In this supplement we give proofs of all the main results in the text.

A Structural Causal Models (§2)

A.1 Background on Relations and Orders

Definition A.1.1. Let $C$ be a set. Then a subset $R \subseteq C \times C$ is called a **binary relation** on $C$. We write $c R c'$ if $(c, c') \in R$. The binary relation $R$ is **well-founded** if every nonempty subset $D \subseteq C$ has a minimal element with respect to $R$, i.e., if for every nonempty $D \subseteq C$, there is some $d \in D$, such that there is no $d' \in D$ such that $d' R d$. The binary relation $\prec \subseteq C \times C$ is a (strict) total order if it is irreflexive, transitive, and connected: either $c \prec c'$ or $c' \prec c$ for all $c \neq c' \in C$.

Example 1. The edges of a dag form a well-founded binary relation on its nodes. If $V = \{V_n\}_{n \geq 0}$, then the binary relation $\rightarrow$ defined by $V_m \rightarrow V_n$ iff either $0 < m < n$ or $n = 0 < m$ is well-founded but not extendible to an $\omega$-like total order (see Fact 2) and not locally finite: $V_0$ has infinitely many predecessors $V_1, V_2, \ldots$

A.2 Proofs

Proof of Proposition 1. We assume without loss that $\mathbf{U}(V) = \mathbf{U}$ for every $V \in \mathbf{V}$. For each $u \in \chi_U$, well-founded induction along $\rightarrow$ shows unique existence of a $m^M(u) \in \chi_V$ solving $f_V^*(\pi_{Pa(V)}(m^M(u)), u) = \pi_V(m^M(u))$ for each $V$. We claim the resulting function $m^M$ is measurable. One has a clopen basis of cylinders, so it suffices to show each preimage $(m^M)^{-1}(v)$ is measurable. Recall that here $v$ denotes the cylinder set $\pi_V^{-1}(\{v\}) \in \mathcal{B}(\chi_V)$, for $v \in \chi_V$. Once again this can be established inductively. Note that

$$
(m^M)^{-1}(v) = \bigcup_{p \in \chi_{Pa(V)}} \left[ (m^M)^{-1}(p) \cap \pi_U(f_V^{-1}(\{v\}) \cap (\{p\} \times \chi_U)) \right].
$$

which is a finite union (by local finiteness) of measurable sets (by the inductive hypothesis) and therefore measurable. Thus for any $\mathcal{M}$ the pushforward $p^M = m_\times^M(P)$ is a measure on $\mathcal{B}(\chi_V)$ and gives the observational distribution (Definition 4).

Remark on Definition 6. To see that $p_{\mathcal{M}}^M$ thus defined is a measure, note that $p_{\mathcal{M}}^M = p_{\mathcal{M}}^{\mathcal{M}},$ and apply Proposition 1, where the model $\mathcal{M}_A$ is defined in Definition 4.2.1. This is similar in spirit to the construction of “twinned networks” [2] or “single-world intervention graphs” [8].

Definition A.2.1. Given $\mathcal{M}$ as in Def. [2] and a collection of interventions $A$ form the following counterfactual model $\mathcal{M}_A = (\mathbf{U}, A \times \mathbf{V}, \{f_{(\alpha,V')}(\alpha,V)\}_{(\alpha,V)}, P)$, over endogenous variables $A \times \mathbf{V}$. The counterfactual model has the influence relation $\rightarrow'$, defined as follows. Where $\alpha', \alpha \in A$ let $(\alpha', V') \rightarrow' (\alpha, V)$ iff $\alpha' = \alpha$ and $V' \rightarrow V$. The exogenous space $\mathbf{U}$ and noise distribution $P$ of $\mathcal{M}_A$ are the same as those of $\mathcal{M}$, the exogenous parents sets $\{U(V)\}_V$ are also identical, and the functions are $\{f_{(\alpha,V)}\}_{(\alpha,V)}$ defined as follows. For any $W := w \in A, V \in \mathbf{V}, p \in \chi_{Pa(V)}$, and

\( u \in \chi_{U(V)} \) let

\[
f_{(W:=w, V)}((W:=w, p), u) = \begin{cases} \pi_V(w), & V \in W \\ f_V(p, u), & V \notin W. \end{cases}
\]

**B Proofs from §3**

**Remark on exact characterizations of \( \mathcal{G}_3, \mathcal{G}_2 \).** Rich probabilistic languages interpreted over \( \mathcal{G}_3 \) and \( \mathcal{G}_2 \) were axiomatized in [3]. This axiomatization, along with the atomless restriction, gives an exact characterization for the hierarchy sets. Standard form, defined below, gives an alternative characterization exhibiting each \( \mathcal{G}_3 \) as a particular atomless probability space (Corollary B.1.1). For \( \mathcal{G}_2^{X \rightarrow Y} \) (or \( \mathcal{G}_2 \) in the two-variable case) we need the characterization for the proof of the hierarchy separation result, so it is given explicitly as Lemma [B.3.1] in the section below on 2VE-spaces.

**B.1 Standard Form**

Fix \( \prec \). Note that the map \( \omega_3 \) restricted to \( \mathcal{M}_\prec \) does not inject into \( \mathcal{G}_3 \), as any trivial reparametrizations of exogenous noise are distinguished in \( \mathcal{M}_\prec \). It is therefore useful to identify a “standard” subclass \( \mathcal{M}_\text{std} \) on which \( \omega_3 \) is injective with image \( \mathcal{G}_3 \), and in which we lose no expressivity.

**Notation.** Let \( \text{Pred}(V) = \{ V' : V' \prec V \} \) and denote a deterministic mechanism for \( V \) mapping a valuation of its predecessors to a value as \( f_V \in \chi_{\text{Pred}(V)} \rightarrow \chi_V \). Write an entire collection of such mechanisms, one for each variable, as \( f = \{ f_V \}_V \). A set \( B \subset V \) is ancestrally closed if \( B = \bigcup_{V \in B} \text{Pred}(V) \). For any ancestrally closed \( B \) let \( \xi(B) = \{ (V, p) : V \in B, p \in \chi_{\text{Pred}(V)} \} \). Note that \( f(B) = \chi_{\{ (V, p) \in \xi(B) : f_V \}} \) encodes the set of all possible such collections of deterministic mechanisms, and we write, e.g., \( f \in F(B) \). Abbreviate \( \xi(V), F(V) \) for the entire endogenous variable set \( V \) as \( \xi, F \) respectively. We also use \( f \) to abbreviate the set

\[
\bigcap_{V \in B} \pi^{-1}_{\text{Pred}(V)}(\{ f(p) \}) \in \mathcal{B}(\chi_A \times V)
\]

so we can write, e.g., \( P_f^V(\xi) \) for the probability in \( \mathcal{M} \) that the effective mechanisms \( f \) have been selected (by exogenous factors) for the variables \( B \).

**Definition B.1.1.** The SCM \( \mathcal{M} = \langle U, V, \{ f_V \}_V, P \rangle \) of Def. [B.1] is standard form over \( \prec \), and we write \( \mathcal{M} \in \mathcal{M}_\text{std} \), if we have that \( \rightarrow = \prec \) for its influence relation, \( U = \{ U \} \) for a single exogenous variable \( U \) with \( \chi_U = F, P \in \mathcal{P}(F) \) for its exogenous noise space, and for every \( V \), we have that \( U(V) = U = \{ U \} \) and the mechanism \( f_V \) takes \( p \). \( (f_V)_V \rightarrow f_V(p) \) for each \( p \in \chi_{\text{Pred}(V)} \) and joint collection of deterministic functions \( \{ f_V \}_V \in F = \chi_U \).

Each unit \( u \) in a standard form model amounts to a collection \( \{ f_V \}_V \) of deterministic mechanisms, and each variable is determined by a mechanism specified by the “selector” endogenous variable \( U \).

**Lemma B.1.1.** Let \( \mathcal{M} \in \mathcal{M}_\prec \). Then there exists \( \mathcal{M}_\text{std} \in \mathcal{M}_\text{std} \) such that \( \omega_3(\mathcal{M}) = \omega_3(\mathcal{M}_\text{std}) \).

**Proof.** To give \( \mathcal{M}_\text{std} \) define a measure \( P \in \mathcal{P}(F) \) as in Def. [B.1.1] on a basis of cylinder sets by the counterfactorial in \( \mathcal{M} \)

\[
P_1(\pi^{-1}_{(V_1, p_1)}(\{ v_1 \}) \cap \cdots \cap \pi^{-1}_{(V_n, p_n)}(\{ v_n \})) = P_{\langle f \rangle} \pi^{-1}_{\text{Pred}(V_1):=p_1, V_1}(\{ v_1 \}) \cap \cdots \cap \pi^{-1}_{\text{Pred}(V_n):=p_n, V_n}(\{ v_n \})).
\]

To show that \( \omega_3(\mathcal{M}) = \omega_3(\mathcal{M}_\text{std}) \) it suffices to show that any two models agreeing on all counterfactuals of the form (B.2) must agree on all counterfactuals in \( A \). Suppose \( V_i \in A, v_i \in V_i \in \chi_V \) for \( i = 1, \ldots, n \). Let \( B = \bigcup \text{Pred}(V_i) \) and given \( f = \{ f_V \}_V \), define \( f'_V = w \) to be a constant function mapping to \( \pi^{-1}_V(w) \) if \( V \in W \) and \( f'_V = w \) otherwise. Write \( f = V = v \) if \( f_V = \pi_V(v) = v \) for each \( v \in \chi_V \) such that \( f_V(\pi(V_V)(V)) = \pi_V(v) \) for all \( V \). Finally, note that

\[
\bigcup_{V \in B} f'_V = \bigcup_{\chi_V \times V \in F(B)} \{ f'_V \} = \bigcup_{\chi_V \times V \in F(B)} \{ f'_V \}
\]

for each \( i \).
where each set in the finite disjoint union is of the form \( \|B.1\| \). Thus the measure of the left-hand side can be written as a sum of measures of such sets, which use only counterfactuals of the form \( \|B.2\| \), showing agreement of the measures (by Fact 1).

**Corollary B.1.** \( \mathfrak{S}_3^\prec \) bijects with the set of atomless measures in \( \mathfrak{P}(F) \), which we denote \( \mathfrak{S}_3^\prec_{\text{std}} \). We write the map as \( \varpi^\prec_{\text{std}} : \mathfrak{S}_3^\prec \rightarrow \mathfrak{S}_3^\prec_{\text{std}} \).

Where the order \( \prec \) is clear, the above result permits us to abuse notation, using e.g. \( \mu \) to denote either an element of \( \mathfrak{S}_3^\prec \) or its associated point \( \varpi^\prec_{\text{std}}(\mu) \) in \( \mathfrak{S}_3^\prec_{\text{std}} \). We will henceforth indulge in such abuse.

**Proof of Fact 2** The follows easily from Lem. B.1.2 below, adapted from Suppes and Zanotti [9] Thm. 1. This shows that every atomless distribution is generated by some SCM; furthermore, it can be written as a sum of measures of such sets, which use only counterfactuals of the form \( \|B.2\| \), showing agreement of the measures (by Fact 1).

**Definition B.1.2.** Say that \( \nu \in \mathfrak{P}(F(V)) \) is acausal if \( \nu(\pi^{-1}(\{v_1\}) \cap \pi^{-1}(\{v_2\})) = 0 \) for every \((V, p), (V, p') \in \xi \) and \( v_1 \neq v_2 \in \chi_V \).

**Lemma B.1.2.** Let \( \mu \in \mathfrak{P}(\chi_V) \) be atomless. Then there is a \( M \in \mathfrak{M}_{\text{std}}^\prec \) (see Def. B.1.1) with an acausal noise distribution such that \( \mu = (\omega_1 \circ \omega_2 \circ \omega_3)(M) \).

**Proof.** Consider \( \nu \in \mathfrak{P}(F(V)) = \mathfrak{P}(X_{(V, p)} \chi_V) \) determined on a basis as follows:

\[
\nu(\pi^{-1}_1(\{v_1\}) \cap \cdots \cap \pi^{-1}_{n}(\{v_n\})) = \mu(\pi^{-1}_1(\{v_1\}) \cap \cdots \cap \pi^{-1}_{n}(\{v_n\})).
\]

This is clearly acausal and atomless.

**B.2 Proofs from §3.2**

**Proof of Prop. 3 (Collapse set \( \mathcal{C}_1 \) is empty).** Let \( \mu \in \mathfrak{S}_1 \) and \( \nu \in \mathfrak{S}_3^\prec_{\text{std}} \) with \( (\omega_1 \circ \omega_2 \circ \omega_3)(\nu) = \mu \). By Lemma B.1.2 we may assume \( \nu \) is acausal. Let \( X \) be the first, and \( Y \) the second variable with respect to \( \prec \). Note there are \( x^*, y^* \) such that \( \mu(\pi^{-1}_X(\{x^*\}) \cap \pi^{-1}_Y(\{y^*\})) > 0 \); let \( x^1 \neq x^* \), \( y^1 \neq y^* \). Consider \( \nu' \) defined as follows where \( F_3 \) stands for any set of the form \( \pi^{-1}_{(V_1, p_1)}(\{v_1\}) \cap \cdots \cap \pi^{-1}_{(V_n, p_n)}(\{v_n\}) \subset \mathbb{F}(V) \), for \( V_i \in \mathcal{V}, p_i \in \chi_{\mathcal{P}(V_i)}, v_i \in \chi_{V_i} \), and \( F_1 \) is the corresponding \( \pi^{-1}_{V_1}(\{v_1\}) \cap \cdots \cap \pi^{-1}_{V_n}(\{v_n\}) \subset \chi_V \).

\[
\nu'(\pi^{-1}_{(X, i)}(\{x\}) \cap \pi^{-1}_{(Y, (x^1))}(\{y_i\}) \cap \pi^{-1}_{(Y, (x^0))}(\{y\}) \cap F_3) =
\begin{cases}
\mu(\pi^{-1}_X(\{x^*\}) \cap \pi^{-1}_Y(\{y^*\}) \cap F_1), & x = x^*, y_i = y^* \neq y^1 \\
0, & x = x^*, y_i = y^* \neq y^1 \\
0, & x = x^*, y_i = y^* = y^1 \\
\mu(\pi^{-1}_X(\{x^*\}) \cap \pi^{-1}_Y(\{y^1\}) \cap F_1), & x = x^*, y_i = y^* = y^1 \\
\mu(\pi^{-1}_X(\{x\}) \cap \pi^{-1}_Y(\{y\}) \cap F_1), & x = x^*, y_i = y^* = y^1
\end{cases}
\]

We claim that \( \mu = \mu' \) where \( \mu' = (\omega_1 \circ \omega_2)(\nu') \); it suffices to show agreement on sets of the form \( \pi_X^{-1}(\{x\}) \cap \pi_Y^{-1}(\{y\}) \cap F_1 \). If \( x = x^1 \) then the last case above occurs; if \( x = x^* \) and \( y = y^1 \) then we are in the fourth case; if \( x = x^* \) and \( y = y^* \) then exclusively the first case applies. In all cases the measures agree. Let \( (\nu_\alpha)_\alpha = \omega_2(\nu) \) and \( (\nu'_\alpha)_\alpha = \omega_2(\nu') \) be the Level 2 projections of \( \nu, \nu' \) respectively. Note that \( \nu_{X:=x^1}(y^* \neq y^1) < \nu'_{X:=x^1}(y^1) \). This shows that the standard-form measures \( \nu, \nu' \) project down to different points in \( \mathfrak{S}_2 \) (in particular differing on the \( Y \)-marginal at the index corresponding to the intervention \( X := x^1 \)) while projecting to the same point in \( \mathfrak{S}_1 \). Thus \( \mu \notin \mathcal{C}_1 \) and since \( \mu \) was arbitrary, \( \mathcal{C}_1 = \emptyset \).
for any $\prec' \neq \prec$. It suffices to show that $\mu(\xi) = \mu'(\xi)$ for any $n$ and $f = \{f_V\}_{V \in S_n}$; recall that in the measures, $f$ denotes a set of the form (B.1). Let $(\mu_\alpha)_\alpha = \omega_2(\mu) \in \Theta_2$ and $(\mu_\alpha')_\alpha = \omega_2(\mu')$, with $(\mu_\alpha) = (\mu_\alpha')_\alpha$. Since $\mu_\alpha^{\text{Pred}(V)} \equiv \mu(\pi_V^{-1}(\{\{\}\})) = 1$ for any $V \in S_n \setminus \{X\}$, $p \neq (0,\ldots,0)$, probability bounds show $\mu'(\xi)$ vanishes unless $f_V(p) = 1$ for each such $p$, in which case

$$\mu'(\xi) = \mu'(\bigcap_{i=1}^n \pi_{\{V_i,\{V_i,\ldots,\{V_i,\ldots,0\}}\}(\{v_i\}))$$

for some $v_i \in \chi', \ldots, V_n$, with $V_1 \prec \ldots \prec V_n$. We claim this is reducible—again using probabilistic reasoning alone—to a linear combination of quantities fixed by $(\mu_\alpha)_\alpha$, the Level 2 projection of $\mu'$, which is the same as the projection $(\mu_\alpha)_\alpha$ of $\mu$. This can be seen by an induction on the number $m = |M|$ where $M = \{i : v_i = 1\}$; note (B.3) becomes

$$\mu'(\bigcap_{i \in M} \pi_{(V_i,\{V_i,\ldots,\{V_i,\ldots,0\}}\}(\{0\}))$$

and the inductive hypothesis implies each summand can be written in the sought form while the first term becomes $\mu'(\bigcap_{i \notin M} \pi_{(V_i,\{V_i,\ldots,\{V_i,\ldots,0\}}\}(\{0\})) = \mu'_i\bigcap_{i \notin M} \pi_{(V_i,\{V_i,\ldots,\{V_i,\ldots,0\}}\}(\{0\})) = \mu_i\bigcap_{i \notin M} \pi_{(V_i,\{V_i,\ldots,\{V_i,\ldots,0\}}\}(\{0\}))$. Here () abbreviates the empty intervention $\varnothing := \{\}$. Thus any Level 3 quantity reduces to Level 2, on which the two measures agree by hypothesis.

**B.3 Remarks on §3.3**

**Lemma B.3.1.** Let $(\mu_\alpha)_\alpha \in X_{\alpha \in A_2^X \times Y} \Psi(\chi_X,\chi_Y)$. Then $(\mu_\alpha)_\alpha \in \Theta_2^{X \rightarrow Y}$ iff

$$\mu_X := x(x) = 1$$

for every $x \in \chi_X$ and

$$\mu_X := x(y) \geq \mu_i(x,y)$$

for every $x \in \chi_X$, $y \in \chi_Y$. Here $x, y$ abbreviates the basic set $\pi_x^{-1}(\{x\}) \cap \pi_y^{-1}(\{y\})$.

**Proof.** It is easy to see that (B.4), (B.5) hold for any $(\mu_\alpha)_\alpha$. For the converse, consider the two-variable model over endogenous $Z = \{X, Y\}$ with $X \prec Y$; note that $|F(Z)| = 8$. A result of Tian et al. [10] gives that this model is characterized exactly by (B.4), (B.5), so for any such $(\mu_\alpha)_\alpha$ there is a distribution on $F(Z)$ such that this model induces $(\mu_\alpha)_\alpha$. It is straightforward to extend this distribution to an atomless measure on $F(V)$.

**C Proofs from §4**

**Proof of Prop. 4.** This amounts to the continuity of projections in product spaces and marginalizations in weak convergence spaces. The latter follows easily from results in §3.1.3 of [4] or [3].

**Proof of Thm. 2.** We show how Theorem 3.2.1 of [4] can be applied to derive the result. Specifically, let $\Omega = X_{\alpha \in A_2^X} \chi_Y$. Let $\mathcal{I}$ be the usual clopen basis, and let $W$ be the set of Borel measures $\mu \in \Psi(\Omega)$ that factor as a product $\mu = \times \alpha \mu_\alpha$, where each $\mu_\alpha \in \mathcal{G}_1$ and $(\mu_\alpha)_\alpha \in \Theta_2$. This choice of $W$ corresponds exactly to our notion of experimental verifiability.

It remains to check that a set is open in $W$ iff the associated set is open in $\Theta_2$ (homeomorphism). It suffices to show their convergence notions agree. Suppose $(\nu_\alpha)_\alpha$ is a sequence, each $\nu_\alpha \in W$, converging to $\nu = \times \alpha \mu_\alpha \in W$. We have for each $\alpha$ that $\nu_\alpha = \times \alpha \mu_\alpha \alpha$ such that $(\mu_\alpha)_\alpha \in \Theta_2$. By Theorem 3.1.4 in [4], which is straightforwardly generalized to the infinite product, for each fixed $\alpha$ we have $(\mu_\alpha)_\alpha \Rightarrow \nu_\alpha$. This is exactly pointwise convergence in the product space $\Theta_2$, and the same argument in reverse works for the converse.
D Proofs from §5

We will use the following result to categorize sets in the weak topology.

**Lemma D.0.1.** If \( X \subset \partial \) is a basic clopen, the map \( p_X : (\mathfrak{S}, \tau^w) \to ([0, 1], \tau) \) sending \( \mu \mapsto \mu(X) \) is continuous and open (in its image), where \( \tau \) is as usual on \([0, 1] \subset \mathbb{R}\).

**Proof.** Continuous: the preimage of the basic open \((r_1, r_2) \cap p_X(\mathfrak{S})\) where \(r_1, r_2 \in \mathbb{Q}\) is \(\{\mu : \mu(X) > r_1\} \cap \{\mu : \mu(X) < r_2\} = \{\mu : \mu(X) > r_1\} \cap \{\mu : \mu(\partial \setminus X) > 1 - r_2\}\), a finite intersection of the subbasic sets \(\{\mathfrak{S}\} \) from \([1]\). See also Kechris [6, Corollary 17.21].

Open: if \( X = \varnothing \) or \( \partial \), then \( p_X(\mathfrak{S}) = \{0\} \) or \( \{1\} \) resp., both open in themselves. Else \( p_X(\mathfrak{S}) = [0, 1] \); we show any \( Z \) is continuous and open. Consider a mutually disjoint, covering \( D = \{\bigcap_{i=0}^n Y_i : Y_0 \in \{X, \partial \setminus X\}, \text{ each } Y_i \in \{X_i, \partial \setminus X_i\}\} \) and space \( \Delta = \{(\mu(D))_{D \in D} : \mu \in \mathfrak{S} \}\subset \mathbb{R}^{2^{n+1}} \). Just as in the Lemma, we have \( \mathfrak{p}_S : \Delta \to [0, 1] \), for each \( S \subset D \) taking \( (\mu(D))_D \mapsto \sum_{D \in S} \mu(D) \). Note \( Z = \mathfrak{p}_{\{D : D \cap X = \varnothing\}}(\bigcap_{i=1}^n \mathfrak{p}_{\{D : D \cap X = \varnothing\}}((r_i, 1])) \) so it suffices to show \( \mathfrak{p}_S \) is continuous and open; this is straightforward.

**Full proof of Lem. [2]** We show a stronger result, namely that the complement of the good set is nowhere dense. By rearrangement and laws of probability we find that the second inequality in (2) is equivalent to

\[
\mu_x(y') < \mu_1(x') + \mu_0(x, y') \\
1 - \mu_x(y) < \mu_1(x') + \mu_0(x, y') - \mu_0(x, y)
\]

\[
\mu_x(y) > \mu_0(x, y).
\]

Lemma [B.3.1] then entails the non-strict analogues of all four inequalities in (2). (3) are met for any \( (\mu_\alpha)_\alpha \in \mathfrak{S}_X \to Y \), so we show that converting each to an equality yields a nowhere dense set, whose finite union is also nowhere dense. Note that we have a continuous and surjective observational projection \( \pi_0 : \mathfrak{S}_X \to \mathfrak{S}(X, Y) \), and the first inequality in (3) is met iff \( (\mu_\alpha)_\alpha \in (p_{x,y} \circ \pi_0)^{-1}(\{0\}) \) where \( p_{x,y} \) is the map from Lemma [D.0.1] and \( x, y \) denotes the set \( \pi_0^{-1}(\{x\}) \cap \pi_Y^{-1}(\{y\}) \subset \chi(X,Y) \). This is nowhere dense as it is the preimage of the nowhere dense set \( \{0\} \subset [0, 1] \) under a map which is continuous by Lemma [D.0.1]. The second inequality in (3) is wholly analogous after rearrangement.

As for (2), define a function \( d : \mathfrak{S}_X \to [0, 1] \) taking \( (\mu_\alpha)_\alpha \mapsto \mu_{X=x}(y') - \mu_{X=x}(y') \); this function \( d \) is continuous by Lemma [D.0.1] and the continuity of addition and projection. Note that the first inequality of (2) holds iff \( d(\mu_\alpha)_\alpha = 0 \). For any \( \mu \in \mathfrak{S}_X \) such that \( (\pi_0 \circ \pi_Y)(\mu) = (\mu_\alpha)_\alpha \), note that \( d((\mu_\alpha)_\alpha) = \mu(x', y'_z) \) where \( x', y'_z \) abbreviates the basic set \( \pi_0^{-1}(\{x'\}) \cap \pi_Y^{-1}(\{y'\}) \) is \( \mathcal{B}(X \times Y) \). Thus \( d \) is surjective, so that \( d^{-1}(\{0\}) \) is nowhere dense since \( \{0\} \subset [0, 1] \) is nowhere dense.

**Proof of Lem. [2]** Abbreviate \( \mu_3 \) as \( \mu \), and without loss take \( \mu \in \mathfrak{S}_\text{std} \). Note that (2), (3) entail

\[
0 < \mu(x', y'_z) < \mu(x'), \quad 0 < \mu(x', y'_x) < \mu(x').
\]

and therefore

\[
0 < \mu(\pi_0^{-1}(\{x'\}) \cap \pi_Y^{-1}(\{y'\})) < \mu(\pi_0^{-1}(\{x'\})\{1\})
\]

for each \( x' \in \chi X = \{0, 1\} \). In turn this entails that there are some values \( y_0, y_1 \in \{0, 1\} \) such that \( \mu(\Omega_1) > 0, \mu(\Omega_2) > 0 \) where the disjoint sets \( \{\Omega_i\}_i \) are defined as

\[
\Omega_1 = \pi_0^{-1}(\{x'\}) \cap \pi_Y^{-1}(\{y_0\}) \cap \pi_Y^{-1}(\{y_1\})
\]

\[
\Omega_2 = \pi_0^{-1}(\{x'\}) \cap \pi_Y^{-1}(\{y_0\}) \cap \pi_Y^{-1}(\{y_1\})
\]

where in the second line, \( y'_0 = 1 - y_0 \) and \( y'_1 = 1 - y_1 \). Note that for \( i = 1, 2 \) we have conditional measures \( \mu_i(S_i) = \frac{\mu(S_i)}{\mu(\Omega_i)} \) for \( S_i \in \mathcal{B}(\Omega_i) \); further, \( \Omega_i \) is Polish, since each is clopen. This implies
We claim that applying (D.7) and (D.1), (D.6) becomes
\[
\mu(g(X_1)) = \frac{\mu(\Omega_2)}{\mu(\Omega_1)} \mu(X_1).
\]
Consider \( \mu' = \varpi_3(M') \) for a new \( M' \in \mathcal{M}_\prec \), given as follows. Its exogenous valuation space is \( \chi_U = \Omega' \) where we define the sample space \( \Omega' = F(V) \times \{ T, H \} \); that is, a new exogenous variable representing a coin flip is added to some representation of the choice of deterministic standard form mechanisms. Fix constants \( \varepsilon_1, \varepsilon_2 \in (0, 1) \) with \( \varepsilon_1 \cdot \mu(\Omega_1) = \varepsilon_2 \cdot \mu(\Omega_2) \) and define its exogenous noise distribution \( P \) by
\[
P(X \times \{ S \}) = \begin{cases} (1 - \varepsilon_1) \cdot \mu(X), & X \subseteq \Omega_1, S = T \\ \varepsilon_1 \cdot \mu(X), & X \subseteq \Omega_1, S = H \\ (1 - \varepsilon_2) \cdot \mu(X), & X \subseteq \Omega_2, S = T \\ \varepsilon_2 \cdot \mu(X), & X \subseteq \Omega_2, S = H \\ \mu(X), & X \subseteq F(V) \setminus (\Omega_1 \cup \Omega_2), S = T \\ 0, & X \subseteq F(V) \setminus (\Omega_1 \cup \Omega_2), S = H \end{cases}
\]
(\[D.2\])

Where \( f \in F(V) \) and \( V \in V \) write \( f_V \) for the deterministic mechanism (of signature \( \chi_{\text{Pred}(V)} \to \chi_V \)) for \( V \) in \( f \). (Note that each \( f \) is just an indexed collection of such mechanisms \( f_V \).) The function \( f'_V \) in \( M' \) is defined at the initial variable \( X \) as \( f'_X(f, S) = f_X \) for both values of \( S \), and for \( V \neq X \) is defined as follows, where \( p \in \text{Pred}(V) \):
\[
f'_V(p, (f, S)) = \begin{cases} (g(f))_V(p), & f \in \Omega_1, S = H, \pi_X(p) = x \\ (g^{-1}(f))_V(p), & f \in \Omega_2, S = H, \pi_X(p) = x \\ f_V(p), & \text{otherwise} \end{cases}
\]
(\[D.3\])

We claim that \( \varpi_2(\mu') = \varpi_2(\mu) \). It suffices to show for any \( Z := z \in A \) and \( w \in \chi_W \), \( W \) finite, we have
\[
\mu(\theta) = \mu'(\theta), \text{ where } \theta = \bigcap_{W \in W} \pi_z^{-1}(\{ \pi_W(w) \}).
\]
(\[D.4\])

Assume \( \pi_Z(w) = \pi_Z(z) \) for every \( Z \in Z \cap W \), since both sides of (\[D.4\]) trivially vanish otherwise. Where \( f \in F(V) \) write, e.g., \( f \vdash \theta \) if \( m^{M'}(f) \vdash \theta \), where \( M' \) is a standard form model (Def. [B.1.1]) for \( \omega' \in \Omega' \) write \( \omega' \vdash \theta \) if \( m^{M'(\omega')} \vdash \theta \). By the last two cases of (\[D.3\]) we have
\[
\mu'(\theta) = \sum_{S = T, H} P(\{(f, S) \in \Omega': (f, S) \vdash \theta\})
\]
\[
= \mu(\{(f, S) \in F(V) \setminus (\Omega_1 \cup \Omega_2) : f \vdash \theta\}) + \sum_{S = T, H} \sum_{i=1,2} P(\{(f, S) \in \Omega_i : f \in \Omega_i, (f, S) \vdash \theta\}).
\]
(\[D.5\])

Applying the first four cases of (\[D.2\]) and the third case of (\[D.3\]), the second term of (\[D.3\]) becomes
\[
\sum_i \varepsilon_i \cdot \mu(\{(f, H) \vdash \theta\}) + (1 - \varepsilon_i) \cdot \mu(\{(f, H) \vdash \theta\})
\]
(\[D.6\])

Either \( X \in Z \) and \( \pi_X(z) = x \), or not. In the former case: defining \( X_i = \{ f \in \Omega_i : f \vdash \theta \} \) for each \( i = 1, 2 \), the first two cases of (\[D.3\]) yield that
\[
\{ f \in \Omega_1 : (f, H) \vdash \theta \} = \{ f \in \Omega_1 : g(f) \vdash \theta \} = g^{-1}(X_2)
\]
\[
\{ f \in \Omega_2 : (f, H) \vdash \theta \} = \{ f \in \Omega_2 : g(f) \vdash \theta \} = g(X_1).
\]
(\[D.7\])

Applying (\[D.7\]) and (\[D.1\]), (\[D.6\]) becomes
\[
\varepsilon_1 \cdot \mu(\Omega_1) \cdot \mu(X_2) + (1 - \varepsilon_1) \cdot \mu(X_1) + \varepsilon_2 \cdot \mu(\Omega_2) \cdot \mu(X_1) + (1 - \varepsilon_2) \cdot \mu(X_2)
\]
\[
= \mu(X_1) + \mu(X_2),
\]
(\[D.8\])
the final cancellation by choice of \( \varepsilon_1, \varepsilon_2 \). In the latter case: since \( n:^{\text{dir}}(x') \in \pi_i^{-1}(\{x'\}) \) for any \( x' \in \Omega_1 \cup \Omega_2 \), the third case of (D.3) gives \( \{ f \in \Omega : (f, H) \models' \theta \} = X_i. \) Thus (D.6) becomes (D.8) in either case. Putting in (D.8) as the second term in (D.5), we find \( \mu(\theta) = \mu'(\theta) \).

Now we claim \( \mu(\zeta) \neq \mu'(\zeta) \) for \( \zeta = \zeta_0 \cap \zeta_1 \) where \( \zeta_1 = \pi_{1,1.1}'(\{y_1\}) \) and \( \zeta_0 = \pi_{1,0,1}'(\{y_0\}) \).

We have
\[
\mu'(\zeta) = \mu(\{ f \in \Omega \setminus (\Omega_1 \cup \Omega_2) : f \models' \zeta \}) + \sum_{i=1,2} \varepsilon_i \cdot \mu(\{ f \in \Omega_i : (f, H) \models' \zeta \}) + (1 - \varepsilon_i) \cdot \mu(\{ f \in \Omega_i : f \models' \zeta \}).
\]

(D.9)

First suppose that \( x = 0 \). If \( f \in \Omega_1 \), then note that \( (f, H) \models' \zeta_0 \) iff \( g(f) \models \zeta_0 \), but this is never so, since \( g(f) \in \Omega_2 \). If \( f \in \Omega_2 \), then \( (f, H) \models' \zeta_1 \) iff \( f \models \zeta_1 \), which is never so again by choice of \( \Omega_2 \). If \( x = 1 \) then we find that \( (f, H) \not\models' \zeta_1 \) (if \( f \in \Omega_1 \)) and \( (f, H) \not\models' \zeta_0 \) (if \( f \in \Omega_2 \)). Thus \( (f, H) \not\models' \zeta \) for any \( f \in \Omega_1 \cup \Omega_2 \) and (D.9) becomes
\[
\mu(\{ f \in \Omega : f \models' \zeta \}) - \sum_{i=1,2} \varepsilon_i \cdot \mu(\{ f \in \Omega_i : f \models' \zeta \}) = \mu(\{ f \in \Omega : f \models' \zeta \}) - \varepsilon_1 \cdot \mu(\Omega_1) < \mu(\zeta).
\]

It is straightforward to check (via casework on the values \( y_0, y_1 \)) that \( \mu \) and \( \mu' \) disagree also on the PNS; \( \mu(y_x, y'_x) \neq \mu'(y_x, y'_x) \) as well as its converse. As for the probability of sufficiency (Definition 10), note that
\[
P(y_x \mid x', y') = \frac{P(y_x, x', y'_x) + P(y_x, y'_x, x')}{P(x', y')}
\]

and it is again easily seen (given the definition of the \( \Omega_i \)) that \( \mu(y_x, x', y'_x) \neq \mu'(y_x, x', y'_x) \) while the two measures agree on the denominator; similar reasoning shows disagreement on the probability of enablement, since
\[
P(y_x \mid y') = \frac{P(y_x, y'_x, x') + P(y_x, y'_x, x)}{P(y')}.
\]

\[\square\]

References


