigcap_{n=1}^m p_n^{f(n)} \neq \emptyset$  and  $\bigcap_{n=1}^m p_n^{f(n)} \subseteq p_i$  for all  $i \geq m + 1$ , it also follows that  $\bigcap_{n=1}^\infty p_n^{f(n)} \neq \emptyset$ , which contradicts our assumption. So  $S$  is a single point  $x$ ,  $W^* = W \cup \{w^*\}$ , and  $\mathcal{F}^* = \{\{n \leq N\} : N \in \mathbb{N}\}$ .  $\mathcal{F}^*$  is identical to  $\mathcal{F}$ , except that there is a point in the complement of every proposition in  $\mathcal{F}^*$ . For a coherent credence function  $c$  on  $(W, \mathcal{F})$ ,  $c^*$  on  $(W^*, \mathcal{F}^*)$  is identical to  $c$  and is a countably coherent credence function on the compact opinion space  $(W^*, \mathcal{F}^*)$ . For example, for credence function  $c$  in Example 4.2,  $c^*$  extends to the countably additive omniscient credence function  $v_{w^*}$  on the  $\sigma$ -algebra generated by  $\mathcal{F}^*$ .

Using Theorem 4.10, Lemma 4.11, and Lemma 4.12, the proof strategy for extending Theorem 2.9 is more precisely as follows. First, we establish dominance results for countably coherent credence functions. Second, we transform each coherent credence function on  $(W, \mathcal{F})$  into a countably coherent credence function on  $(W^*, \mathcal{F}^*)$ , as in Lemma 4.12. Lastly, we use the dominance results for countably coherent credence functions to establish dominance results for coherent credence functions in certain cases where there is “accuracy dominance stability” in compactifying.

### 4.3 W-stable opinion spaces

In this section, we establish the equivalence between coherence and avoiding weak dominance for certain countably infinite opinion spaces (Theorem 4.19), as well as additional results extending Theorem 2.9 to the countably infinite setting (corollary 4.22 and Theorem 4.23). We first note that under certain circumstances, countably coherent credence functions are not weakly dominated (Proposition 4.15 and Proposition 4.16); then I use the compactification construction from the previous section and a property of an opinion space—*W-stability* (Definition 4.17)—to establish that for certain opinion spaces, mere coherence is also sufficient to avoid weak dominance.

I first prove that if a countably coherent credence function  $c$  has finite expected inaccuracy, then  $c$  is not weakly dominated.

**Definition 4.14.** For  $c$ , a countably coherent credence function, and  $\mathcal{I}$ , a generalized legitimate inaccuracy measure, we say that  $c$  has *finite expected inaccuracy relative to  $\mathcal{I}$*  if  $c$  has a countably additive extension  $\bar{c}$  defined on the opinion space  $(W, \sigma(\mathcal{F}))$  such that  $\mathbb{E}_{\bar{c}} \mathcal{I}(c, \cdot) < \infty$ .<sup>20</sup> For  $c$ , a coherent but not countably coherent credence function, and  $\mathcal{I}$ , a generalized legitimate inaccuracy measure, we say that  $c$  has *finite expected inaccuracy relative to  $\mathcal{I}$*  if  $c^*$  has finite expected inaccuracy relative to  $\mathcal{I}$ .

**Proposition 4.15.** Let  $(W, \mathcal{F})$  be an opinion space and  $\mathcal{I}$  be a generalized legitimate inaccuracy measure. If  $c$  is a countably coherent credence function on  $(W, \mathcal{F})$  with finite expected inaccuracy relative to  $\mathcal{I}$ , then  $c$  is not weakly dominated relative to  $\mathcal{I}$ .

<sup>20</sup> Consider the measure space  $(W, \sigma(\mathcal{F}), \mu)$ . Note that  $\mathfrak{d}(v_w(p_i), d_i) = 1_{p_i}(w)\mathfrak{d}(1, d_i) + (1 - 1_{p_i}(w))\mathfrak{d}(0, d_i)$ , so each term in  $\mathcal{I}(d, \cdot)$  is measurable for any credence function  $d$ , and thus the infinite sum is measurable because the finite sum and limit of measurable functions are measurable. Thus, we can take the expectation of  $\mathcal{I}(d, \cdot)$  with respect to  $\mu$  for any credence function  $d$ .

Here is another dominance result for countably coherent credence functions where we assume  $\mathcal{F}$  is *point-finite* ( $|\{p \in \mathcal{F} : w \in p\}| < \infty$  for all  $w \in W$ ) but weaken the assumption that  $c$  has finite expected inaccuracy considerably, namely, to *somewhere finitely inaccurate* (there is a  $w \in W$  such that  $\mathcal{J}(c, w) < \infty$ ). We also restrict to the generalized Brier score  $\mathcal{B}$ .

**Proposition 4.16.** Let  $(W, \mathcal{F})$  be a point-finite opinion space. If a credence function  $c$  is countably coherent and somewhere finitely inaccurate relative to  $\mathcal{B}$ , then  $c$  is not weakly dominated relative to  $\mathcal{B}$ .

I now introduce the notion of *W-stability*, which will allow the use of Propositions 4.15 and 4.16 to prove extensions of Theorem 2.9. Intuitively, an opinion space is *W-stable* if coherence is preserved under compactification.

**Definition 4.17.** Let  $(W, \mathcal{F})$  be *W-stable relative to  $\mathcal{J}$*  if for any coherent credence function  $c$  on  $(W, \mathcal{F})$ , if  $c$  is weakly dominated relative to  $\mathcal{J}$ , then  $c^*$  on  $(W^*, \mathcal{F}^*)$  is weakly dominated relative to  $\mathcal{J}$ .

**Remark 4.18.** Not all opinion spaces are *W-stable relative to every generalized legitimate inaccuracy measure*. Indeed, consider the opinion space from Example 4.2. The credence function  $c$  in that example, which assigns 0 to each proposition, is strongly dominated. However, where  $w^*$  is the single world added to the complement of each proposition when compactifying (see Example 4.13) and  $v_{w^*} : \mathcal{F}^* \rightarrow [0, 1]$  is the omniscient credence function at world  $w^*$ , we have that  $c^* = v_{w^*}$ , and so  $\mathcal{B}(c^*, w^*) = 0$ . Further, for any credence function  $d \neq c^*$ , we have  $\mathcal{B}(d, w^*) > 0$  because  $\mathcal{B}(d, w^*) = 0$  if and only if  $d = v_{w^*} = c^*$ . Thus, once the underlying space is compactified,  $c$  is no longer weakly dominated relative to  $\mathcal{B}$ .

Using Proposition 4.15, we establish one of the main results: sufficient and partly necessary conditions on an opinion space for coherence to be equivalent to not being weakly dominated.

**Theorem 4.19.** Let  $\mathcal{J}$  be a generalized legitimate inaccuracy measure and  $(W, \mathcal{F})$  a *W-stable opinion space relative to  $\mathcal{J}$*  where all coherent credence functions have finite expected inaccuracy relative to  $\mathcal{J}$ . Then the following are equivalent:

1.  $c$  is coherent;
2.  $c$  is not weakly dominated.

**Proof.** We prove that if  $c$  is coherent, then  $c$  is not weakly dominated. Let  $(W^*, \mathcal{F}^*)$  be the compactification of  $(W, \mathcal{F})$ . If  $c$  is coherent on  $(W, \mathcal{F})$ , then  $c^*$  is countably coherent by Lemma 4.12. Further,  $c^*$  has finite expected inaccuracy by definition and the assumption that  $c$  has finite expected inaccuracy. Thus, by Proposition 4.15,  $c^*$  is not weakly dominated. But since  $(W, \mathcal{F})$  is *W-stable*, this implies that  $c$  is not weakly dominated. The other direction follows from Theorem 3.4. □

**Remark 4.20.** It is trivial to see that W-stability is necessary for the equivalence of coherence and not being weakly dominated. How far finite expected inaccuracy can be weakened is an open question.

**Remark 4.21.** If  $\mathcal{F}$  is defined with summable weights, that is,  $\{a_i\}_{i=1}^{\infty}$  such that  $\sum_{i=1}^{\infty} a_i < \infty$ , then there is a  $C < \infty$  such that  $\mathcal{F}(c, w) < C$  for all credence functions  $c$  and  $w \in W$ . So, in particular, all coherent credence functions have finite expected inaccuracy relative to  $\mathcal{F}$ .

If we add in an additional finiteness assumption, then we get the full equivalence of Theorem 2.9.

**Corollary 4.22.** In Theorem 4.19, if in addition, all coherent credence functions  $c$  have  $\mathcal{F}(c, w) < \infty$  for all  $w \in W$ , then the following are equivalent:

1.  $c$  is coherent;
2.  $c$  is not weakly dominated;
3.  $c$  is not strongly dominated.

We combine W-stability and Proposition 4.16 to get another set of sufficient conditions on  $(W, \mathcal{F})$  for Theorem 2.9 to go through for the generalized Brier score.

**Theorem 4.23.** Let  $(W, \mathcal{F})$  be a W-stable opinion space with  $(W^*, \mathcal{F}^*)$  point-finite such that all coherent credence functions on  $(W, \mathcal{F})$  are somewhere finitely inaccurate relative to  $\mathcal{B}$ . Then the following are equivalent:

1.  $c$  is coherent;
2.  $c$  is not weakly dominated relative to  $\mathcal{B}$ ;
3.  $c$  is not strongly dominated relative to  $\mathcal{B}$ .

**Remark 4.24.** We can drop the assumption that all coherent credence functions are somewhere finitely inaccurate in Theorem 4.23 if we strengthen W-stable to compact so that  $(W, \mathcal{F}) = (W^*, \mathcal{F}^*)$ . Indeed, compactness alongside point-finiteness implies that coherent credence functions on  $(W, \mathcal{F}) = (W^*, \mathcal{F}^*)$  are somewhere finitely inaccurate: if there were a coherent (and thus countably coherent) credence function infinitely inaccurate at all worlds, then it would be strongly dominated by an omniscient credence function, contradicting Proposition 4.27 later in this section.

#### 4.3.1 Partitions

As an application of Theorem 4.19, I establish Theorem 2.9 for countably infinite partitions.<sup>21</sup> In parts of the existing literature (e.g., Joyce 2009), credence functions are assumed to be defined on a (finite) partition of  $W$  to begin with, and so such a result

<sup>21</sup> It has been noted that de Finetti's (1974) original proof of Theorem 2.9 assumes the Brier score extends to countably infinite opinion sets. However, the only proof I have seen is a sketch of the necessity of coherence for countably infinite partitions by Joyce (1998, 6n).

might be especially relevant to extending the accuracy dominance argument for probabilism to countably infinite opinion sets.

**Lemma 4.25.** A countably infinite partition is  $W$ -stable relative to any generalized legitimate inaccuracy measure.

**Theorem 4.26.** Let  $(W, \mathcal{F})$  be a countably infinite partition and  $\mathcal{I}$  be a generalized legitimate inaccuracy measure. Then the following are equivalent:

1.  $c$  is coherent;
2.  $c$  is not weakly dominated;
3.  $c$  is not strongly dominated.

#### 4.4 $S$ -stable opinion spaces

In this section, I establish the equivalence between coherence and avoiding strong dominance for certain countably infinite opinion spaces (Theorem 4.29). The conditions are in terms of the analogous stability condition— $S$ -stability (Definition 4.28)—but a different finiteness assumption, and the proof strategy is the same as for Theorem 4.19.

I begin by establishing that on compact countably infinite opinion spaces, coherent and thus countably coherent credence functions (recall Theorem 4.10) are not strongly dominated.

**Proposition 4.27.** Let  $(W, \mathcal{F})$  be a compact opinion space and  $\mathcal{I}$  be a generalized legitimate inaccuracy measure. If  $c$  is coherent (and thus countably coherent), then  $c$  is not strongly dominated relative to  $\mathcal{I}$ .

I now introduce  $S$ -stability and the main theorem of this section.

**Definition 4.28.** Let  $(W, \mathcal{F})$  be  $S$ -stable relative to  $\mathcal{I}$  if for any coherent credence function  $c$  on  $(W, \mathcal{F})$ , if  $c$  is strongly dominated relative to  $\mathcal{I}$ , then  $c^*$  on  $(W^*, \mathcal{F}^*)$  is strongly dominated relative to  $\mathcal{I}$ .

**Theorem 4.29.** Let  $\mathcal{I}$  be a generalized legitimate inaccuracy measure and  $(W, \mathcal{F})$  be an  $S$ -stable opinion space relative to  $\mathcal{I}$ . Assume that  $\mathcal{I}(c, w) < \infty$  for each coherent credence function  $c$  and  $w \in W$ . Then the following are equivalent:

1.  $c$  is coherent;
2.  $c$  is not strongly dominated.

**Proof.** Assume  $c$  is coherent.  $c^*$  defined on the compact opinion space  $(W^*, \mathcal{F}^*)$  is countably coherent by Lemma 4.12. Thus, by Proposition 4.27,  $c^*$  is not strongly dominated. But since  $\mathcal{F}$  is  $S$ -stable, this implies that  $c$  is not strongly dominated relative to  $\mathcal{I}$ . The other direction follows from Theorem 3.4.  $\square$

Opinion Space	Inaccuracy Measure	Conclusion
—	—	incoherent $\implies$ weakly dominated (by coherent credence function)
—	$\mathcal{I}(c, w) < \infty$ for all $w$	incoherent $\implies$ strongly dominated (by coherent credence function)
W-stable	finite expected inaccuracy	coherent $\iff$ not weakly dominated
W-stable	finite expected inaccuracy + $\mathcal{I}(c, w) < \infty$ for all $w$	coherent $\iff$ not weakly dominated $\iff$ not strongly dominated
W-stable + ( $W^*, \mathcal{F}^*$ ) point-finite	$\mathcal{I} = \mathcal{B}$ + somewhere finitely inaccurate	coherent $\iff$ not weakly dominated $\iff$ not strongly dominated
partition	—	coherent $\iff$ not weakly dominated $\iff$ not strongly dominated
S-stable	$\mathcal{I}(c, w) < \infty$ for all $w$	coherent $\iff$ not strongly dominated

**Figure 4.1.** Summary of the results presented in Sections 3 and 4. Each column summarizes a main result by specifying conditions under which a particular conclusion about a credence function  $c$  holds. The column titled “Opinion Space” specifies conditions on the opinion space on which  $c$  is defined, the column titled “Inaccuracy Measure” specifies conditions on the generalized legitimate inaccuracy measure relative to which dominance is defined, and the column titled “Conclusion” specifies what can be deduced regarding the relationship between coherence and dominance for  $c$  under the given conditions. “—” in a box means no additional condition is imposed.

**Remark 4.30.** It is trivial to see that S-stability is necessary for the equivalence of coherence and avoiding strong dominance. How much the assumption that coherent credence functions satisfy  $\mathcal{I}(c, w) < \infty$  for all  $w$  can be weakened is an open question.

**Remark 4.31.** Schervish et al. (2009) take a different approach to dropping the assumption that the opinion set is finite: they apply weak and strong dominance notions to finite subsets of opinion sets of arbitrary cardinality. They also explore connections between the two notions of dominance considered here—weak and strong dominance—and what they call *coherence*<sub>1</sub>, which amounts to avoiding being susceptible to a finite *Dutch book*.<sup>22</sup>

**Remark 4.32.** Theorem 4.29 is related to Theorem 1 of Schervish et al. (2014). However, 1) their assumptions are in some ways weaker and in some ways stronger

<sup>22</sup> Thanks to Teddy Seidenfeld for pointing me to this work of Schervish et al. (2009). An additional point worth noting about their work is that they further generalize the finite results of Predd et al. (2009) by i) allowing a wider variety of inaccuracy measures, including those that are merely proper as opposed to strictly proper, and ii) by scoring conditional probabilities. A natural direction for future work is to use these relaxations in the finite case to relax the assumptions made here. Similarly, Steeger (2019) considers the property of avoiding strong dominance with respect to the Brier score for every finite subset of opinion sets of arbitrary cardinality (see “sufficient coherence” on p. 38 of Steeger 2019).

than those in Theorem 4.29,<sup>23</sup> and 2) although Schervish et al. (2014) establish that coherence is sufficient for avoiding strong dominance in certain cases, unlike Theorems 4.19 and 4.29, their results do not show that coherence is sufficient for avoiding even weak dominance in certain cases or that incoherence always entails being weakly dominated (and sometimes strongly dominated) by a coherent credence function (see Remark 3.6).

For a summary of the results established thus far, see Figure 4.1.

#### 4.5 Further directions

Although Theorems 4.19 and 4.29 come close to characterizing the countably infinite opinion spaces on which not being weakly and strongly dominated, respectively, are equivalent to coherence, how far the finiteness assumptions in the theorems can be weakened is an open question. This is a natural next line of inquiry. In addition, it would be useful to determine characterizations of W- and S-stability in terms of the inaccuracy measure that would make it relatively easy to check whether an opinion set is W- or S-stable. For example, might it be that an opinion space is W- and S-stable relative to a generalized legitimate inaccuracy measure if the generalized legitimate inaccuracy measure only outputs finite scores for credence functions on that opinion space? Also, there are natural ways to generalize the previous results to more closely match the finite results: allow different one-dimensional Bregman divergences for different propositions and allow unbounded one-dimensional Bregman divergences.

Another direction one could go in exploring the sufficiency of coherence for avoiding dominance is as follows: instead of characterizing the countably infinite opinion sets on which Theorem 2.9 goes through, one could characterize the kinds of coherent credence functions for which Theorem 2.9 goes through on any countably infinite opinion set.<sup>24</sup> Doing so might show that although coherence is not enough to avoid dominance in all cases, coherence along with additional plausible constraints is sufficient. In particular, although restricting to finitely supported credence functions is not enough to establish the sufficiency of coherence for avoiding strong dominance (as shown in Example 4.2), whether *countable* coherence is equivalent to avoiding weak or strong dominance on the restricted class is an open question.

### 5. The uncountable case

So far, we have been concerned with credences defined on countable opinion sets. We now consider what can be said in favor of probabilism when credences are defined on uncountable opinion sets, though much of what is said will be preliminary. When extending from the finite to the countably infinite setting, we used inaccuracy measures that naturally restrict to legitimate inaccuracy measures in the finite case. When extending from the countable to the uncountable setting, we will use inaccuracy measures defined as integration against a measure on the uncountable set of

<sup>23</sup> Schervish et al. (2014) require that the *prevision* for the inaccuracy of the credence function be finite and that inaccuracy be pointwise finite, whereas I only assume the latter. On the other hand, I require the opinion set to be S-stable, whereas they do not.

<sup>24</sup> Thanks to Thomas Icard and Milan Mossé for suggesting this alternative direction of study.

propositions. This is a natural generalization of (generalized) legitimate inaccuracy measures, for (generalized) legitimate inaccuracy measures are defined as integration against a particular kind of measure on a countable set of propositions. Indeed, upon inspection, one can see that in the finite and countably infinite setting, a (generalized) legitimate inaccuracy measure defined in terms of weights  $\{a_i\}$  and one-dimensional Bregman divergence  $\mathfrak{d}$  is given by integrating  $\mathfrak{d}(v_w(\cdot), c(\cdot))$  as a function of  $\mathcal{F} = \{p_1, p_2, \dots\}$  against the measure  $\mu$  on  $\mathcal{F}$  defined by  $\mu(A) = \sum_{p_i \in A} a_i$  for  $A \in \mathcal{P}(\mathcal{F})$ . We generalize this construction for uncountable  $\mathcal{F}$  by defining the legitimate inaccuracy measures to be given by integration of a one-dimensional Bregman divergence  $\mathfrak{d}(v_w(\cdot), c(\cdot))$  as a function of  $\mathcal{F}$  against a measure  $\mu$  defined on  $\mathcal{F}$ .

Due to the measure-theoretic construction of the inaccuracy measures we consider, we restrict our attention to measurable credence functions and equate credence functions that are equal almost everywhere. On some measure spaces, such as the weighted counting measure spaces (with all nonzero weights) underlying (generalized) legitimate inaccuracy measures, we lose nothing since every credence function is measurable and only the empty set is measure zero. However, in other cases, these assumptions are substantive. I discuss this issue further after stating the main theorem of this section (Theorem 5.3).

We now formally extend the accuracy framework to the measure theoretic setting.

**Definition 5.1.** Let  $\mathcal{F} \subseteq \mathcal{P}(W)$  be an opinion set,  $(\mathcal{F}, \mathcal{A}, \mu)$  a measure space,<sup>25</sup> and  $c : \mathcal{F} \rightarrow \mathbb{R}^+$ . If  $c$  is  $\mathcal{A}$ -measurable and  $\mu(\{p : c(p) \notin [0, 1]\}) = 0$ , we call  $c$  a  $\mu$ -credence function. We say that a  $\mu$ -credence function  $c$  is  $\mu$ -coherent if there is a coherent (in the usual sense) credence function  $c'$  on  $\mathcal{F}$  with  $c = c'$   $\mu$ -a.e. We say a  $\mu$ -credence function  $c$  is  $\mu$ -incoherent if there is no coherent credence function  $c'$  such that  $c = c'$   $\mu$ -a.e.

**Definition 5.2.** Let  $\mathcal{F}$  be an opinion set (of arbitrary cardinality) over a set  $W$  of worlds. Let  $(\mathcal{F}, \mathcal{A}, \mu)$  be a  $\sigma$ -finite measure space over the opinion set  $\mathcal{F}$ . Let  $\mathcal{C}$  be the space of all  $\mu$ -credence functions. Assume  $\mathcal{I} : \mathcal{C} \times W \rightarrow [0, \infty]$  is such that, for all  $(c, w) \in \mathcal{C} \times W$ , we have

$$\mathcal{I}(c, w) = B_{\varphi, \mu}(v_w, c),$$

where  $B_{\varphi, \mu}$  is a Bregman distance relative to  $\varphi$ <sup>26</sup> and  $(\mathcal{F}, \mathcal{A}, \mu)$  (see Definition A.1). In particular, each  $v_w$  is a  $\mu$ -credence function. Then we call  $\mathcal{I}$  an integral inaccuracy measure on  $(\mathcal{F}, \mathcal{A}, \mu)$ .

Here is a dominance result for integral inaccuracy measures. The proof is essentially a measure theoretic version of the proof of Theorem 3.4.

**Theorem 5.3.** Let  $\mathcal{I}$  be an integral inaccuracy measure on a finite measure space  $(\mathcal{F}, \mathcal{A}, \mu)$ .<sup>27</sup> Then for every  $\mu$ -credence function  $c$ , if  $c$  is  $\mu$ -incoherent, then there is a  $\mu$ -coherent  $\mu$ -credence function  $c'$  that strongly dominates  $c$  relative to  $\mathcal{I}$ .

<sup>25</sup> So  $\mathcal{A}$  is a  $\sigma$ -algebra on  $\mathcal{F}$ , and  $\mu$  is a (countably additive) measure on  $(\mathcal{F}, \mathcal{A})$ .

<sup>26</sup> Again, we assume that the one-dimensional Bregman divergence  $\mathfrak{d}$  generated by  $\varphi$  is bounded.

<sup>27</sup> We assume finiteness for technical reasons.

It is worth noting that this result does *not* show that every incoherent credence function is strongly dominated because not every incoherent credence function is a  $\mu$ -incoherent credence function: there can be incoherent credence functions that are  $\mu$ -a.e. equivalent to coherent credence functions, even ones that are undominated insofar as any coherent credence function is undominated.<sup>28</sup> This is to be expected because, as discussed earlier, in moving to a measure-theoretic framework, measure-zero differences between credence functions will not be detected as far as accuracy is concerned, and thus incoherent credence functions that deviate from coherence by a measure-zero set will have the same inaccuracy scores as coherent credence functions. Thus, an accuracy dominance argument in this measure-theoretic setting will at most establish that one ought to have a credence function that is coherent off a measure-zero set (i.e., is  $\mu$ -coherent). Theorem 5.3 is a first step toward such an argument. I leave for future work a consideration of different ways to extend inaccuracy measures to the uncountable setting that might establish not just  $\mu$ -coherence but (strict) coherence.

Here is an example of how Theorem 5.3 can be used to give an accuracy argument in a concrete uncountable setting. Assume we have a coin with unknown bias  $\theta \in [0, 1]$  and a set of propositions of the form “ $a \leq \theta \leq b$ ” for each  $a, b \in [0, 1]$  with  $a \leq b$ . Then a credence function on this uncountable opinion set can be represented by a function

$$c : X \rightarrow [0, 1],$$

where  $X = \{(a, b) : 0 \leq a \leq b \leq 1\} \subseteq [0, 1]^2$ . We put the Lebesgue measure on  $X$  to generalize the additive constraint often assumed in the finite case. We let

$$\mathcal{F}(c, w) = \int_X \mathfrak{d}(v_w(\mathbf{x}), c(\mathbf{x}))\lambda(d\mathbf{x})$$

for a bounded one-dimensional Bregman divergence  $\mathfrak{d}$ . Then the assumptions of Theorem 5.3 hold, so we get the following dominance result: for any  $\lambda$ -credence function  $c$ , if  $c$  is  $\lambda$ -incoherent, then there is a  $\lambda$ -coherent  $\lambda$ -credence function that strongly dominates  $c$ .

## 6. Discussion

I now briefly consider the difficult question of what normative conclusions to draw from the results of Sections 3–5.

### 6.1 Extending the finite accuracy dominance argument for probabilism

To begin, it has been shown by Schervish et al. (2009, Theorem 1) that a credence function on an opinion set of arbitrary cardinality is coherent if and only if its restriction to any finite subset of the opinion set is not strongly dominated. One might wonder whether this result suffices for an accuracy-based argument for probabilism on an infinite opinion set.<sup>29</sup> Why care, in addition, about dominance relations when all infinitely many of the agent’s credences are scored at once?

<sup>28</sup> Thanks to an anonymous referee for raising this issue.

<sup>29</sup> Easwaran (2016) suggests appealing to a similar notion of “local accuracy dominance” when dealing with infinite sets of full beliefs.

We need to separate the questions of whether there is an accuracy-based argument for probabilism in the infinite setting and whether there is an *extension* to the infinite setting of the accuracy dominance arguments for probabilism already established in the finite setting. Although there are differences between the various accuracy dominance arguments for probabilism in the finite setting, they share the key feature of appealing to the following mathematical fact: relative to some legitimate way of scoring the inaccuracy of credence functions over a given opinion set, every incoherent credence function is accuracy dominated by a coherent credence function, and no coherent credence function is accuracy dominated by any other credence function.<sup>30</sup> Motivating the appeal to this mathematical fact is a commitment to the epistemic value of having an *overall* accurate epistemic state, where the epistemic state in question is represented by a credence function.

By suggesting that the epistemic value underlying the finite accuracy dominance arguments is that of overall accuracy, I do not mean to suggest that inaccuracy measures ought to be sensitive to what Pettigrew (2016) calls the “irreducibly global features” of a credence function (see 49–50). To suggest so would likely conflict with allowing an inaccuracy measure to be additive, which I do here. Indeed, Pettigrew motivates additivity by suggesting that inaccuracy need *not* be sensitive to global features because credence functions are not unified doxastic states but simply agglomerations of individual credences.<sup>31</sup> The thought here, however, is that in the finite accuracy dominance argument, *all* credences in the agglomeration contribute to the accuracy scores of the representing credence function (perhaps differentially weighted), which are then analyzed for dominance relations. Doing so is motivated by the epistemic ideal of an accuracy undominated epistemic state, where every credal state in the epistemic state is accounted for in scoring. By similarly requiring every credal state in an *infinite* agglomeration to contribute to the accuracy scores of the representing credence function that are then analyzed for dominance relations, I prove results that allow one to extend the accuracy dominance argument while retaining this underlying motivating ideal.

Now, Schervish et al.’s Theorem 1 alone does not establish the cited crucial mathematical fact when credence functions are defined over an infinite opinion set, for it suggests no way to score credence functions over infinite opinion sets in the first place. Therefore, an appeal to their result would not be motivated by a commitment to the epistemic value of having an overall accurate epistemic state, a commitment that underwrites the existing accuracy dominance arguments for probabilism. Moreover, while their Theorem 1 establishes that for an incoherent credence function  $c$ , the restriction of  $c$  to any finite subset on which it is incoherent is strictly dominated by a coherent credence function, the dominating credence function may depend on the finite subset.<sup>32</sup> Thus, their result does not provide even an alternative sense in which every incoherent credence function is dominated by a single coherent credence function, which is the kind of result appealed to in the argument for finite accuracy dominance.

<sup>30</sup> One notable exception is that unlike Joyce (2009), Joyce (1998) does not appeal to the second half of this key fact.

<sup>31</sup> I am grateful to an anonymous referee for raising this objection.

<sup>32</sup> In fact, this dependence may hold even in the finite case. See Schervish et al. (2009, example 10).

Of course, even if a local dominance result like that of Schervish et al. cannot be used to extend to the infinite setting what is often referred to as the *accuracy dominance argument for probabilism*, it could be used to establish a different accuracy-based argument for probabilism in the infinite setting. Thus, we should not assume that from an accuracy perspective, probabilism in the infinite setting will stand or fall based on the results in this article—that depends, for instance, on whether it is irrational *in itself* to be in an epistemic state where restrictions of one’s representing credence function are accuracy dominated. My point is simply that whatever accuracy-based arguments one can give for probabilism in the infinite setting, it is important to determine whether the influential set of arguments that are often referred to under the single heading of “the accuracy dominance argument for probabilism” must be restricted to finite opinion sets. In light of these considerations, to make clear that I will be concerned with comparing the “total” inaccuracy scores of credence functions, let us call the inaccuracy (resp., accuracy) of an agent’s entire epistemic state—that is, the inaccuracy of the agent’s full credence function—the *total inaccuracy* (resp., total accuracy) of the agent’s epistemic state.

## 6.2 The status of the accuracy dominance argument for probabilism on infinite opinion sets

Given the results of the last three sections, what should we conclude about the accuracy dominance argument for probabilism and more generally about the accuracy framework applied to credence functions defined on infinite opinion sets? When an agent’s epistemic state contains only finitely many credences, the concern for total accuracy leads to a dominance justification for probabilism (using Theorem 2.9). However, we saw earlier that when the agent’s epistemic state includes even countably infinitely many credences, the concern for total accuracy does not so clearly lead to a dominance justification for probabilism because coherence is not sufficient to avoid accuracy dominance in all cases (as in Examples 4.1 and 4.2). I suggest that the main normative challenges at this point are i) to clarify this asymmetry in the relationship between coherence and total accuracy dominance for finite and infinite opinion sets and ii) to determine what this asymmetry means for the accuracy framework more generally.

One response is to deny that real-world agents ever have infinitely many credences at a time, and thus this entire discussion is mere ideal theory. The empirical claim that real-world agents are restricted to finite opinion sets is not obvious, however. For example, if there were a coin in front of me of unknown bias, could I not have credences in each of the propositions “the bias of the coin is  $x$ ” for  $x \in [0, 1]$ ?<sup>33</sup> Or could I not have credence  $2^{-n}$  in the proposition “the coin would land heads  $n$  times in a row if I flipped it  $n$  times in a row” for each  $n \in \mathbb{N}$ ? More generally, it seems we can have credences parameterized by some infinite set such as the natural numbers. Thus, this kind of objection to considering infinite opinion sets at all seems unlikely to work or, at the very least, requires further defense.

<sup>33</sup> Pettigrew (2016, 222) gives this example to motivate dropping his assumption that the opinion set is finite.

Alternatively, one could argue that any asymmetry is unimportant with regard to probabilism: while the results show that some coherent credence functions are ruled out as irrational on the basis of total accuracy dominance, all incoherent credence functions are ruled out on the same basis (in light of Theorems 3.4 and 5.3). Thus, probabilism is justified similarly in both the finite and infinite case: one ought to have at least coherent credences so as to avoid total accuracy dominance. The only difference is that in the infinite case, coherence is not enough to avoid total accuracy dominance. I find this response promising. However, one complication is that although Theorems 3.4 and 5.3 together guarantee that incoherent credence functions are dominated by coherent credence functions, no matter the infinite opinion set, they do not guarantee dominance by an *undominated* coherent credence function in general.<sup>34</sup> But as discussed in Footnote 17, it has been argued that being accuracy dominated is not irrational in itself, but rather, being dominated by a credence function that is itself not dominated is irrational.<sup>35</sup> Thus, this line of response would require proving additional results or appealing to a controversial decision-theoretic principle.

A more radical response would be to claim that upon examination, there is simply no reasonable formal measure of total inaccuracy for an agent with infinitely many credences. In particular, no generalized legitimate inaccuracy measure, as I have been calling them, is in fact a legitimate measure of total inaccuracy in the infinite setting. Thus, to justify probabilism on any opinion set using accuracy considerations, there is no other option but to appeal to a local dominance result like that of Schervish et al. (2009). I see three challenges with this response. First, if there is no way to evaluate epistemic states with infinitely many credences for total inaccuracy, i) is the ability to justify norms on epistemic states in terms of total inaccuracy restricted in scope, and ii) if so, might this restriction in scope have negative consequences for the more general project of justifying epistemic norms by appealing to total inaccuracy? For example, if concern for total accuracy can only motivate epistemic norms in the finite case so that one must appeal to other epistemic virtues in the infinite case anyway, why care so much about total accuracy in the finite case?

Second, there are infinite opinion sets for which the natural extensions of the inaccuracy measures used in the finite case seem to behave as they do in the finite case with respect to accuracy dominance (e.g., for countably infinite partitions). Do we reject the legitimacy of generalized legitimate inaccuracy measures even in these cases where the measures behave as we would like? If so, why? And if not, then work must be done to spell out precisely for which infinite opinion sets there is no reasonable way to measure total inaccuracy. Presumably, Theorems 4.19 and 4.29 would be helpful toward this end.

Third, if one is to deny the legitimacy of what I have called generalized legitimate inaccuracy measures, then one needs to explain what exactly is wrong with them,

<sup>34</sup> It is an important open question whether Theorems 3.4 and 5.3 can be strengthened to conclude that the dominating coherent credence function can always be assumed to be undominated.

<sup>35</sup> See Pettigrew (2016, 20–21) for a counterexample to the stronger principle that dominance alone is irrational.

given their structural and axiomatic similarity to the inaccuracy measures often used to measure the total accuracy of credence functions on finite opinion sets.<sup>36</sup> Although one might think that a simple response is that no credence function should be infinitely inaccurate according to a reasonable inaccuracy measure, this response is not so obviously compelling. First, as Pettigrew (2016) points out, it is hard to assess the plausibility of this claim because whether inaccuracy can be infinite or not “is not something which our concept of accuracy contains much information about” (37). Another issue with this response, also raised by Pettigrew, is that the logarithmic inaccuracy measure—often seen as a reasonable way to measure inaccuracy in the finite case (see, e.g., McCutcheon 2019)—can output an infinite inaccuracy score even for finite opinion sets. So if one rejects generalized legitimate inaccuracy measures by virtue of their outputting infinite values, this will have consequences for what the legitimate inaccuracy measures are in the finite case. Lastly, insisting on only inaccuracy measures that output a finite inaccuracy score would justify focusing on a subclass of the generalized legitimate inaccuracy measures rather than justify rejecting all of them.<sup>37</sup> In fact, it is an interesting open question whether this restriction to the finite generalized legitimate inaccuracy measures picks out a set of generalized legitimate inaccuracy measures for which the finite dominance result goes through to all countably infinite opinion sets.<sup>38</sup>

To conclude, I do not claim to have provided a full analysis of the options available for responding to the results in this paper, but I hope to have shown that further philosophical work is in order if we are to understand their implications for probabilism and accuracy-based justifications of epistemic norms more generally.

## 7. Conclusion

As discussed in the previous section, there is plenty of normative work to be done using the results established in this article. In light of the failure of coherence being sufficient to avoid strong dominance on certain countably infinite opinion sets, the most pressing question seems to be this: is there an accuracy dominance argument for probabilism on at least all countable opinion sets? If not, what does this mean for the accuracy project as a whole? Can we give some sort of privileged status to certain kinds of opinion sets or inaccuracy measures for which coherence is equivalent to

---

<sup>36</sup> For example, generalized legitimate inaccuracy measures satisfy all of Pettigrew’s (2016, 65) axioms on a legitimate measure of inaccuracy except for his decomposition axiom because it is not clear how to even define a key part of the decomposition axiom—what he calls the “well calibrated counterpart” of a credence function—in the infinite setting. See also footnote 14. Clearly, generalized legitimate inaccuracy measures do not satisfy Joyce’s (2009) “coherent admissibility” in general. However, generalized legitimate inaccuracy measures do satisfy coherent admissibility when restricted to countably infinite partitions (Theorem 4.26), and Joyce (2009) restricts them to credence functions on partitions in his argument in the finite case. Further analysis of the axiomatic properties of generalized legitimate inaccuracy measures is left for future work.

<sup>37</sup> For example, if the weights defining a generalized legitimate inaccuracy measure are summable, then inaccuracy is always finite.

<sup>38</sup> Walsh’s (ms) result discussed earlier (see Theorem 3.7) may provide some evidence in the affirmative.

not being dominated (e.g., partitions)? What is the normative status of the stronger condition of countable coherence? Further, although the measure theoretic framework introduced in Section 5 to score the inaccuracy of credence functions over opinion sets of arbitrary cardinality seems like a natural extension of the finite and countably infinite frameworks, is it well motivated that inaccuracy does not track the behavior of a credence function on measure zero sets? The hope with this paper is to start a conversation about these questions by first establishing relevant mathematical results.

**Acknowledgements.** Thanks to participants of the Berkeley-Stanford Logic Circle (April 2019), Probability and Logic Conference (July 2019), Berkeley Formal Epistemology Reading Course (October 2019), and Stanford Logic and Formal Philosophy Seminar (November 2019), to whom earlier versions of this article were presented. For helpful comments and discussion, thanks to Kenny Easwaran, Craig Evans, Wesley Holliday, Thomas Icard, Kiran Luecke, Calum McNamara, Sven Neth, Richard Pettigrew, Eric Raidl, Teddy Seidenfeld, and James Walsh. Special thanks to two anonymous referees at *Philosophy of Science* for their very useful feedback.

## References

- Banerjee, Arindam, Xin Guo, and Hui Wang. 2005. "On the Optimality of Conditional Expectation as a Bregman Predictor." *IEEE Transactions on Information Theory* 51 (7):2664–9.
- Borkar, Vivek, Vijay Konda, and Sanjoy Mitter. 2003. "On de Finetti Coherence and Kolmogorov Probability." *Statistics & Probability Letters* 66 (4):417–21.
- Csiszár, Imre. 1995. "Generalized Projections for Non-Negative Functions." *Acta Mathematica Hungarica* 68 (1–2):161–86.
- De Bona, Glauber, and Julia Staffel. 2018. "Why Be (Approximately) Coherent?" *Analysis* 78 (3):405–15.
- de Finetti, Bruno. 1974. *Theory of Probability*. New York: John Wiley.
- Easwaran, Kenny. 2013. "Expected Accuracy Supports Conditionalization—and Conglomerability and Reflection." *Philosophy of Science* 80 (1):119–42.
- Easwaran, Kenny. 2016. "Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probabilities." *Nous* 50 (4):816–53.
- Gneiting, Tilmann, and Adrian E. Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association* 102 (477):359–78.
- Greaves, Hilary, and David Wallace. 2005. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind* 115 (459):607–32.
- Horn, Alfred, and Alfred Tarski. 1948. "Measures in Boolean Algebras." *Transactions of the American Mathematical Society* 64 (November):467–97.
- Horwich, Paul. 1982. *Probability and Evidence*. Cambridge: Cambridge University Press.
- Hrbacek, Karel, and Thomas Jech. 1999. *Introduction to Set Theory*. New York: Marcel Dekker.
- Huttenegger, Simon M. 2013. "In Defense of Reflection." *Philosophy of Science* 80 (3):413–33.
- Joyce, James. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65 (4):575–603.
- Joyce, James. 2009. "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In *Degrees of Belief*, edited by F. Huber and C. Schmidt-Petri, 263–97. New York: Springer.
- Kelley, Mikayla. 2019. "Accuracy Dominance on Infinite Opinion Sets." PhD diss., University of California, Berkeley.
- Leitgeb, Hannes, and Richard Pettigrew. 2010a. "An Objective Justification of Bayesianism I: Measuring Inaccuracy." *Philosophy of Science* 77 (2):201–35.
- Leitgeb, Hannes, and Richard Pettigrew. 2010b. "An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy." *Philosophy of Science* 77 (2):236–72.
- Levinstein, Benjamin A. 2019. "An Objection of Varying Importance to Epistemic Utility Theory." *Philosophical Studies* 176 (11):2919–31.

- Lindley, Dennis V. 1987. "Scoring Rules and the Inevitability of Probability." In *System Design for Human Interaction*, edited by A. P. Sage, 182–208. Piscataway, NJ: IEEE Press.
- Maher, Patrick. 2002. "Joyce's Argument for Probabilism." *Philosophy of Science* 69 (1):73–81.
- McCutcheon, Randall G. 2019. "In Favor of Logarithmic Scoring." *Philosophy of Science* 86 (2):286–303.
- Nielsen, Michael. 2022. "On the Best Accuracy Arguments for Probabilism." *Philosophy of Science* 89 (3): 621–30.
- Pettigrew, Richard. 2016. *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.
- Pettigrew, Richard. 2022. "Accuracy-First Epistemology without Additivity." *Philosophy of Science* 89 (1):128–51.
- Predd, J. B., R. Seiringer, E. H. Lieb, D. N. Osherson, H. V. Poor, and S. R. Kulkarni. 2009. "Probabilistic Coherence and Proper Scoring Rules." *IEEE Transactions on Information Theory* 55 (10):4786–92.
- Rosenkrantz, Roger D. 1981. *Foundations and Applications of Inductive Probability*. Atascadero, CA: Ridgeview Press.
- Savage, Leonard. 1971. "Elicitation of Personal Probabilities and Expectations." *Journal of the American Statistical Association* 66 (336):783–801.
- Schervish, Mark, Teddy Seidenfeld, and Joseph B. Kadane. 2009. "Proper Scoring Rules, Dominated Forecasts, and Coherence." *Decision Analysis* 6 (4):202–21.
- Schervish, Mark, Teddy Seidenfeld, and Joseph B. Kadane. 2014. "Dominating Countably Many Forecasts." *Annals of Statistics* 42 (2):728–56.
- Staffel, Julia. 2019. *Unsettled Thoughts: A Theory of Degrees of Rationality*. Oxford: Oxford University Press.
- Steeger, Jeremy. 2019. "Probabilism for Stochastic Theories." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 66, 34–44.
- Tao, Terence. 2011. *An Introduction to Measure Theory*. Providence, RI: American Mathematical Society.
- Walsh, Sean. ms. "Probabilism in infinite dimensions." Unpublished manuscript.

## A. Appendix

### A.1 Proof of Theorem 3.4

I review the necessary background before proving Theorem 3.4.

#### A.1.1 Generalized Projections

Csiszár (1995) showed that what he calls *generalized projections* onto convex sets with respect to Bregman distances exist under very general conditions. I review his relevant results here (but assume knowledge of basic measure theory).

**Definition A.1.** Fix a  $\sigma$ -finite measure space  $(X, \mathcal{X}, \mu)$ . The *Bregman distance* of nonnegative ( $\mathcal{X}$ -measurable) functions  $s$  and  $t$  is defined by

$$B_{\varphi, \mu}(s, t) = \int \mathfrak{D}(s(x), t(x)) \mu(dx) \in [0, \infty],$$

where  $\mathfrak{D}(s(x), t(x)) = \varphi(s(x)) - \varphi(t(x)) - \varphi'(t(x))(s(x) - t(x))$  for some strictly convex, differentiable function  $\varphi$  on  $(0, \infty)$ .<sup>39</sup> Note that  $B_{\varphi, \mu}(s, t) = 0$  if  $s = t$   $\mu$ -a.e. See Csiszár (1995, 165) for details.

**Remark A.2.** Notice that a generalized quasi-additive Bregman divergence  $\mathfrak{D}$  with weights  $\{a_i\}_{i=1}^{\infty}$  whose generated one-dimensional Bregman divergence  $\mathfrak{D}$  is given in terms of  $\varphi$  has a corresponding Bregman distance  $B_{\varphi, \mu}$  with

<sup>39</sup> For  $B_{\varphi, \mu}$  to be a distance measure, we do not need to assume that  $\varphi(1) = \varphi'(1) = 0$  by the remark following equation (1.9) in Csiszár (1995).

1. the measure space being  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ , where  $\mu(A) = \sum_{i \in A} a_i$  for each  $A \in \mathcal{P}(\mathbb{N})$ ;
2.  $\bar{\varphi}$  on  $(0, \infty)$  being a strictly convex, differentiable extension of  $\varphi$  on  $[0, 1]$ .<sup>40</sup>

Thus, nonnegative ( $\mathcal{P}(\mathbb{N})$ -measurable) functions are elements of  $\mathbb{R}^{+\infty}$ . Note, importantly, that the corresponding generalized legitimate inaccuracy measure  $\mathcal{F}$  determined by  $\mathfrak{D}$  is also given by the corresponding Bregman distance. That is,

$$\mathcal{F}(c, w) = B_{\bar{\varphi}, \mu}(v_w, c).$$

To simplify notation, let  $B$  denote a Bregman distance  $B_{\bar{\varphi}, \mu}$ . Let  $S$  be the set of nonnegative measurable functions on  $(X, \mathcal{X}, \mu)$ . For any  $E \subseteq S$  and  $t \in S$ , we write

$$B(E, t) = \inf_{s \in E} B(s, t).$$

If there exists  $s^* \in E$  with  $B(s^*, t) = B(E, t)$ , then  $s^*$  is unique and is called the *B-projection of  $t$  onto  $E$*  (see Csiszár 1995, Lemma 2). As Csiszár notes, these projections may not exist. However, a weaker kind of projection exists in a large number of cases. To describe them, we need to introduce a kind of convergence called *loose in  $\mu$ -measure convergence*.

**Definition A.3.** We say a sequence  $\{s_n\}$  of elements from  $S$  converges *loosely in  $\mu$ -measure* to  $t$ , denoted by  $s_n \rightsquigarrow_{\mu} t$ , if for every  $A \in \mathcal{X}$  with  $\mu(A) < \infty$ , we have

$$\lim_{n \rightarrow \infty} \mu(A \cap \{p : |s_n(p) - t(p)| > \epsilon\}) = 0 \text{ for all } \epsilon > 0.$$

**Definition A.4.**

- i. Given  $E \subseteq S$  and  $t \in S$ , we say that a sequence  $\{s_n\}$  of elements from  $E$  is a *B-minimizing sequence* if  $B(s_n, t) \rightarrow B(E, t)$ .
- ii. If there is an  $s^* \in S$  such that every B-minimizing sequence converges to  $s^*$  loosely in  $\mu$ -measure, then we call  $s^*$  the *generalized B-projection of  $t$  onto  $E$* .

The result that is integral to proving Theorem 3.4 is the following (see Csiszár’s Theorem 1, Lemma 2, and corollary of Theorem 1).

**Theorem A.5.** (Csiszar 1995) Let  $E$  be a convex subset of  $S$  and  $t \in S$ . If  $B(E, t)$  is finite, then there exists  $s^* \in S$  such that

$$B(s, t) \geq B(E, t) + B(s, s^*) \text{ for every } s \in E,$$

and  $B(E, t) \geq B(s^*, t)$ . It follows that the generalized B-projection of  $t$  onto  $E$  exists and equals  $s^*$ .

### A.1.2 Extending Partial Measures

We also use an extension result of Horn and Tarski (1948) in the proof of Theorem 3.4. Following Horn and Tarski, we introduce *partial measures* and recall that they can be extended to finitely additive probability functions. Recall the definition of a finitely additive probability function in Definition 2.2 (although we drop the assumption that  $\mathcal{F}$  is finite).

---

<sup>40</sup> Using that  $\varphi'$  exists and is finite at  $x = 1$  as we assumed  $\mathfrak{d}$  is bounded, we can extend  $\varphi$  as follows: for  $x \in [1, \infty)$ , let  $\bar{\varphi}(x) = q(x) = x^2 + bx + c$ , where  $b$  and  $c$  are chosen so that  $\varphi(1) = q(1)$  and  $\varphi'(1) = q'(1)$ . Then, using the fact that  $\bar{\varphi}$  is differentiable at 1 by construction, and a function is strictly convex if and only if its derivative is strictly increasing, it is easy to see that  $\bar{\varphi}$  is differentiable and strictly convex on  $(0, \infty)$ .

**Remark A.6.** It is a simple corollary of the definition of a finitely additive probability function  $c$  over an algebra  $\mathcal{F}$  that for any  $p, p' \in \mathcal{F}$ : if  $p \subseteq p'$ , then  $c(p) \leq c(p')$ .

Here is another useful fact about finitely additive probability functions:

**Proposition A.7.** If  $c$  is a finitely additive probability function on an algebra  $\mathcal{F}$  and  $a_0, \dots, a_{m-1} \in \mathcal{F}$ , then

$$\sum_{k=0}^{m-1} c(a_k) = \sum_{k=0}^{m-1} c\left(\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} a_{p_i}\right), \tag{4}$$

where  $S^{m,k}$  is the set of all sequences  $p = (p_0, \dots, p_k)$  with  $0 \leq p_0 < \dots < p_k < m$ .

To introduce the notion of a partial measure, we need the following definition:

**Definition A.8.** Let  $\varphi_0, \dots, \varphi_{m-1}$  and  $\psi_0, \dots, \psi_{n-1}$  be elements of  $\mathcal{F}$ . Then we write

$$(\varphi_0, \dots, \varphi_{m-1}) \subseteq (\psi_0, \dots, \psi_{n-1})$$

to mean

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \subseteq \bigcup_{p \in S^{n,k}} \bigcap_{i \leq k} \psi_{p_i} \text{ for every } k < m, \tag{5}$$

where  $S^{r,k}$  ( $r = m, n$ ) is as in Proposition A.7.<sup>41</sup>

**Definition A.9.** A function  $c$ , defined on a subset  $S$  of an algebra  $\mathcal{F}$  over  $W$ , that maps to  $\mathbb{R}$  is called a *partial measure* if it satisfies the following properties:

1.  $c(x) \geq 0$  for  $x \in S$ ;
2. if  $\varphi_0, \dots, \varphi_{m-1}, \psi_0, \dots, \psi_{n-1} \in S$  and

$$(\varphi_0, \dots, \varphi_{m-1}) \subseteq (\psi_0, \dots, \psi_{n-1}),$$

then

$$\sum_{k=0}^{m-1} c(\varphi_k) \leq \sum_{k=0}^{n-1} c(\psi_k); \text{ and}$$

3.  $W \in S$  and  $c(W) = 1$ .

The following result is the point of introducing the previous definitions.

**Theorem A.10.** (Horn and Tarski 1948) Let  $c$  be a partial measure on a subset  $\mathcal{F}$  of an algebra  $\mathcal{A}$ . Then there is a finitely additive probability function  $c^*$  on  $\mathcal{A}$  that extends  $c$ .

### A.1.3 Proof

We now establish the necessity of coherence to avoid dominance.

**Theorem 3.4.** Let  $\mathcal{F}$  be a countably infinite opinion set,  $\mathcal{J}$  a generalized legitimate inaccuracy measure, and  $c$  an incoherent credence function. Then:

<sup>41</sup> Note that if  $m > n$ , this condition implies  $\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} = \bigcup_{p \in S^{n,k}} \bigcap_{i \leq k} \psi_{p_i} = \emptyset$  for  $k \geq n$ .

1.  $c$  is weakly dominated relative to  $\mathcal{F}$  by a coherent credence function; and
2. if  $\mathcal{F}(c, w) < \infty$  for each  $w \in W$ , then  $c$  is strongly dominated relative to  $\mathcal{F}$  by a coherent credence function.

**Proof.** Let  $\mathcal{F}$  be a generalized legitimate inaccuracy measure and thus defined by a Bregman distance  $B_{\phi, \mu}$  (see Remark A.2). We write  $B$  for  $B_{\phi, \mu}$ . Let  $S$  be the set of nonnegative functions on  $\mathcal{F}$ . Let  $E \subseteq S$  be the set of coherent credence functions on  $\mathcal{F}$ . Then clearly  $E$  is convex.

Let  $c$  be an incoherent credence function.

**Case 1:**  $\mathcal{F}(c, w) = \infty$  for all  $w \in W$ . Then because  $\mathcal{F}(v_w, w) = 0$  for all  $w \in W$ , any omniscient credence function weakly dominates  $c$ .

**Case 2:**  $\mathcal{F}(c, w') < \infty$  for some  $w' \in W$ . We can show that there is a coherent credence function  $\pi_c$  such that

$$\mathcal{F}(c, w) > \mathcal{F}(\pi_c, w) \text{ for any } w \text{ such that } \mathcal{F}(c, w) < \infty.$$

Since  $v_{w'} \in E$ , we see that

$$B(E, c) \leq B(v_{w'}, c) = \mathcal{F}(c, w') < \infty.$$

Thus, we can apply Theorem A.5 to get a  $\pi_c \in S$  such that

$$B(s, t) \geq B(E, c) + B(s, \pi_c) \text{ for every } s \in E. \tag{6}$$

In particular, (6) holds when  $s$  is the omniscient credence function at world  $w$  for any  $w \in W$ , and so we see that

$$\mathcal{F}(c, w) \geq B(E, c) + \mathcal{F}(\pi_c, w) \tag{7}$$

for all  $w$ , where all numbers in (7) are finite whenever  $\mathcal{F}(c, w) < \infty$ .

Next, we can show that  $\pi_c$  is in fact coherent. This is due to the following claim:  $E$  is closed under loose convergence in  $\mu$ -measure, where  $\mu$  is a weighted counting measure on  $\mathcal{P}(\mathbb{N})$  defined with weights  $\{a_i\}_{i=1}^\infty$ . To see this, let  $c_n \in E$  for each  $n$  and  $c \in S$ . Assume  $c_n \rightarrow c$  loosely in  $\mu$ -measure. We can show  $c \in E$ ; that is,  $c$  is coherent. Note that  $c$  is coherent on  $\mathcal{F}$  if and only if  $c' : \mathcal{F} \cup \{W\} \rightarrow [0, 1]$  is coherent on  $\mathcal{F} \cup \{W\}$ , where  $c' = c$  on  $\mathcal{F}$  and  $c'(W) = 1$ . Thus, it suffices to assume  $c$  and  $c_n$  for all  $n$  are defined on  $\mathcal{F} \cup \{W\}$  with  $c(W) = c_n(W) = 1$  for all  $n$ .

It is easy to see that loose convergence in a weighted counting measure (where all weights are nonzero) implies pointwise convergence on  $\mathcal{F}$ , so

$$c(p) = \lim_{n \rightarrow \infty} c_n(p) \in [0, 1]$$

for each  $p \in \mathcal{F} \cup \{W\}$ . To show  $c \in E$ , it suffices to show that  $c$  can be extended to a finitely additive probability function on  $\mathcal{P}(W)$ .

First, we can show that  $c$  is a partial measure on  $\mathcal{F} \cup \{W\}$ . Definitions A.9.1 and A.9.3 clearly hold for  $c$ , so we just need to show that Definition A.9.2 holds. Let  $\varphi_0, \dots, \varphi_{m-1}, \psi_0, \dots, \psi_{m'-1} \in \mathcal{F} \cup \{W\}$  and

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \subseteq \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i}$$

for every  $k < m$ . Because the  $c_n$  are coherent and thus extend to finitely additive probability functions on algebras containing  $\mathcal{F}$ , we have by Proposition A.7 and Remark A.6 that

$$\sum_{k=0}^{m-1} c_n(\varphi_k) = \sum_{k=0}^{m-1} c_n \left( \bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \right) \leq \sum_{k=0}^{m'-1} c_n \left( \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i} \right) = \sum_{k=0}^{m'-1} c_n(\psi_k)$$

using that

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} = \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i} = \emptyset$$

for  $k \geq m'$ . Sending  $n$  to infinity and using the pointwise convergence of  $c_n$  to  $c$  on  $\mathcal{F} \cup \{W\}$ , we obtain that

$$\sum_{k=0}^{m-1} c(\varphi_k) \leq \sum_{k=0}^{m'-1} c(\psi_k).$$

Thus,  $c$  is a partial measure on  $\mathcal{F} \cup \{W\}$ . By Theorem A.10, it follows that there is a finitely additive probability function  $c^*$  on an algebra  $\mathcal{F}^* \supseteq \mathcal{F}$  that extends  $c$ , and so  $c \in E$ , which concludes the proof that  $E$  is closed under loose  $\mu$ -convergence.

By Theorem A.5,  $\pi_c$  is the generalized  $B$ -projection of  $c$  onto  $E$ . Also, because

$$B(E, c) = \inf_{s \in E} (s, c) < \infty,$$

there is a  $B$ -minimizing sequence  $\{s_n\} \subseteq E$  such that  $B(s_n, c) \rightarrow B(E, c)$  by the definition of infimum. By the definition of a generalized projection,  $s_n \rightsquigarrow_\mu \pi_c$ . Since  $E$  is closed under loose convergence, it follows that  $\pi_c \in E$ . Further, by Theorem A.5,

$$B(E, c) \geq B(\pi_c, c) > 0$$

because  $\pi_c \neq c$  (because  $c$  is incoherent) and  $B(s, t) = 0$  if and only if  $s = t$  (because  $\mu$  is a weighted counting measure with all nonzero weights). Thus, for every  $w$  such that  $\mathcal{F}(c, w) < \infty$ , we deduce that

$$\mathcal{F}(c, w) \geq B(E, c) + \mathcal{F}(\pi_c, w) > \mathcal{F}(\pi_c, w).$$

This proves that  $c$  is weakly dominated by  $\pi_c$ , and  $c$  is strongly dominated by  $\pi_c$  if  $\mathcal{F}(c, w) < \infty$  for all  $w \in W$ . □

### A.2 Proofs from Section 4

We assume throughout this section that  $\mathcal{F}$  is countably infinite.

**Proposition 4.7.** Let  $\mathcal{F}$  be a countably infinite opinion set that is countably discriminating (thus,  $V_{\mathcal{F}}$  is countable). Then a credence function  $c$  is countably coherent if and only if there are  $\lambda_{v_w} \in [0, 1]$  with  $\sum_{v_w \in V_{\mathcal{F}}} \lambda_{v_w} = 1$  such that

$$c(p) = \sum_{v_w \in V_{\mathcal{F}}} \lambda_{v_w} u_w(p)$$

for all  $p \in \mathcal{F}$ .

**Proof.** I adapt the proof of Proposition 1 of Predd et al. (2009). Let  $\mathcal{F} = \{p_1, p_2, \dots\}$ . Let  $\mathcal{X}$  be the collection of all nonempty sets of the form  $\bigcap_{i=1}^{\infty} p_i^*$ , where  $p_i^*$  is either  $p_i$  or  $p_i^c$ . Then  $\mathcal{X}$  partitions  $W$ . Also,  $\mathcal{X}$  is in bijection with  $V_{\mathcal{F}}$ , the set of omniscient credence functions.

Indeed, let  $f$  map  $v_w$  to  $\bigcap_{i=1}^{\infty} p_i^*$ , where  $p_i^* = p_i$  if  $v_w(p_i) = 1$  and  $p_i^* = p_i^c$  otherwise. Then for each  $w$ ,  $w \in f(v_w)$ , and so  $f(v_w) \in \mathcal{X}$ . Note that  $f$  is onto. Indeed, let  $w \in \bigcap_{i=1}^{\infty} p_i^*$ , where  $\bigcap_{i=1}^{\infty} p_i^* \in \mathcal{X}$ . Then  $f(v_w) = \bigcap_{i=1}^{\infty} p_i^*$ . Also,  $f$  is injective. Indeed, assume  $f(v_w) = f(v_{w'})$ . Then

$$f(v_w) = \bigcap_{i=1}^{\infty} p_i^1 = \bigcap_{i=1}^{\infty} p_i^2 = f(v_{w'})$$

for  $p_i^1 = p_i$  or  $p_i^1 = p_i^c$  for all  $i \in \mathbb{N}$  and  $j \in \{1, 2\}$ . If  $p_i^1 \neq p_i^2$  for some  $i$ , then without loss of generality, we may assume  $p_i^1 = p_i$  and  $p_i^2 = p_i^c$ . So  $w \in p_i^1$  but  $w \notin p_i^2$ , and thus  $w \notin \bigcap_{i=1}^{\infty} p_i^2$ . But  $w \in \bigcap_{i=1}^{\infty} p_i^1$  by definition of  $f$  and so  $\bigcap_{i=1}^{\infty} p_i^1 \neq \bigcap_{i=1}^{\infty} p_i^2$ , which is a contradiction. It follows that  $p_i^1 = p_i^2$  for all  $i$ , but then by definition of  $f$ , this implies  $v_w(p_i) = 1$  if and only if  $v_{w'}(p_i) = 1$  for all  $i$ , and so  $v_w = v_{w'}$ .

It is easy to see that because  $\mathcal{F}$  is countably discriminating,  $V_{\mathcal{F}}$  is countable. It follows that  $\mathcal{X}$  is countable. Enumerate the elements of  $V_{\mathcal{F}}$  and  $\mathcal{X}$  by  $v_{w_1}, v_{w_2}, \dots$  and  $e_1, e_2, \dots$ , respectively, such that  $f^{-1}(e_j) = v_{w_j}$ . We have that  $p_i$  is the disjoint union of  $e_j$  such that  $e_j \subseteq p_i$  or, equivalently, the  $e_j$  where  $f^{-1}(e_j)(p_i) = 1$ . Note that i) for any countably additive probability function  $\mu$  on a  $\sigma$ -algebra containing  $\mathcal{F}$  (and thus containing  $\mathcal{X}$ ) and any  $p_i \in \mathcal{F}$ :

$$\mu(p_i) = \sum_{j=1}^{\infty} \mu(e_j) f^{-1}(e_j)(p_i).$$

Now we can prove the equivalence. Assume  $c$  is countably coherent. By the definition of countable coherence,  $c$  extends to a countably additive probability function  $\mu$  on a  $\sigma$ -algebra containing  $\mathcal{F}$ . Then by i),

$$c(p_i) = \mu(p_i) = \sum_{j=1}^{\infty} \mu(e_j) f^{-1}(e_j)(p_i)$$

for all  $p_i \in \mathcal{F}$ . But because  $\mu(e_j)$  are nonnegative and sum to 1 (because the  $e_j$  values partition  $W$ , and  $\mu$  is a countably additive probability function), we have that  $c$  has the form stated.

Now assume  $c(p_i) = \sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i)$  for all  $i$ , where  $\sum_{j=1}^{\infty} \lambda_j = 1$ . Let  $\sigma(\mathcal{F})$  be the smallest  $\sigma$ -algebra on  $W$  containing  $\mathcal{F}$ . Then it is easy to check that the function on  $\sigma(\mathcal{F})$  defined by  $\bar{v}_{w_j}(p) = 1$  if and only if  $w_j \in p$  extends  $v_{w_j}$  and is a countably additive probability function on  $\sigma(\mathcal{F})$ . Then  $\sum_{j=1}^{\infty} \lambda_j \bar{v}_{w_j}$  is a countably additive probability function on  $\sigma(\mathcal{F})$  since a countable sum of countably additive probability functions with coefficients that sum to 1 is a countably additive probability function. Because

$$c(p_i) = \sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) = \sum_{j=1}^{\infty} \lambda_j \bar{v}_{w_j}(p_i)$$

for all  $i$ , it follows that  $c$  extends to a countably additive probability function on a  $\sigma$ -algebra containing  $\mathcal{F}$ . □

**Lemma 4.11.** For  $(W, \mathcal{F})$ , an opinion space,  $(W^*, \mathcal{F}^*)$  is compact.

**Proof.** Let  $\{\Psi(p_n)^{f(n)}\}_{n=1}^{\infty}$  be a sequence of elements of  $\mathcal{F}^*$  or their complements, as in Definition 4.9. Case 1: For each  $N$ , there is some  $w_N \in W$  such that  $w_N \in \bigcap_{n=1}^N \Psi(p_n)^{f(n)}$ . Then, since i)  $\Psi(p) \cap W = p$  and ii)  $\Psi(p)^c \cap W = p^c$  for any  $p \in \mathcal{F}$ , it follows that  $w_N \in \bigcap_{n=1}^N p_n^{f(n)}$  for each  $N$ . If there is some  $w' \in W$  with  $w' \in \bigcap_{n=1}^{\infty} p_n^{f(n)}$ , then, by i) and ii), it follows that  $w' \in \bigcap_{n=1}^{\infty} \Psi(p_n)^{f(n)}$ . Otherwise, by construction, we defined some  $x_s$  to be such that  $x_s \in \bigcap_{n=1}^{\infty} \Psi(p_n)^{f(n)}$ . In either case, we are done. Case 2: There is some  $N$  such that  $\bigcap_{n=1}^N \Psi(p_n)^{f(n)} \subseteq W^* \setminus W$ . I claim this implies that  $\bigcap_{n=1}^N \Psi(p_n)^{f(n)} = \emptyset$ . Indeed, if there were some  $w \in W^* \setminus W$  such that  $w \in \bigcap_{n=1}^N \Psi(p_n)^{f(n)}$ , then that is because  $\{p_n^{f(n)}\}_{n=1}^N$  is an initial sequence of some sequence  $\{\bar{p}_n^{f(n)}\}_{n=1}^{\infty}$  such that  $\bigcap_{n=1}^{\infty} \bar{p}_n^{f(n)} \neq \emptyset$  for each  $l$  and thus, in particular,  $\bigcap_{n=1}^N \bar{p}_n^{f(n)} \neq \emptyset$ . Thus, there is some  $w \in W$  such that  $w \in \bigcap_{n=1}^N \Psi(p_n)^{f(n)}$  by i) and ii), which is a contradiction. Thus, we have established that  $(W^*, \mathcal{F}^*)$  is compact. □

**Lemma 4.12.** Let  $(W, \mathcal{F})$  be an opinion space and  $c$  be a coherent credence function on  $(W, \mathcal{F})$ . Let  $(W^*, \mathcal{F}^*)$  be the compactification of  $(W, \mathcal{F})$ , and define  $c^*(\Psi(p)) := c(p)$  for each  $p \in \mathcal{F}$ . Then  $c^*$  is a countably coherent credence function on  $(W^*, \mathcal{F}^*)$ , and  $\mathcal{C}(c, w) = \mathcal{C}(c^*, w)$  for  $w \in W$ .

**Proof.** Because  $(W^*, \mathcal{F}^*)$  is compact, we only need to show that  $c^*$  is coherent by Theorem 4.10. Thus, it suffices to show that  $c^*$  can be extended to a finitely additive probability function on  $\mathcal{A}(\mathcal{F}^*)$ . Since  $c$  is coherent, there is a finitely additive probability function  $\bar{c}$  such that:

1.  $\bar{c}(p) = c(p)$  for  $p \in \mathcal{F}$ ;
2.  $\bar{c}(p \cup q) = \bar{c}(p) + \bar{c}(q)$  for  $p, q \in \mathcal{F}$  with  $p \cap q = \emptyset$ ; and
3.  $\bar{c}(W) = 1$ .

First, define  $\Psi(p^c) := \Psi(p)^c$  for each  $p \in \mathcal{F}$ . Then each element in  $\mathcal{A}(\mathcal{F}^*)$  can be represented by  $\bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij})$ , where  $q_{ij}$  or its complement is in  $\mathcal{F}$ . We define

$$c^* \left( \bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) \right) := \bar{c} \left( \bigcup_{i=1}^N \bigcap_{j=1}^M q_{ij} \right).$$

Using that  $p = \Psi(p) \cap W$  and  $p^c = \Psi(p)^c \cap W$ , we can show that  $\bar{c}^*$  is a well-defined finitely additive probability function on  $\mathcal{A}(\mathcal{F}^*)$  extending  $c^*$ . First, we can show that  $\bar{c}^*$  is well defined. Assume that

$$\bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) = \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} \Psi(r_{ij}).$$

This clearly implies that

$$\bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) \cap W = \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} \Psi(r_{ij}) \cap W,$$

which, noting that  $p = \Psi(p) \cap W$  and  $p^c = \Psi(p)^c \cap W$ , establishes that

$$\bigcup_{i=1}^N \bigcap_{j=1}^M q_{ij} = \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} r_{ij},$$

and so

$$\bar{c}^* \left( \bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) \right) = \bar{c} \left( \bigcup_{i=1}^N \bigcap_{j=1}^M q_{ij} \right) = \bar{c} \left( \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} r_{ij} \right) = \bar{c}^* \left( \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} \Psi(r_{ij}) \right).$$

Thus,  $\bar{c}^*$  is well defined. Clearly,  $\bar{c}^*$  extends  $c^*$ . Now, since  $W \subseteq W^*$ , if

$$\bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) \cap \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} \Psi(r_{ij}) = \emptyset,$$

then

$$\bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) \cap W \cap \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} \Psi(r_{ij}) \cap W = \emptyset$$

and so

$$\bar{c} \left( \bigcup_{i=1}^N \bigcap_{j=1}^M q_{ij} \cup \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} r_{ij} \right) = \bar{c} \left( \bigcup_{i=1}^N \bigcap_{j=1}^M q_{ij} \right) + \bar{c} \left( \bigcup_{i=1}^{N'} \bigcap_{j=1}^{M'} r_{ij} \right).$$

Then, noting the definition of  $\bar{c}^*$  in terms of  $\bar{c}$ , we establish finite additivity. Lastly, if

$$W^* = \bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}),$$

then

$$W = \bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) \cap W,$$

and so

$$\bar{c}^* \left( \bigcup_{i=1}^N \bigcap_{j=1}^M \Psi(q_{ij}) \right) = \bar{c} \left( \bigcup_{i=1}^N \bigcap_{j=1}^M q_{ij} \right) = \bar{c}(W) = 1.$$

This establishes that  $c^*$  is coherent on  $(W^*, \mathcal{F}^*)$ , and so because  $(W^*, \mathcal{F}^*)$  is compact,  $c^*$  is countably coherent. Further,  $w \in p$  if and only if  $w \in \Psi(p)$  for each  $w \in W$ , so  $v_w$  defined on  $\mathcal{F}$  is the same as  $v_w$  defined on  $\mathcal{F}^*$  for each  $w \in W$ . Since  $c(p) = c^*(\Psi(p))$  for all  $p \in \mathcal{F}$ , this establishes that  $\mathcal{J}(c, w) = \mathcal{J}(c^*, w)$  for each  $w \in W$ . □

**Proposition 4.15.** Let  $(W, \mathcal{F})$  be an opinion space and  $\mathcal{I}$  be a generalized legitimate inaccuracy measure. If  $c$  is a countably coherent credence function on  $(W, \mathcal{F})$  with finite expected inaccuracy relative to  $\mathcal{I}$ , then  $c$  is not weakly dominated relative to  $\mathcal{I}$ .

**Proof.** Since  $c$  is countably coherent, let  $\bar{c}$  be a countably additive probability function on  $\sigma(\mathcal{F})$  extending  $c$  such that  $\mathbb{E}_{\bar{c}}\mathcal{I}(c, \cdot) < \infty$ . Note that  $\mathfrak{d}$  is a strictly proper inaccuracy measure for singleton opinion sets (see Remark 2.7 and Pettigrew [2016, Theorem 4.3.5] for a precise definition), which implies by definition that  $p\mathfrak{d}(1, x) + (1 - p)\mathfrak{d}(0, x)$  is uniquely minimized at  $x = p$ .<sup>42</sup> It follows that

$$\mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, c_i) = c_i\mathfrak{d}(1, c_i) + (1 - c_i)\mathfrak{d}(0, c_i) < c_i\mathfrak{d}(1, x) + (1 - c_i)\mathfrak{d}(0, x) = \mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, x)$$

for any  $x \neq c_i$ .

Assume toward a contradiction that there is a credence function  $d$  with  $d \neq c$  and  $\mathcal{I}(d, w) \leq \mathcal{I}(c, w)$  for each  $w$  with strict inequality for some  $w$ . Then  $\mathbb{E}_{\bar{c}}\mathcal{I}(d, \cdot) \leq \mathbb{E}_{\bar{c}}\mathcal{I}(c, \cdot) < \infty$ , so both  $\mathcal{I}(d, \cdot)$  and  $\mathcal{I}(c, \cdot)$  are integrable with respect to the measure space  $(W, \sigma(\mathcal{F}), \bar{c})$ . Then let  $i$  be any index such that  $d_i \neq c_i$ . There must be at least one because  $c \neq d$ . Then

$$\mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, c_i) < \mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, d_i).$$

If  $i$  is such that  $d_i = c_i$ , then clearly,  $\mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, c_i) = \mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, d_i)$ . So because  $\mathbb{E}_{\bar{c}}\mathcal{I}(c, \cdot) < \infty$  and  $\mathbb{E}_{\bar{c}}\mathcal{I}(d, \cdot) < \infty$ , we have

$$\mathbb{E}_{\bar{c}}\mathcal{I}(c, \cdot) = \sum_{i=1}^{\infty} a_i \mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, c_i) < \sum_{i=1}^{\infty} a_i \mathbb{E}_{\bar{c}}\mathfrak{d}(v_w, d_i) = \mathbb{E}_{\bar{c}}\mathcal{I}(d, \cdot),$$

which implies that  $\mathbb{E}_{\bar{c}}(\mathcal{I}(c, \cdot) - \mathcal{I}(d, \cdot)) < 0$ . Thus, there is some nonempty set  $E \in \sigma(\mathcal{F})$  with  $\bar{c}(E) > 0$  on which  $\mathcal{I}(c, \cdot) - \mathcal{I}(d, \cdot) < 0$  (since the Lebesgue integral is positive). But this contradicts our assumption that  $d$  weakly dominates  $c$ , and so we are done.  $\square$

**Proposition 4.16.** Let  $(W, \mathcal{F})$  be a point-finite opinion space. If a credence function  $c$  is countably coherent and somewhere finitely inaccurate relative to  $\mathcal{B}$ , then  $c$  is not weakly dominated relative to  $\mathcal{B}$ .

**Proof.** Assume  $d$  weakly dominates  $c$ . Note that i)  $c$  is somewhere finitely inaccurate if and only if  $\mathcal{B}(c, w) < \infty$  for all  $w \in W$  if and only if  $\sum_{i=1}^{\infty} c_i^2 < \infty$ . It follows by weak dominance that  $\mathcal{B}(d, w) < \infty$  for all  $w \in W$ , and therefore  $\sum_{i=1}^{\infty} d_i^2 < \infty$ . Let  $\mathcal{B}(c, w) = \mathfrak{D}(v_w, c)$  for  $\mathfrak{D}$  a generalized quasi-additive Bregman divergence.

Because  $(W, \mathcal{F})$  is point-finite, it is also countably discriminating as there are only countably many finite subsets of  $\mathcal{F}$ . So by Proposition 4.7,  $c = \sum_{j=1}^{\infty} \lambda_j v_{w_j}$  for  $\lambda_j \in [0, 1]$  with  $\sum_{j=1}^{\infty} \lambda_j = 1$  and  $\mathcal{V}_{\mathcal{F}} = \{v_{w_j}\}_{j=1}^{\infty}$ . First, note that  $\mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) = 0$  and  $(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) - c(p_i))^2 = 0$  for all  $i$ , so

$$\mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c\right) - \mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d\right) = \sum_{i=1}^{\infty} a_i \left[ \left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) - c_i\right)^2 - \left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) - d_i\right)^2 \right].$$

<sup>42</sup> In more detail, a special case of Theorem 4.3.5 in Pettigrew (2016) establishes that because  $\mathfrak{d}$  is a one-dimensional Bregman divergence, there is a function  $s: \{0, 1\} \times [0, 1] \rightarrow [0, \infty]$  that satisfies i)  $ps(1, x) + (1 - p)s(0, x)$  is uniquely minimized at  $x = p$  for all  $p \in [0, 1]$ , ii)  $s$  is continuous, and iii)  $s(v_w, x) = \mathfrak{d}(v_w, x)$  for all  $w \in W$  and  $x \in [0, 1]$ .

Using that

$$\begin{aligned} \left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) - c_i\right)^2 - \left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) - d_i\right)^2 &= c_i^2 - d_i^2 + 2(d_i - c_i) \sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) \\ &= \sum_{j=1}^{\infty} \lambda_j c_i^2 - \sum_{j=1}^{\infty} \lambda_j d_i^2 + \sum_{j=1}^{\infty} \lambda_j 2(d_i - c_i) v_{w_j}(p_i) \\ &= \sum_{j=1}^{\infty} \lambda_j [c_i^2 - d_i^2 + 2(d_i - c_i) v_{w_j}(p_i)] \\ &= \sum_{j=1}^{\infty} \lambda_j [(v_{w_j}(p_i) - c_i)^2 - (v_{w_j}(p_i) - d_i)^2] \end{aligned}$$

for each  $i$  because  $\sum_{j=1}^{\infty} \lambda_j = 1$ , we have that

$$\begin{aligned} \mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c\right) - \mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d\right) &= \sum_{i=1}^{\infty} a_i \sum_{j=1}^{\infty} \lambda_j [(v_{w_j}(p_i) - c_i)^2 - (v_{w_j}(p_i) - d_i)^2] \tag{8} \\ &= \sum_{i=1}^{\infty} a_i \left(\sum_{j:w_j \neq p_i} \lambda_j\right) (c_i^2 - d_i^2) + a_i \left(\sum_{j:w_j = p_i} \lambda_j\right) ((1 - c_i)^2 - (1 - d_i)^2) \\ &= \sum_{i=1}^{\infty} a_i (c_i^2 - d_i^2) + 2a_i \left(\sum_{j:w_j = p_i} \lambda_j\right) (d_i - c_i) \\ &= \sum_{i=1}^{\infty} a_i (-c_i^2 - d_i^2) + 2a_i \left(\sum_{j:w_j = p_i} \lambda_j\right) d_i \end{aligned}$$

because  $c_i = \sum_{j:w_j = p_i} \lambda_j$ . We have  $\sum_{i=1}^{\infty} c_i^2 + d_i^2 < \infty$  by item (i). Thus,

$$0 \leq \sum_{i=1}^{\infty} a_i 2 \left(\sum_{j:w_j = p_i} \lambda_j\right) d_i < \infty \tag{9}$$

because

$$0 \geq \mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c\right) - \mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d\right).$$

Having established (9), we can claim that we can use the dominated convergence theorem (see, e.g., Theorem 1.4.49 in Tao 2011) to switch the limits in equation (8). Indeed,

$$\sum_{i=1}^{\infty} a_i \sum_{j=1}^N \lambda_j [(v_{w_j}(p_i) - c_i)^2 - (v_{w_j}(p_i) - d_i)^2] = \sum_{i=1}^{\infty} a_i \left(\sum_{1 \leq j \leq N} \lambda_j\right) (c_i^2 - d_i^2) + 2a_i \left(\sum_{j:w_j = p_i, 1 \leq j \leq N} \lambda_j\right) (d_i - c_i).$$

Letting

$$g_N(i) = a_i \left(\sum_{1 \leq j \leq N} \lambda_j\right) (c_i^2 - d_i^2) + 2a_i \left(\sum_{j:w_j = p_i, 1 \leq j \leq N} \lambda_j\right) (d_i - c_i)$$

and noting that  $-\left(\sum_{j:w_j = p_i, 1 \leq j \leq N} \lambda_j\right) c_i \geq -c_i^2$  because  $c_i = \sum_{j:w_j = p_i} \lambda_j$ , we see that

$$|g_N(i)| \leq a_i (2c_i^2 + d_i^2) + 2 \left(\sum_{j:w_j = p_i} \lambda_j\right) d_i.$$

Each of  $c_i^2, d_i^2$ , and  $(\sum_{j:w_j = p_i} \lambda_j) d_i$  is summable in  $i$  and  $\sup_i a_i < \infty$ . So, the dominated convergence theorem applies, and we can switch limits.

Thus, we have

$$\begin{aligned}
 0 &\geq \mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c\right) - \mathfrak{D}\left(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d\right) \\
 &= \sum_{j=1}^{\infty} \lambda_j \sum_{i=1}^{\infty} a_i [(v_{w_j}(p_i) - c(p_i))^2 - (v_{w_j}(p_i) - d(p_i))^2] \\
 &= \sum_{j=1}^{\infty} \lambda_j (\mathfrak{D}(v_{w_j}, c) - \mathfrak{D}(v_{w_j}, d)) \geq 0,
 \end{aligned}$$

where we used that  $\mathfrak{D}(v_{w_j}, c) = \mathcal{B}(c, w_j) < \infty$  and  $\mathfrak{D}(v_{w_j}, d) = \mathcal{B}(d, w_j) < \infty$  for each  $j$  by i) to break up the summation in the second line. Thus, we can conclude that  $c = d$  because  $\mathfrak{D}(c, d) = 0$  if and only if  $c = d$ . □

**Theorem 4.23.** Let  $(W, \mathcal{F})$  be a  $W$ -stable opinion space with  $(W^*, \mathcal{F}^*)$  point-finite such that all coherent credence functions on  $(W, \mathcal{F})$  are somewhere finitely inaccurate relative to  $\mathcal{B}$ . Then the following are equivalent:

1.  $c$  is coherent;
2.  $c$  is not weakly dominated relative to  $\mathcal{B}$ ;
3.  $c$  is not strongly dominated relative to  $\mathcal{B}$ .

**Proof.** If  $c$  is coherent, then  $c^*$  is countably coherent on a point-finite opinion set. Further,  $c^*$  is somewhere finitely inaccurate relative to  $\mathcal{B}$ , as  $c$  is somewhere finitely inaccurate by assumption. Thus, by Proposition 4.16,  $c^*$  is not weakly dominated relative to  $\mathcal{B}$ . By  $W$ -stability,  $c$  is not weakly dominated relative to  $\mathcal{B}$ . Clearly, if  $c$  is not weakly dominated, then  $c$  is not strongly dominated. Finally, we show that if  $c$  is incoherent, then  $c$  is strongly dominated. First, if  $c$  is not somewhere finitely inaccurate, then any omniscient credence function strongly dominates  $c$  because  $\mathcal{F}(v_w, w') < \infty$  for every  $w, w' \in W$  by point-finiteness. If  $c$  is somewhere finitely inaccurate, then  $\mathcal{F}(c, w) < \infty$  for all  $w \in W$  by point-finiteness. Thus, Theorem 3.4 establishes that  $c$  is strongly dominated relative to  $\mathcal{B}$ . □

**Lemma 4.25.** A countably infinite partition is  $W$ -stable relative to any generalized legitimate inaccuracy measure.

**Proof.** Let  $\mathcal{F} = \{p_1, p_2, \dots\}$  be a partition. Assume that a coherent credence function  $c$  on  $\mathcal{F}$  is weakly dominated by some credence function  $d$ . We can assume  $d$  is coherent by Theorem 3.4, and so  $\sum_{m=1}^{\infty} d_m \leq 1$ . We can show that  $c^*$  is weakly dominated by  $d^*$ , thereby establishing that a partition is  $W$ -stable.

First,  $\mathcal{F}(c^*, w) = \mathcal{F}(c, w)$  and  $\mathcal{F}(d^*, w) = \mathcal{F}(d, w)$  for all  $w \in W$  by Lemma 4.12. Thus, by the assumption of weak dominance,

$$\mathcal{F}(c^*, w) \geq \mathcal{F}(d^*, w) \text{ for all } w \in W$$

with a strict inequality for some  $w \in W$ . We therefore only need to check what happens for  $w \in W^* \setminus W$ . The compactification of a partition consists of adding one point  $w^*$  that is in the complement of all  $p^* \in \mathcal{F}^*$ ; so  $W^* = W \cup w^*$ . If  $\mathcal{F}(c^*, w^*) = \infty$ , then clearly,  $d^*$  weakly dominates  $c^*$ . So, we assume  $\mathcal{F}(c^*, w^*) < \infty$ . Because  $\mathcal{F}(d^*, w^*)$  and  $\mathcal{F}(d^*, w)$  differ by a single term for any  $w \in W$  and  $\mathcal{F}(d^*, w) < \infty$  for some  $w \in W$  by the assumption of weak dominance, we have  $\mathcal{F}(d^*, w^*) < \infty$ . Now, consider

$$\mathcal{F}(c^*, w^*) - \mathcal{F}(d^*, w^*) = \sum_{m=1}^{\infty} a_m (\varphi(d_m) - \varphi'(d_m)d_m) - \sum_{m=1}^{\infty} a_m (\varphi(c_m) - \varphi'(c_m)c_m),$$

which we claim is greater than or equal to 0. Indeed, assume toward a contradiction that

$$\sum_{m=1}^{\infty} a_m(\varphi(d_m) - \varphi'(d_m)d_m) < \sum_{m=1}^{\infty} a_m(\varphi(c_m) - \varphi'(c_m)c_m).$$

Then, because  $d_n \rightarrow 0$  as  $\sum_{n=1}^{\infty} d_n \leq 1$ ,  $c_n \rightarrow 0$  as  $\sum_{n=1}^{\infty} c_n \leq 1$ , and  $\varphi'(0) = \lim_{x \rightarrow 0} \varphi'(x) = 0$  (recall Remark 3.2), we have that  $\varphi'(d_n) - \varphi'(c_n) \rightarrow 0$ ; and so we can find a  $K$  such that

$$|\varphi'(d_n) - \varphi'(c_n)| < \left| \sum_{m=1}^{\infty} a_m(\varphi(d_m) - \varphi'(d_m)d_m) - \sum_{m=1}^{\infty} a_m(\varphi(c_m) - \varphi'(c_m)c_m) \right|$$

for  $n \geq K$ . Thus, for any  $n \geq K$  and any  $w_n \in p_n$ ,

$$\mathcal{I}(c, w_n) - \mathcal{I}(d, w_n) = \sum_{m=1}^{\infty} a_m(\varphi(d_m) - \varphi'(d_m)d_m) - \sum_{m=1}^{\infty} a_m(\varphi(c_m) - \varphi'(c_m)c_m) + \varphi'(d_n) - \varphi'(c_n) < 0,$$

contradicting that  $d$  weakly dominates  $c$ . So indeed,  $d^*$  weakly dominates  $c^*$ . □

**Theorem 4.26.** Let  $(W, \mathcal{F})$  be a countably infinite partition and  $\mathcal{I}$  be a generalized legitimate inaccuracy measure. Then the following are equivalent:

1.  $c$  is coherent;
2.  $c$  is not weakly dominated;
3.  $c$  is not strongly dominated.

**Proof.** The result follows from Corollary 4.22, Lemma 4.25, and the fact that  $\mathcal{I}(c^*, \cdot)$  is bounded on  $W^*$  for each coherent credence function  $c$ . To see the latter, note that because  $c$  is coherent, it follows that  $\sum_{i=1}^{\infty} c_i = \sum_{i=1}^{\infty} c_i^* \leq 1$ . For  $w \in W$  such that  $w \in p_i$ , recalling that  $\varphi(0) = 0$  (Remark 3.2),

$$\mathcal{I}(c^*, w) = a_i \mathfrak{d}(1, c_i^*) + \sum_{j \neq i} a_j \mathfrak{d}(0, c_j^*) = a_i \mathfrak{d}(1, c_i^*) + \sum_{j \neq i} a_j (c_j^* \varphi'(c_j^*) - \varphi(c_j^*)) \leq C + D \sum_j c_j^* \leq C + D$$

for some constants  $C, D$  independent of  $c^*$ . Similarly, as seen in the proof of Lemma 4.25,  $W^* \setminus W = \{w^*\}$ , where

$$\mathcal{I}(c^*, w^*) = \sum_{j=1}^{\infty} a_j \mathfrak{d}(0, c_j^*) = \sum_{j=1}^{\infty} a_j (c_j^* \varphi'(c_j^*) - \varphi(c_j^*)) \leq C$$

for some constant  $C$  independent of  $c^*$  or  $w$ . It follows that i) all coherent credence functions have finite expected inaccuracy, and ii)  $\mathcal{I}(c, w) < \infty$  for  $c \in \mathcal{C}$  and  $w \in W$ . Thus, Lemma 4.25 and Corollary 4.22 establish the result. □

**Proposition 4.27.** Let  $(W, \mathcal{F})$  be a compact opinion space and  $\mathcal{I}$  be a generalized legitimate inaccuracy measure. If  $c$  is coherent (and thus countably coherent), then  $c$  is not strongly dominated relative to  $\mathcal{I}$ .

**Proof.** Let  $\mathcal{I}_n(c', w) := \sum_{i=1}^n a_i \mathfrak{d}(v_w(p_i), c'(p_i))$  for each  $n \in \mathbb{N}$ ,  $w \in W$ , and credence function  $c'$  on  $\mathcal{F}$ . Consider a credence function  $d \neq c$ . Define

$$T^n = \{(v_w(p_1), \dots, v_w(p_n)) : \mathcal{I}_k(c, w) < \mathcal{I}_k(d, w) \text{ for some } k \geq n, w \in W\}$$

and  $T = \{e\} \cup \bigcup_{n=1}^{\infty} T^n$ , where  $e$  is the empty sequence. For each  $s, t \in T$ , we set  $s < t$  if and only if  $s$  is an initial sequence of  $t$ , and we set the height of  $t \in T$  to be the length of the tuple. Then  $T$  is a binary tree.

We claim  $T$  is infinite. Fix  $n \in \mathbb{N}$ . Then there is a  $t \in T$  with height  $n$  if and only if  $T^n \neq \emptyset$  if and only if  $\mathcal{I}_k(c, w) < \mathcal{I}_k(d, w)$  for some  $k \geq n$  and  $w \in W$ . Let  $k$  be the maximum of  $n$  and the smallest  $i$  such that  $c(p_i) \neq d(p_i)$ . Then, because  $c$  restricted to any subset of  $\mathcal{F}$  is coherent, by Theorem 2.9,  $\mathcal{I}_k(c, w') < \mathcal{I}_k(d, w')$  for some  $w' \in W$ , and so  $(v_{w'}(p_1), \dots, v_{w'}(p_n)) \in T^n$ .

By König's lemma (see, e.g., Hrbacek and Jech 1999, sec. 12.3), there exists an infinite branch

$$B = \bigcap_{n=1}^{\infty} \{v_{w_n}(p_1), \dots, v_{w_n}(p_n)\}$$

through  $T$ , where

$$(v_{w_n}(p_1), \dots, v_{w_n}(p_n)) < (v_{w_m}(p_1), \dots, v_{w_m}(p_m))$$

whenever  $n < m$ . For each  $i$ , let  $p_i^* = p_i$  if  $v_{w_i}(p_i) = 1$  and  $p_i^* = p_i^c$  if  $v_{w_i}(p_i) = 0$ . Then  $w_n \in \bigcap_{i=1}^n p_i^*$  because  $v_{w_i}(p_i) = 1$  if and only if  $v_{w_n}(p_i) = 1$  for  $i < n$  as  $(v_{w_i}(p_1), \dots, v_{w_i}(p_i)) < (v_{w_n}(p_1), \dots, v_{w_n}(p_n))$ . Thus,  $\bigcap_{i=1}^n p_i^* \neq \emptyset$  for each  $n$ , and so by compactness, there is some  $w \in \bigcap_{i=1}^{\infty} p_i^*$ . Then

$$(v_w(p_1), \dots, v_w(p_n)) = (v_{w_n}(p_1), \dots, v_{w_n}(p_n)) \in T^n$$

for each  $n \in \mathbb{N}$ . By the definition of  $T^n$ , for each  $n \in \mathbb{N}$ , we have

$$\mathcal{J}_{k_n}(c, w) < \mathcal{J}_{k_n}(d, w)$$

for some  $k_n \geq n$ . Sending  $n$  to infinity,  $\mathcal{J}(c, w) \leq \mathcal{J}(d, w)$ , and thus  $d$  does not strongly dominate  $c$ .  $\square$

### A.3 Proof of Theorem 5.3

**Theorem 5.3.** Let  $\mathcal{J}$  be an integral inaccuracy measure on a finite measure space  $(\mathcal{F}, \mathcal{A}, \mu)$ .<sup>43</sup> Then for every  $\mu$ -credence function  $c$ , if  $c$  is  $\mu$ -incoherent, there is a  $\mu$ -coherent  $\mu$ -credence function  $c'$  that strongly dominates  $c$  relative to  $\mathcal{J}$ .

**Proof.** Let  $\mathcal{J}(c, w) = B_{\phi, \mu}(v_w, c)$ . We write  $B$  for  $B_{\phi, \mu}$ . Let  $S$  be the set of nonnegative  $\mathcal{A}$ -measurable functions on  $\mathcal{F}$ . Let  $E \subseteq S$  be the set of  $\mu$ -coherent  $\mu$ -credence functions over  $\mathcal{F}$ . Then  $E$  is convex. Let  $c$  be a  $\mu$ -incoherent  $\mu$ -credence function. Because  $\mu$  is finite and  $\mathfrak{d}$  is bounded,

$$B(E, c) < \infty.$$

Thus, we can apply Theorem A.5 to get a  $\pi_c \in S$  such that

$$B(s, c) \geq B(E, c) + B(s, \pi_c) \text{ for every } s \in E. \tag{10}$$

In particular, (10) holds when  $s$  is the omniscient credence function at world  $w$  for each  $w$ , so we obtain

$$\mathcal{J}(c, w) \geq B(E, c) + \mathcal{J}(\pi_c, w) \tag{11}$$

for all  $w$ , where all numbers in (11) are finite. We show that  $\pi_c$  is in fact a  $\mu$ -coherent  $\mu$ -credence function. It suffices to show that  $\pi_c$  is  $\mu$ -a.e. equal to a coherent credence function on  $\mathcal{F}$  (because  $\pi_c \in S$ , it is  $\mathcal{A}$ -measurable). To do so, we prove the following claim:  $E$  is closed under loose convergence in  $\mu$ -measure.

To see this, let  $c_n \in E$  for each  $n$  and  $c \in S$ . Assume  $c_n \rightarrow c$  loosely in  $\mu$ -measure. The first thing to notice is that because  $\mu$  is finite, loose  $\mu$ -convergence implies  $\mu$ -a.e. convergence on a subsequence  $\{a_n\}_{n=1}^{\infty}$  of  $\{n\}_{n=1}^{\infty}$ ,<sup>44</sup> so that

$$c(p) = \lim_{n \rightarrow \infty} c_{a_n}(p) \in [0, 1]$$

for each  $p \in \mathcal{G}$  with  $\mu(\mathcal{G}^c) = 0$ . Since the  $c_{a_n}$  are  $\mu$ -coherent, we can change each  $c_{a_n}$  on a (measurable) measure zero set  $\mathcal{X}_n$  to get coherent  $\mu$ -credence functions  $c_{a_n}$ . Further, we replace  $\mathcal{G}$  with  $\mathcal{G} \setminus (\bigcup_{n=1}^{\infty} \mathcal{X}_n)$ .

<sup>43</sup> We assume finiteness for technical reasons.

<sup>44</sup> It is a standard fact that convergence in measure implies a.e. convergence on a subsequence. Now notice that loose convergence implies convergence in measure when the measure is finite.

Assuming these adjustments have been made, we have that  $c_{a_n} \rightarrow c$  on  $\mathcal{G}$  with  $\mu(\mathcal{G}^c) = 0$ , and each  $c_{a_n}$  is coherent. We can now show  $c \in E$  by showing it is equal to a coherent credence function on  $\mathcal{F}$  when restricting to  $\mathcal{G}$ .

First, we extend  $c$  (resp.,  $c_{a_n}$ ) to  $\bar{c}$  (resp.,  $\bar{c}_{a_n}$ ), where  $\bar{c}$  (resp.,  $\bar{c}_{a_n}$ ) is a credence function on  $\mathcal{G} \cup \{W\}$  such that  $c = \bar{c}$  (resp.,  $c_{a_n} = \bar{c}_{a_n}$ ) on  $\mathcal{G}$  and  $\bar{c}(W) = 1$  (resp.,  $\bar{c}_{a_n}(W) = 1$ ). Then notice that  $c$  (resp.,  $c_{a_n}$ ) is coherent on  $\mathcal{G}$  if and only if  $\bar{c}$  (resp.,  $\bar{c}_{a_n}$ ) is coherent on  $\mathcal{G} \cup \{W\}$ . Thus, we work with  $\bar{c}$  and  $\bar{c}_{a_n}$  instead, noting that  $\bar{c} = \lim_n \bar{c}_{a_n}$  on  $\mathcal{G} \cup \{W\}$ . To show  $\bar{c} \in E$ , we first show that  $\bar{c}$  is a partial measure on  $\mathcal{G} \cup \{W\}$ .

Definitions A.9.1 and A.9.3 clearly hold for  $\bar{c}$ , so we just need to show that Definition A.9.2 holds. Let  $\varphi_0, \dots, \varphi_{m-1}, \psi_0, \dots, \psi_{m'-1} \in \mathcal{G} \cup \{W\}$  and

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \subseteq \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i}$$

for every  $k < m$ . Since  $\bar{c}_{a_n}$  are coherent on  $\mathcal{G} \cup \{W\}$  and thus extend to measures on an algebra containing  $\mathcal{G} \cup \{W\}$ , we have by Corollary A.7 that

$$\sum_{k=0}^{m-1} \bar{c}_{a_n}(\varphi_k) = \sum_{k=0}^{m-1} \bar{c}_{a_n} \left( \bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \right) \leq \sum_{k=0}^{m'-1} \bar{c}_{a_n} \left( \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i} \right) = \sum_{k=0}^{m'-1} \bar{c}_{a_n}(\psi_k)$$

using that

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} = \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i} = \emptyset$$

for  $k \geq m'$ . Sending  $n$  to infinity and using the pointwise convergence of  $\bar{c}_{a_n}$  to  $\bar{c}$  on  $\mathcal{G} \cup \{W\}$ , we conclude that

$$\sum_{k=0}^{m-1} \bar{c}(\varphi_k) \leq \sum_{k=0}^{m'-1} \bar{c}(\psi_k).$$

Thus,  $\bar{c}$  is a partial measure on  $\mathcal{G} \cup \{W\}$ . By Theorem A.10, it follows that there is a finitely additive probability function  $c^*$  on  $\mathcal{A}(\mathcal{F})$  such that  $c^* = \bar{c}$  on  $\mathcal{G} \cup \{W\}$ . Thus,  $c^*|_{\mathcal{F}}$  is a coherent credence function on  $\mathcal{F}$ , and

$$c = \bar{c}|_{\mathcal{F}} = c^*|_{\mathcal{F}}$$

$\mu$ -a.e. (specifically off  $\mathcal{G}^c$ ). Further, we already assumed  $c$  is  $\mathcal{A}$ -measurable and  $\{p : c(p) \in [0, 1]\} \subseteq \mathcal{G}$ . Thus,  $c$  is a  $\mu$ -coherent  $\mu$ -credence function.

The proof is finished just as in the proof of Theorem 3.4. By Theorem A.5,  $\pi_c$  is the generalized projection of  $c$  onto  $E$ . Since

$$B(E, c) = \inf_{s \in E} (s, c) < \infty,$$

there is a  $B$ -minimizing sequence  $\{s_n\}$  of elements in  $E$  such that  $B(s_n, c) \rightarrow B(E, c)$  by the definition of infimum. By the definition of a generalized projection,  $s_n \rightsquigarrow_{\mu} \pi_c$ . Since  $E$  is closed under loose convergence, it follows that  $\pi_c \in E$ . Further, because  $c$  is  $\mu$ -incoherent, we know  $c \neq \pi_c$  (up to  $\mu$ -a.e. equivalence), so we see that  $B(E, c) \geq B(\pi_c, c) > 0$  because  $B(s, t) = 0$  if and only if  $s = t$   $\mu$ -a.e. Since  $\mathcal{F}(c, w) < \infty$  for all  $w$ , we deduce that

$$\mathcal{F}(c, w) \geq B(E, c) + \mathcal{F}(\pi_c, w) > \mathcal{F}(\pi_c, w)$$

for all  $w \in W$ . This proves that  $c$  is strongly dominated by  $\pi_c$ , and we are done. □