Efficient Mechanisms under Unawareness*

Kym Pram† Burkhard C. Schipper‡

Preliminary & Incomplete: March 21, 2023
Please do not circulate yet.

Abstract

We study the design of mechanisms under asymmetric awareness and asymmetric information. With limited awareness, an agent's message space is type-dependent because an agent cannot misrepresent herself as a type that she is unaware of. Nevertheless, we show that the revelation principle holds. The revelation principle is of limited use though because a mechanism designer is hardly able to commit to outcomes for type profiles of which he is unaware. Yet, the mechanism designer can at least commit to properties of social choice functions like efficiency given ex post awareness. Assuming quasi-linear utilities, private values, and welfare isotonicity in awareness, we show that if a social choice function is utilitarian ex post efficient, then it is implementable under pooled agents' awareness in conditional dominant strategies. That is, it is possible to reveal all asymmetric awareness among agents and implement the welfare maximizing social choice function in conditional dominant strategies without the need of the social planner being fully aware ex ante. To this end, we develop dynamic versions of the Groves and Clarke mechanisms along which true types are revealed and subsequently elaborated at endogenous higher awareness levels. We explore how asymmetric awareness affects budget balance and participation constraints.

Keywords: dynamic mechanism design, dominant strategy implementation, Vickrey-Clarke-Groves mechanisms, utilitarian ex post efficiency, unknown unknowns.

JEL Classification Numbers: D83

*We thank Sarah Auster, Tomasz Sadzik, and Joel Watson for useful discussions. Burkhard gratefully acknowledges financial support via ARO Contract W911NF2210282.

†University of Nevada, Reno. Email: kpram@unr.edu

‡University of California, Davis. Email: bcschipper@ucdavis.edu

1 Introduction

Mechanism design studies the design of institutions governing collective decisions such as markets or political systems in the presence of asymmetric information. However, agents may not just face asymmetric information but also asymmetric awareness. For instance, agents may be unaware of some events affecting preferences or endowments of others. They may also be unaware of events affecting costs of producing private or public goods. Consequently, they may form beliefs only about events that they are aware of. Unawareness refers to the lack of conception rather than the lack of information.

Central to mechanism design is the revelation principle. At the first glance, designing "optimal" institutions seems to be daunting task as we would have search over all possible institutions no matter how complicated. Fortunately, in standard mechanism design, the revelation principle allows us to focus w.l.o.g. on mechanisms in which agents report directly their type to the institution. Is the revelation principle still valid when agents have asymmetric awareness. With limited awareness, an agent's message space is type-dependent because an agent cannot misrepresent herself as a type that she is unaware of. Starting with Green and Laffont (1986), there is a literature showing the failure of the revelation principle with type-dependent message sets (see also Bull and Watson, 2007, Deneckere and Severinov, 2008, and Strausz, 2017). Nevertheless, we show that the revelation principle holds under unawareness if the mechanism design can commit to social choice functions. This is because Green and Laffont's (1986) Nested Range Condition always holds under unawareness.

The mechanism designer herself may be unaware of events. This poses at least two challenges: First, how could the mechanism designer take incentive compatibility constraints into account for types of which he is unaware? Second, how could the mechanism designer commit to outcomes for type profiles of which he is unaware? It is well known that revelation principle may fail if the mechanism designer cannot commit to the mechanism (e.g., Bester and Strausz, 2001). Yet, the mechanism designer can at least commit to properties of outcome functions like efficiency given ex post awareness. In order to analyze the implementation of efficient social choice functions, we restrict to a prominent class of quasilinear utilities with private values. We generalize the notion of utilitarian ex post efficiency of social choice functions to allow for asymmetric awareness. We also assume that utilitarian welfare is isotone in awareness. At a philosophical level, this assumption reflects some Fortschrittsglaube (Meek Lange, 2011) that more awareness is better for society. This assumption is consistent with the utilitarian ideas embodied in economics in general and the idea of utilitarian ex post efficiency in mechanism design in particular. For instance, John Stuart Mill singled out the growth in awareness and knowledge of mankind as the predominant force of social progress.[1]

[1]Mill (1868, pp. 523-525) writes "Now, the evidence of history and that of human nature combine, by a striking instance of consilience, to show that there really is one social element which is thus predominant, and almost paramount, among the agents of the social progression. This is, the state of the speculative faculties of mankind; including the nature of the beliefs which by any means they have arrived at, concerning themselves and the world by which they are surrounded. ... Thus (to take the

We desire to implement efficient social choice functions at the *highest awareness level possible*. The problem is that no agent or the mechanism designer might be aware of everything. However, we seek to at least pool awareness of all agents and aim to implement the social choice that would be utilitarian ex post efficient at this pooled awareness level. For instance, consider a social choice function that prescribes to each illness the best possible treatment available at that awareness level. Before becoming aware of antibiotics, this would for instance involve prescribing mercury for syphilis while this is clearly not the best choice once being aware of the toxic properties of mercury and the availability of modern antibiotics. So clearly, it is desirable to implement outcomes at the highest awareness level possible in society, which is the awareness level that pools awareness of all agents. Again, this desideratum reflects the Fortschrittsglaube. It requires us to consider dynamic mechanisms in which agents do not only report their type (and awareness) to the mechanism designer/mediator but also the mechanism designer then disseminates awareness among agents and consequently changes and unifies the awareness levels of agents in an endogenous fashion. We achieve this with a dynamic direct elaboration mechanism. In the initial stage, agents report their type given their awareness. After that, their awareness may be raised when the mechanism designer provides feedback about the pooled awareness level. At this point, agents have the opportunity to elaborate on their previously reported type at the higher awareness level. This process of elaborations by agents and subsequent communication of pooled awareness by the mediator continuous until no agents wants to elaborate any further at which point the mechanism stops and the outcome is implemented.

Since we consider dynamic mechanisms, we require a solution concept to dynamic games. We aim to retain the belief-freeness of dominant strategy implementation as it allows us to be silent on complicated updates of probabilistic beliefs throughout the dynamic mechanism in the dace of raising awareness. Nevertheless, we require dynamically optimal elaborations throughout the dynamic mechanism in which agents report optimally conditional on the awareness communicated back to them by the mediator. To wit, we make use of *conditional* dominant strategies. That is, we focus on strategies that are dominant conditional on the state and every information set reached in the dynamic mechanism. The idea of conditional dominance is known from game theory (Shimoji and Watson, 1998) and has been applied as solution concept to games with unawareness in extensive form (Meier and Schipper, 2022). It is precisely the changes of awareness in the dynamic direct elaboration mechanism that necessitates the use of conditional dominance rather than just dominance.[2]

most obvious case first,) the impelling force to most of the improvements effected in the arts of life, is the desire of increased material comfort; but as we can only act upon external objects in proportion to our knowledge of them, the state of knowledge at any time is the limit of the industrial improvements possible at that time; and the progress of industry must follow, and depend on, the progress of knowledge. ... we are justified in concluding, that the order of human progression in all respects will mainly depend on the order of progression in the intellectual convictions of mankind, that is, on the law of the successive transformations of human opinions."

[2]E.g., Brandenburger and Friedenberg (2011) show that iterated extensive-form admissibility coincides with iterated admissibility in the associated strategic form in standard games without unawareness.

We show that if a social choice function is utilitarian ex post efficient, then it is implementable under pooled agents' awareness in conditional dominant strategies in a dynamic direct elaboration mechanism. That is, it is possible to reveal all asymmetric awareness among agents and implement the welfare maximizing social choice function in conditional dominant strategies without the need of the social planner/mechanism designer/mediator being fully aware ex ante. We consider several versions of dynamic direct elaboration mechanisms. Our first mechanism can be viewed as a dynamic version of the Groves mechanism (Groves, 1973) as it uses almost the same transfer functions as in the Groves mechanism. Recall that transfers of a Groves mechanism typically consist of two terms: The sum of opponents' utilities and a term that depends on opponent's reports only. In our dynamic Groves mechanism, this latter terms can only depend on *initial* reports of opponents. Since in our dynamic mechanism, awareness contained in initial reports gets broadcasted to other agents who subsequently can elaborate on their prior reports, restricting the second term of transfers to initial reports avoids that the second term of the Groves transfers indirectly depend on the agent's own reports.

The last mentioned subtlety highlights also a similarity to mechanism design with interdependent valuations. In some sense, in our setting each agent's valuation depends on the awareness of all other agents. While we know from Jehiel and Moldovanu (2001) that it is impossible to implement ex post efficient social choice functions in settings with interdependent valuations, we are able to implement ex post efficient social choice functions in our setting because awareness creates just a particular dependency that is mitigated by two features: First, an agent can lie only with what she is aware of, which means she can only pretend to be less aware. Second, we assume utilitarian ex post isotony in awareness. More awareness is better for society. While these two features make lots of sense in the context of unawareness but it is unclear what they would represent in a standard setting with interdependent valuations.

Next, we investigate properties beyond efficiency. We show that if our efficient dynamic direct elaboration Groves mechanism satisfies budget balance, then utilitarian welfare must be constant in awareness. This motivates us to look beyond our dynamic Groves mechanisms. In particular, because one of the terms of our transfer function can only depend on initial reports, our dynamic Groves mechanisms do not permit for the analogue of the pivot mechanism à la Clarke (1971). We define efficient dynamic direct elaboration Clarke mechanisms. I this mechanism, the constant term of the Groves transfers is now allowed to depend on final reports of other players. Yet, we add an additional adjustment term that incentivizes reporting of awareness and is budget neutral. Under the standard assumption of positive valued benefit functions, we show that this dynamic direct elaboration Clarke mechanism implements utilitarian ex post efficient social choice functions under pooled awareness in conditional dominant strategies with no deficit. The downside is that an agent who is aware of their unawareness may not want to participate

In contrast, for games with unawareness, Meier and Schipper (2022) show that conditioning on information sets is crucial since it implies conditioning on the awareness captured by the information set. Although dominant strategy implementation uses very weak dominance rather than weak dominance, it is still the case that in games with unawareness conditioning on awareness is crucial.

in the mechanism for the fear of being penalized when others raise awareness more than herself does.[3] ...

Since mechanism design is closely related to contract theory, our paper contributes to the recent literature on contracting under unawareness (e.g., Lee, 2008, van Thadden and Zhao, 2012, Auster, 2013, Filiz-Ozbay, 2012, Grant, Kline, and Quiggin, 2012, Auster and Pavoni, 2021, Lei and Zhao, 2021, Francetich and Schipper, 2022). For instance, Francetich and Schipper (2022) show that screening contracts may not provide sufficient incentives to agents to reveal their awareness. Consequently, the principal may not be able to consider the full set of incentive constraints. In contrast, we show that we can go beyond incomplete contracts and reveal awareness in dynamic direct elaboration mechanisms that provide sufficient incentives for truthful reporting. The literature on unawareness in contracting is very different from the earlier literature on indescribable contingencies in contracting. For instance, in Maskin and Tirole (1999), agents are fully aware of all payoff consequences but cannot describe them for some reason ex ante. In contrast, agents may not be fully aware of all payoff consequences in contracting under unawareness.

Our paper was inspired by Herweg and Schmidt (2020) who study a procurement problem with a principal and two agents. The agents may be aware of some design flaws. Herweg and Schmidt (2020) propose an efficient two-stage mechanism. In the first stage, the agents can reveal potential design flaws. In a second stage, the project is adjusted and awarded to exactly one agent via a reverse auction. Our paper differs in many respects: First, in Herweg and Schmidt (2020), agent's private costs are independent of the design flaw (i.e., fixing the design flaw requires a known common cost) in Herweg and Schmidt (2020), while in our model the relevant space of payoff types depend on awareness. Second, Herweg and Schmidt (2020) construct an efficient direct mechanism under common awareness and only allow for asymmetric awareness in an indirect mechanism that achieves the same allocation as their efficient direct mechanism while in our case we allow for asymmetric awareness in our dynamic direct elaboration mechanism. We believe this difference may be due to the fact that in Herweg and Schmidt (2020) agents report design flaws whose specific payoff consequence need to be verified. In our case, agents just raise awareness (like of the potential existence of design flaws) but the uncertainty over the payoff consequences of the contingencies that agents became aware are captured in the resulting payoff type space. Third, Herweg and Schmidt (2020) focus on ex post implementation while we focus on conditional dominant strategy implementation. Typically under private values, ex post implementation is equivalent to dominant strategy implementation. Yet, Herweg and Schmidt (2020, Endnote 20) remark that it is not the case in their model as they use the Nash bargaining solution in renegotiations when design flaws are revealed at a later stage. Finally, we consider more generally an abstract mechanism design problem rather than the particular application to procurement.

Our general setting makes use of unawareness type spaces introduced in Heifetz,

[3] Awareness of unawareness refers to a state of mind in which the agent considers it possible that she is unaware of something without being able to point to what she is unaware; see Schipper (2022).

Meier, and Schipper (2013a), which are probabilistic analogues of unawareness structures (Heifetz, Meier, and Schipper, 2006), and games with unawareness (Heifetz, Meier, and Schipper, 2013b, Meier and Schipper, 2014).

The paper is organized as follows: In the next two sections, we introduced unawareness type spaces and unawareness of payoff-types, respectively. In Section 4, we discuss the revelation principle under unawareness. Ex post efficiency under unawareness is introduced in Section 5. This is followed by an exposition of dynamic direct elaboration mechanisms in Section 6. We prove the conditional dominant strategy implementation of utilitarian ex post efficient social choice functions under pooled awareness with dynamic direct elaboration Groves mechanisms in Section 7. The dynamic direct elaboration Clarke mechanism is studied in Section 8.

2 Unawareness Type Spaces

Consider unawareness type spaces à la Heifetz, Meier, and Schipper (2013a). Denote by $\langle \mathbf{S}, \succeq \rangle$ the nonempty finite lattice of nonempty measurable disjoint state-spaces. For any collection $S_1, ..., S_\ell \in \mathbf{S}$ of spaces, we let $\bigvee_{k=1}^{\ell} S_k$ denote the join. The join and meet always exists since any finite lattice is complete. We denote by $\bar{S} := \bigvee_{S \in \mathbf{S}} S$ and $\underline{S} := \bigwedge_{S \in \mathbf{S}} S$ the join and meet of the entire lattice $\mathbf{S}$.

For any $S \in \mathbf{S}$, we denote by $\mathcal{F}(S)$ the σ-field associated with S. Let $\Omega := \bigcup_{S \in \mathbf{S}} S$ be the union of all state spaces. For any $S, S' \in \mathbf{S}$ with $S' \succeq S$, there is a measurable surjective projection $r_S^{S'} : S' \longrightarrow S$ for which r_S^S is the identity for any $S \in \mathbf{S}$. Moreover, projections commute, i.e., for any $S, S', S'' \in \mathbf{S}$ with $S'' \succeq S' \succeq S$, we have $r_S^{S''} = r_S^{S'} \circ r_{S'}^{S''}$.

For a subset of states $D \subseteq S$, for some space $S \in \mathbf{S}$, denote by $D^\uparrow := \bigcup_{S' \succeq S} \left(r_S^{S'}\right)^{-1}(D)$. An *event* has now the form $E = D^\uparrow$ with $D \subseteq S$, for some $S \in \mathbf{S}$. D is called the *base* of the event E and S the *base-space* of the event E denoted by $S(E)$. If $E \neq \emptyset$, then S is uniquely determined by E. Otherwise, we write $\emptyset^S$ for the vacuous event that is based in space S. To understand this, note that the empty set is a subset of any state space. When we take the empty subset of a state space, we can consider also the union of its inverse images in more expressive spaces, which of course is empty as well. This is a vacuous event. But all these vacuous events are different because they have different base-spaces. While this may look strange at first, it makes perfect sense. A vacuous event corresponds to a contradiction, a description that is contradictory like "the sun is shining and the sun is not shining". There is no state of the world where this is true. However, contradictions can be more or less rich depending on how rich is the language with which they are described. The richness of the language is implicitly specified with the base-space, and that's why we must have different vacuous events. We denote by $\mathcal{E}$ the set of all measurable events, i.e., all events $D^\uparrow$ such that $D \in \mathcal{F}_S$ for some $S \in \mathbf{S}$.

For any state space $S \in \mathbf{S}$, let $\Delta(S)$ be the set of probability measures on $(S, \mathcal{F}_S)$. We consider this set itself as a measurable space endowed with the σ-field $\mathcal{F}_{\Delta(S)}$ generated by the sets $\{\mu \in \Delta(S) : \mu(D) \geq p\}$, where $D \in \mathcal{F}_S$ and $p \in [0, 1]$. In order model beliefs at different levels of awareness, we need to relate probability measures on a richer space to probability measures on poorer spaces. Formally, for a probability measure $\mu \in \Delta(S')$, the marginal $\mu_{|S}$ of μ on $S \preceq S'$ is defined by

$$\mu_{|S}(D) := \mu\left(\left(r_S^{S'}\right)^{-1}(D)\right), \quad D \in \mathcal{F}_S.$$

To extend probability measures to events of our lattice structure, let S_μ denote the space on which μ is a probability measure. Whenever for some event $E \in \mathcal{E}$ we have $S_\mu \succeq S(E)$ (i.e., the event E can be expressed in space S_μ) then we abuse notation slightly and write

$$\mu(E) = \mu(E \cap S_\mu).$$

If $S(E) \npreceq S_\mu$ (i.e, the event E is not expressible in the space S_μ because either S_μ is strictly poorer than $S(E)$ or S_μ and $S(E)$ are incomparable), then we say that $\mu(E)$ is undefined.

There is a finite set of agent I. For each agent $i \in I$ there is a *belief mapping* $\beta_i : \Omega \longrightarrow \bigcup_{S \in \mathbf{S}} \Delta(S)$ which is measurable in the sense that for every $S \in \mathbf{S}$ and $Q \in \mathcal{F}_{\Delta(S)}$ we have $\beta_i^{-1}(Q) \cap S \in \mathcal{F}_S$. We impose the following properties:

(i) *Confinement:* If $\omega \in S'$ then $\beta_i(\omega) \in \Delta(S)$ for some $S \preceq S'$.

(ii) If $S'' \succeq S' \succeq S$, $S'', S', S \in \mathbf{S}$, $\omega \in S''$, and $\beta_i(\omega) \in \Delta(S)$ then $\beta_i(\omega_S) = \beta_i(\omega)_{|S}$.

(iii) If $S'' \succeq S' \succeq S$, $S'', S', S \in \mathbf{S}$,, $\omega \in S''$, and $\beta_i(\omega_{S'}) \in \Delta(S)$ then $S_{\beta_i(\omega)} \succeq S$.

(iv) Introspection: For every $\omega \in \Omega$, $\left\{\omega' \in \Omega : \beta_i(\omega')_{|S_{\beta_i(\omega)}} = \beta_i(\omega)\right\} \subseteq E$ (for any measurable event E) implies $\beta_i(\omega)(E) = 1$.

We denote by $\langle \langle \mathbf{S}, \succeq \rangle, (r_S^{S'})_{S,S' \in \mathbf{S}, S' \succeq S}, (\beta_i)_{i \in I} \rangle$ the unawareness type space. See Heifetz, Meier, and Schipper (2013a) for further details.

3 Payoff Environment

For each agent $i \in I$, there is a nonempty complete lattice $\mathbf{T}_i$ of disjoint measurable payoff type spaces, one payoff type space T_i^S for each state-space $S \in \mathbf{S}$. Denote by $\Theta_i := \bigcup_{S \in \mathbf{S}} T_i^S$. There is a payoff type map $\theta_i : \Omega \longrightarrow \Theta_i$ such that $\omega \in S$ implies $\theta_i(\omega) \in T_i^S$. We require that for any $S \in \mathbf{S}$, $\theta_{i|S} : S \longrightarrow T_i^S$ is measurable and surjective. That is, the restriction of θ_i to S is a measurable surjective function to T_i^S. We require that each agent is certain of her own payoff type. I.e., for $\omega', \omega'' \in S_{\beta_i(\omega)}$ such that $\beta_i(\omega'') = \beta_i(\omega')$ implies $\theta_i(\omega'') = \theta_i(\omega')$.

For any $S, S' \in \mathbf{S}$ with $S' \succeq S$, we define a measurable surjective projection $(\rho_i)_S^{S'} : T_i^{S'} \longrightarrow T_i^S$ by $(\rho_i)_S^{S'}(\theta_i(\omega)) = \theta_i(r_S^{S'}(\omega))$. That is, we require that the following diagram commutes:

$$\begin{array}{ccccc} & & \theta_{i|S'} & & \\ & T_i^{S'} & \longleftarrow & S' & \\ (\rho_i)_S^{S'} & \downarrow & & \downarrow & r_S^{S'} \\ & T_i^S & \longleftarrow & S & \\ & & \theta_{i|S} & & \end{array}$$

Denote by $\rho_S^{S'}((t_i)_{i \in I}) := ((\rho_i)_S^{S'}(t_i))_{i \in I}$, $\theta(\omega) := (\theta_i(\omega))_{i \in I}$, and $T^S := \times_{i \in I} T_i^S$.

As for $\mathbf{S}$, for any $D \subseteq T_i^S$, $S \in \mathbf{S}$, $i \in I$, we let $D^{\uparrow} := \bigcup_{S' \succeq S} (\rho_S^{S'})_i^{-1}(D)$.

States in state spaces of the lattice of spaces in $\mathbf{S}$ model the state of the mind of agents. Types in spaces of the lattice of payoff type spaces $\mathbf{T}_i$ just models payoff-relevant events. The payoff type map θ_i "pulls out" from any state of a state space in $\mathbf{S}$ the payoff type of agent i.

For each space $S \in \mathbf{S}$, there is a measurable nonempty set of outcomes X^S. We require $S' \succeq S$ implies $X^S \subseteq X^{S'}$. This formulation allows for unawareness of outcomes. Denote by $X := \bigcup_{S \in \mathbf{S}} X^S$.

Each agent has a utility function $u_i : \bigcup_{S \in \mathbf{S}} X^S \times T_i^S \longrightarrow \mathbb{R}$. From this formulation it is clear that we focus on private values and that types in payoff types spaces of $\mathbf{T}_i$ model conjunctions of payoff relevant events.

Define $\Theta := \times_{i \in I} \Theta_i$. Note that typically $\bigcup_{S \in \mathbf{S}} T^S \subsetneq \Theta$ unless in special cases.

We consider social choice functions $f : \Theta \longrightarrow X$. We require that $\boldsymbol{t} = (t_i)_{i \in I} \in \times_{i \in I} T_i^{S_i}$ for some $S_i \in \mathbf{S}$, $i \in I$, implies $f(\boldsymbol{t}) \in X^{\bigwedge_{i \in I} S_i}$. This assumption ensures that every agent is able to evaluate the outcome.[4]

4 Revelation Principle

A *mechanism* $\langle (A_i)_{i \in I}, g \rangle$ consists of

- a nonempty set of actions A_i for each agent $i \in I$, and
- an outcome function $g : \times_{i \in I} A_i \longrightarrow X$.

Given an unawareness type space $\langle \langle \mathbf{S}, \succeq \rangle, (r_S^{S'})_{S, S' \in \mathbf{S}, S' \succeq S}, (\beta_i)_{i \in I} \rangle$, agents may not be aware of all actions. We model awareness of actions of each agent $i \in I$ by an action correspondence $\mathcal{A}_i : \Omega \longrightarrow 2^{A_i} \setminus \{\emptyset\}$ such that for any nonempty subset of actions $A'_i \subseteq A_i$, $[A'_i] := \{\omega \in \Omega : A'_i \subseteq \mathcal{A}_i(\omega)\}$ is an event in the unawareness type space and

[4]As we will see, this restriction does not preclude us from aiming to implement outcomes under pooled awareness, i.e., elements in $X^{\bigvee_{i \in I} S_i}$. It just necessitates dynamic mechanisms with endogenous awareness in order to implement outcomes under pooled awareness.

$\omega', \omega'' \in [\beta_i(\omega)] \cap S_{\beta_i(\omega)}$ implies $\mathcal{A}_i(\omega') = \mathcal{A}_i(\omega'')$ for all $\omega \in \Omega$. That is, agents may be unaware of actions but each type of an agent is certain of her own actions that she is aware of.

A mechanism $\langle (A_i)_{i \in I}, g \rangle$, unawareness type space $\langle \langle \mathbf{S}, \succeq \rangle, (r_S^{S'})_{S,S' \in \mathbf{S}, S' \succeq S}, (\beta_i)_{i \in I} \rangle$, action correspondences $(\mathcal{A}_i)_{i \in I}$, and payoff environment $\langle (T_i^S)_{S \in \mathbf{S}, i \in I}, ((\rho_i)_S^{S'})_{S,S' \in \mathbf{S}, S' \succeq S, i \in I}, (\theta_i)_{i \in I}, (X^S)_{S \in \mathbf{S}}, (u_i)_{i \in I} \rangle$ induce a Bayesian game with unawareness (see for instance, Meier and Schipper, 2014) in which agent i's payoff is given by $u_i(g(a_1, ..., a_{|I|}), \theta_i(\omega))$ in state ω and action profile $(a_1, ..., a_{|I|})$.

A (pure) *strategy* for agent i is a mapping $\sigma_i : \Omega \longrightarrow A_i$ such that

- $\sigma_i(\omega) \in \mathcal{A}_i(\omega_{\beta_i(\omega)})$,
- $\beta_i(\omega) = \beta_i(\omega')$ implies $\sigma_i(\omega) = \sigma_i(\omega')$.

That is, each agent can only take actions she is aware of. Moreover, she can condition her action on her type.

Let Σ_i denote agent i's set of strategies.

A dominant strategy equilibrium of the Bayesian game with unawareness induced by the mechanism is a profile of strategies $\sigma = (\sigma_i)_{i \in I} \in \Sigma := \times_{i \in I} \Sigma_i$ such that for all $i \in I$ and any $\omega \in \Omega$,

$$u_i(g(\sigma_i(\omega), a_{-i}), \theta_i(\omega_{S_{\beta_i(\omega)}})) \geq u_i(g(a_i, a_{-i}), \theta_i(\omega_{S_{\beta_i(\omega)}}))$$

for all $a_i \in \mathcal{A}_i(\omega_{\beta_i(\omega)})$ and $a_{-i} \in \bigcup_{\omega' \in S_{\beta_i(\omega)}} \times_{j \neq i} \mathcal{A}_j(\omega')$. That is, for any player and every state, the equilibrium action very weakly dominates other actions

A *mechanism implements* a social choice function f in dominant strategies if there exists a dominant strategy equilibrium $\sigma = (\sigma_i)_{i \in I}$ such that $g(\sigma_1(\omega), ..., \sigma_{|I|}(\omega)) = f(\theta_1(\omega), ..., \theta_{|I|}(\omega))$ for all $\omega \in \Omega$. A social choice function f is weak dominant strategy implementable if there exists a mechanism that implements it in weak dominant strategies.

A *direct revelation mechanism* $\langle (\Theta_i)_{i \in I}, d \rangle$ is a mechanism such that

- for each $i \in I$, the set of payoff types, Θ_i, is her set of actions,
- $d : \times_{i \in I} \Theta_i \longrightarrow X$.

Given an unawareness type space $\langle \langle \mathbf{S}, \succeq \rangle, (r_S^{S'})_{S,S' \in \mathbf{S}, S' \succeq S}, (\beta_i)_{i \in I} \rangle$ and payoff environment $\langle (T_i^S)_{S \in \mathbf{S}, i \in I}, ((\rho_i)_S^{S'})_{S,S' \in \mathbf{S}, S' \succeq S, i \in I}, (\theta_i)_{i \in I}, (X^S)_{S \in \mathbf{S}}, (u_i)_{i \in I} \rangle$, each agent may not always be aware of all of her payoff types. For each agent $i \in I$, we model this with a correspondence $\mathcal{T}_i : \Omega \longrightarrow 2^{\Theta_i} \setminus \{\emptyset\}$ such that $\mathcal{T}_i(\omega) := \bigcup_{S \preceq S_\omega} T_i^S$. This is a special case of the action correspondence of Bayesian games with unawareness. A type of an agent can never report a type that she is unaware of.

A social choice function f is *truthfully dominant strategy implementable* if the Bayesian game with unawareness induced by the direct revelation mechanism has a dominant strategy equilibrium $\sigma = (\sigma_i)_{i \in I}$ for which $\sigma_i(\omega) = \theta_i(\omega)$ for all $\omega \in \Omega$ and all $i \in I$.

Theorem 1 *Suppose the social choice function f is implementable in dominant strategies. Then f is truthfully implementable in dominant strategies.*

PROOF. Since f is implementable in weak dominant strategies, there exists a mechanism $\langle (A_i)_{i \in I}, g \rangle$, unawareness type space $\langle \langle \mathbf{S}, \succeq \rangle, (r_S^{S'})_{S, S' \in \mathbf{S}, S' \succeq S}, (\beta_i)_{i \in I} \rangle$, action correspondences $\langle \mathcal{A}_i \rangle_{i \in I}$, and payoff environment $\langle (T_i^S)_{S \in \mathbf{S}, i \in I}, ((\rho_i)_S^{S'})_{S, S' \in \mathbf{S}, S' \succeq S, i \in I}, (\theta_i)_{i \in I}, (X^S)_{S \in \mathbf{S}}, (u_i)_{i \in I} \rangle$ that induce a Bayesian game with unawareness for which there is a dominant strategy equilibrium σ such that the diagram commutes:

$$\begin{array}{ccc} \Omega & \xrightarrow{\theta} & \Theta \\ \sigma \downarrow & & \downarrow f \\ A & \xrightarrow{g} & X \end{array}$$

From the diagram follows now that there exists a direct mechanism $\langle (\Theta_i)_{i \in I}, f \rangle$ such that for the Bayesian game with unawareness induced by the mechanism, the unawareness type space, payoff environment, and correspondences $(\mathcal{T}^S)_{S \in \mathbf{S}, i \in I}$ we have $(\theta_i \circ r_{S_{\beta_i(\cdot)}}^{S_{(\cdot)}})_{i \in I}$ is a dominant strategy equilibriume. Thus, f is truthfully implementable in weak dominant strategies. □

Green and Laffont (1986) observed that in settings with verifiable information, the revelation principle breaks down. That is, if agents have type dependent messages available, the messages themselves become partial proof of their type (see Bull and Watson, 2007, Denecker and Severinov, 2008, and Strausz, 2017, for related work). In mechanisms with unawareness, sets of actions are type dependent as well. Similarly, in direct mechanisms with unawareness, reportable types are type-dependent because a type can only report types that she is aware of. While reporting awareness is not exactly verifiable information in a literal sense, it is clearly not cheap talk as the message may not be available to all types. In particular, if a type is unaware of another type, then she cannot pretend to be that latter type by reporting it. Why does the revelation principle hold under unawareness but is violated in examples with verifiable information by Green and Laffont (1986)? Green and Laffont (1986) identify a condition called Nested Range Condition which characterizes settings with verifiable information in which the revelation principle does hold. (Bull and Watson, 2007, and Deneckere and Severinov, 2008, present related conditions.)

To define the condition, for any payoff type $t_i \in \Theta_i$, let $M_i(t_i)$ denote the set of types that type t_i could report in principle. The setting satisfies the *Nested Range Condition* if for any $t_i, t_i', t_i'' \in \Theta_i$, we have $t_i' \in M_i(t_i)$ and $t_i'' \in M_i(t_i')$ implies $t_i'' \in M_i(t_i)$.

Proposition 1 *Any mechanism under unawareness satisfies the Nested Range Condition.*

Proof. For any agent $i \in I$ and state $\omega \in \Omega$, $\mathcal{T}_i(\omega_{\beta_i(\omega)})$ is the set of types of which type $\beta_i(\omega)$ is aware of. The Nested Range Condition now follows from properties (i) to (iii) of the type mapping, the definition of the payoff type mapping θ_i, and the fact that the lattice order is transitive. A similar argument holds for any indirect mechanism, using correspondence $\mathcal{A}_i$ in place of $\mathcal{T}_i$. □

The intuition is simple: An agent can only report types that she is aware of. If a type t_i is aware of another type t_i' and that type is aware of yet another type t_i'', then former type t_i must also be aware of t_i''. So the transitivity of awareness levels as captured by the lattice order of spaces yields automatically the Nested Range Condition of Green and Laffont (1986).

While the revelation principle continues to hold unawareness, it is a kind of misleading. How exactly would an unaware mechanism designer commit to an outcome that she might be unaware of for a profile of types that she might be unaware of? The implicit assumption must be that for the revelation principle to hold, the mechanism design must be aware of everything. This is highly unrealistic in many settings and avoids the interesting problem of revealing awareness from agents. Consider instead a setting in which the mechanism designer may be unaware of some events. What happens if actions are played or types are reported that the mechanism designer had been unaware of ex ante? Wouldn't this invite tearing up the proposed ex ante mechanism specified for types that the mechanism designer has been aware of or wouldn't agents at least worry that this could happen if they report their types truthfully in a way that raises awareness of the mechanism designer and other players?

Despite being unaware of particular actions and types ex ante and thus being unaware to commit to a mechanism, the mechanism designer is still able to commit to properties of the mechanism. For instance, the mechanism designer could commit that no matter what awareness is raised, an efficient solution will be implemented. Studying such a problem goes immediately beyond the revelation principle and pertains to which social choice functions can be implemented under unawareness. This question is attacked in the next section.

Another problem with the static revelation principle that the outcome of the mechanism may allow agents to become aware of events that they have been unaware of ex ante. With their raised awareness, they may now want to take an action different from the action they have taken before. Or in the direct mechanism, they may want to elaborate on their previously reported type, i.e., provide a more fine-grained description of their type or report a different type altogether. Ideally, the mechanism designer should encourage pooling awareness among agents and take the agents' ability to provide a more fine-grained descriptions of their types with pooled awareness into account when implementing the social choice function. Allowing for implementation under pooled awareness requires us to go beyond static mechanisms. This is the topic of the next sections.

5 Ex Post Efficiency

To facilitate a commonly used notion of efficiency, we assume that each agent's utility function is quasilinear. I.e., for each $i \in I$, $u_i(x, t_i) := v_i(x_0, t_i) + x_i$ for $x = (x_0, x_1, ..., x_{|I|}) \in X_0^S \times \mathbb{R}^{|I|}$ and $t_i \in T_i^S$, $S \in \mathcal{S}$. As usual, x_0 describes the physical properties of the outcome while $(x_i)_{i \in I}$ represents the vector of transfers made *to* agents.

Utility functions are payoff type-dependent. Since payoff-types also encode awareness, utilities may differ for corresponding types across awareness levels. While we do not want to impose a restriction on how individual utilities vary with awareness, we believe that at least for society as a whole a restriction on utilitarian welfare across awareness levels could be defended. We assume that utilitarian ex post welfare is isotone in awareness. That is, welfare should not decrease when awareness of some agents is raised and the welfare maximizing outcome is adjusted accordingly. Behind this assumption is some Fortschrittsglaube. The idea that more awareness allows for welfare enhancing choices does not necessarily apply to each individual but to the society as whole. To define this notion, we introduce some notation.

For any $i \in I$, $S \in \mathbf{S}$, and $t_i \in T_i^S$, let $[t_i] := \bigcup_{S' \succeq S}((\rho_i)_S^{S'})^{-1}(t_i)$. Since θ_i is measurable, we have $\theta_i^{-1}([t_i]) \in \Sigma$ is an event. That is, $[t_i]$ is the event that agent i has payoff type t_i. It makes sense to write $S([t_i])$ for the base-space of the event $[t_i]$. This is the lowest space in which this type is expressible. For any profile of payoff types, $\boldsymbol{t} \in \Theta$, we denote by $\check{S}([\boldsymbol{t}]) = \bigvee_{i \in I} S([t_i])$ and by $\hat{S}([\boldsymbol{t}]) = \bigwedge_{i \in I} S([t_i])$. Furthermore, we $S \in \mathbf{S}$, let $\mathbf{S}(S) := \{S' \in \mathbf{S} : S' \preceq S\}$ be the sublattice of $\mathbf{S}$ with join S. This sublattice is relevant because we need to model which states an agent considers. When an agent's belief is on space S, then she can reason about events in S and in all less expressive spaces. These are events in spaces $\mathbf{S}(S)$.

Utilitarian ex post welfare is isotone in awareness if for all $\boldsymbol{t} = (t_i)_{i \in I} \in \Theta$,

$$\max_{x_0 \in X^{\hat{S}([\boldsymbol{t}])}} \sum_{i \in I} v_i(x_0, t_i) \geq \max_{x_0 \in X^{\hat{S}([\boldsymbol{t'}])}} \sum_{i \in I} v_i(x_0, t_i') \tag{1}$$

for all $\boldsymbol{t'} = (t_i')_{i \in I}$ with $t_i' = \left(\rho_{S_i}^{S([t_i])}\right)_i (t_i)$ for some $S_i \in \mathbf{S}(S([t_i]))$ for every $i \in I$. From now on we assume that utilitarian ex post welfare is isotone in awareness.

Denote by $f_0(\boldsymbol{t})$ the projection of $f(\boldsymbol{t})$ on X_0. That is, $f_0(\boldsymbol{t})$ is the physical outcome prescribed by the social choice function f at the type profile $\boldsymbol{t}$. We require that for any $\boldsymbol{t} \in \Theta$, $f_0(\boldsymbol{t}) \in X_0^{\hat{S}([\boldsymbol{t}])}$ so that every agent's utility for the outcome is defined.

We generalize utilitarian ex post efficiency of social choice functions to asymmetric unawareness as follows: A social choice function f is *utilitarian ex post efficient* if for all $\boldsymbol{t} = (t_i)_{i \in I} \in \Theta$,

$$\sum_{i \in I} v_i(f_0(\boldsymbol{t}), t_i) \geq \sum_{i \in I} v_i(x_0, t_i) \tag{2}$$

for all $x_0 \in X_0^{\hat{S}([\boldsymbol{t}])}$. This definition generalizes utilitarian ex post efficiency by allowing for several awareness levels. Observe that it implies as a special case the utilitarian ex post efficiency of f for each space $S \in \mathbf{S}$. I.e., for every $S \in \mathbf{S}$, $\boldsymbol{t} = (t_i)_{i \in I} \in \times_{i \in I} T_i^S$, $\sum_{i \in I} v_i(f_0(\boldsymbol{t}), t_i) \geq \sum_{i \in I} v_i(x_0, t_i)$ for all $x_0 \in X_0^S$.

Given that we assume utilitarian ex post welfare isotony in awareness, utilitarian ex post efficiency is characterized by a stronger property that we will make use of in proofs.

Remark 1 *The social choice function f is utilitarian ex post efficient if and only if for all $\boldsymbol{t} = (t_i)_{i \in I} \in \Theta$,*

$$\sum_{i \in I} v_i(f_0(\boldsymbol{t}), t_i) \geq \sum_{i \in I} v_i\left(x_0, \left(\rho_{S_i}^{S([t_i])}\right)_i (t_i)\right) \tag{3}$$

for all $x_0 \in X_0^{\bigwedge_{i \in I} S_i}$ and $S_i \in \mathbf{S}(S([t_i]))$, $i \in I$.

Proof. "Only if:" Since the social choice function f is utilitarian ex post efficient, for all $\boldsymbol{t} = (t_i)_{i \in I} \in \Theta$,

$$\sum_{i \in I} v_i(f_0(\boldsymbol{t}), t_i) \geq \sum_{i \in I} v_i(x_0, t_i)$$

for all $x_0 \in X_0^{\hat{S}([\boldsymbol{t}])}$, and

$$\sum_{i \in I} v_i(f_0(\boldsymbol{t'}), t'_i) \geq \sum_{i \in I} v_i(x_0, t'_i)$$

for any $\boldsymbol{t'} = (t'_i) \in \Theta$ with $t'_i = \left(\rho_{S_i}^{S([t_i])}\right)_i (t_i)$ for some $S_i \in \mathbf{S}(S([t_i]))$ for all $i \in I$ and for all $x_0 \in X_0^{\hat{S}([\boldsymbol{t'}])}$.

Since f is utilitarian ex post efficient, utilitarian ex post welfare isotony in awareness implies

$$\sum_{i \in I} v_i(f_0(\boldsymbol{t}), t_i) \geq \sum_{i \in I} v_i(f_0(\boldsymbol{t'}), t'_i).$$

Thus,

$$\sum_{i \in I} v_i(f_0(\boldsymbol{t}), t_i) \geq \sum_{i \in I} v_i(x_0, t'_i)$$

which is precisely inequality (3).

The converse is obvious. □

6 Dynamic Direct Elaboration Mechanisms

We introduce a new class of dynamic mechanisms. Fix a social choice function f. The *dynamic direct elaboration mechanism* implementing f is defined inductively by the following algorithm:

- At period $n = 1$, each agent $i \in I$ must report a type t_i^1.
- At period $n > 1$ and for each agent $i \in I$, if $S([t_i^{n-1}]) \prec \check{S}([\boldsymbol{t}^{n-1}])$, then agent i must report a type $t_i^n \in ([t_i^{n-1}] \cap T_i^{\check{S}([\boldsymbol{t}^{n-1}])})^{\uparrow}$. Otherwise, $t_i^n = t_i^{n-1}$.
- If $t_i^{n+1} = t_i^n$ for all $i \in I$, then $f(\boldsymbol{t}^{n+1})$ is implemented.

Note awareness and information is transmitted both from agents to the mechanism/mediator but also from the mechanism/mediator to the agents. Thus, awareness of agents may change endogenously when interacting in the mechanism. Initially, the agent reports a type that she is aware of. Given the reports by all agents, the mediator/mechanism designer computes the pooled awareness level based on reports in the first period, $\check{S}([\boldsymbol{t}^1])$. In the next period, the mediator asks each agent whose reported awareness level was lower than the pooled one to elaborate on her report by picking a type consistent with her previously reported type but at an awareness level that is at least as expressive as the pooled one. For that reason, one may call this a "direct elaboration" mechanism.

Note that $\check{S}([\boldsymbol{t}^{n-1}])$ may be more expressive than $S([t_i^{n-1}])$ for any $i \in I$.

The mechanism stops when no agent wants to further elaborate on her type. Clearly, if $\mathbf{S}$ is finite, then the mechanism stops at some finite period n. Moreover, when it stops at n, then $\boldsymbol{t}^n \in \times_{i \in I} T_i^{\check{S}([\boldsymbol{t}^n])}$. That is, all agents must have reported twice in a row types at the *same* awareness level.

The dynamic direct mechanism induces a game with unawareness in extensive form akin to Heifetz, Meier, and Schipper (2013b). A strategy of a player assigns to each information set of game with unawareness in extensive form, a report of a type within the set of types allowed by the dynamic direct mechanism. Since awareness can only increase along any path in the game and reported types become more and more elaborate, it is sufficient to denote information sets of a player i just by the most recent report and the join space associated with the most recent reports by all agents. That is, the information set of agent i in period n is $h_i^{n-1} = (t_i^{n-1}, S([\boldsymbol{t}^{n-1}]))$ if $n > 1$ and $h_i^{n-1} = \emptyset$ if $n = 1$.

When writing $h_i^n = (t_i^n, S)$, we implicitly consider an equivalence class of information sets of agent i in which agent i reported t_i^n in period n and received feedback S from the mediator. In principle, the particular path of awareness feedback provided by the mediator may matter for updating of beliefs. Yet, since we will focus on conditional dominant strategy implementation, we do not need to specify in detail how exactly beliefs are updated based on information and awareness received from the mediator throughout the game. It is enough to just specify some features of the awareness update. Write

$S(h_i^n) = S$ if $h_i^n = (t_i, S)$ and $n > 1$ and $S(h^1) = \underline{S}$ (the meet of the lattice $\mathbf{S}$). By construction of the dynamic direct mechanism, $S(h_i^n) \succeq S([t_i^n])$ because the mediator may be able to raise the agent's awareness. More importantly, $S(h_i^n) = S(h_j^n)$ for any $i, j \in I$ and $n \geq 1$. That is, there is common feedback received from the mediator. We do not require private or targeted communication from the mediator to the agent.

Given the unawareness type space $\langle\langle \mathbf{S}, \succeq\rangle, (r_S^{S'})_{S,S' \in \mathbf{S}, S' \succeq S}, (\beta_i)_{i \in I}\rangle$ and a space $S \in \mathbf{S}$, the S-*update* is an unawareness type space $\langle\langle \mathbf{S}, \succeq\rangle, (r_S^{S'})_{S,S' \in \mathbf{S}, S' \succeq S}, (\beta_i(\cdot \mid S))_{i \in I}\rangle$ that satisfies for all $i \in I$:

(1) For all $\omega \in S'$ with $S' \succeq S \vee S_{\beta_i(\omega)}$ implies $S_{\beta_i(\omega \mid S)} = S \vee S_{\beta_i(\omega)}$.

(2) For all $\omega, \omega' \in \Omega$ and $S \in \mathbf{S}$, $\beta_i(\omega' \mid S) = \beta_i(\omega \mid S)$ implies $\beta_i(\omega') = \beta_i(\omega)$.

Property (1) says that in a S-update, for every state in a space more expressive than S, every player has at least awareness level S. Property (2) says that the S-updated belief type does not forget the initial belief type. Note that the S-update of an unawareness type space does not pin down uniquely belief revision as this is not required for our result since we will focus on conditional dominant strategy implementation.[5] It should be clear that conditions and (1) and (2) are consistent with properties (0) to (iv).

In games with unawareness in extensive form, players cannot necessarily foresee all information sets. Define $H_i(S) := \left\{(t_i, S') : t_i \in \bigcup_{S'' \preceq S'} T_i^{S'}, S' \in \mathbf{S}(S)\right\} \cup \{\emptyset\}$. This is the set of agent i's information sets that an agent with awareness level S would be aware of. For any $\omega \in \Omega$, denote by S_ω the space $S \in \mathbf{S}$ for which $\omega \in S$.

A (pure) *strategy* of agent i in the game induced by the dynamic direct elaboration mechanism is a mapping $\sigma_i : \bigcup_{\omega \in \Omega} \{\omega\} \times H_i(S_\omega) \longrightarrow \Theta_i$ such that for $\omega \in \Omega$,

(0) $\sigma_i(\omega, \emptyset) \in \bigcup_{S \preceq S_{\beta_i(\omega)}} T_i^S$,

(i) for all $(t_i, S) \in H_i(S_\omega)$, $\sigma_i(\omega, (t_i, S)) \in \bigcup_{S_{\beta_i(\omega \mid S)} \succeq S' \succeq S} ((\rho_i)_{S([t_i])}^{S'})^{-1}(t_i)$,

(ii) for any $\omega, \omega' \in \Omega$ and $h \in H_i(S_\omega) \cap H_i(S_{\omega'})$, $\beta_i(\omega \mid S(h)) = \beta_i(\omega' \mid S(h))$ implies $\sigma_i(\omega, h) = \sigma_i(\omega', h)$.

Property (i) is a constraint imposed by the dynamic direct elaboration mechanism as it allows only elaborations of the previously reported type. An agent can only provide elaborations of her payoff type that she is aware of herself at the state and history. Note that we have automatically $t_i^n = t_i^{n-1}$ if $S([t_i^{n-1}]) = \check{S}([\boldsymbol{t}^{n-1}])$. Both Properties (0) and (i) imply that an agent can only report a type that she is aware of at the respective state

[5] See Francetich and Schipper (2022) for a screening problem under unawareness in which more properties are imposed on beliefs upon becoming aware. In that paper, it is assumed that belief systems satisfy reverse Bayesianism and logconcavity. Logconcavity implies monotone inverse hazard rates that play an important role in optimal mechanism design. Monotone inverse hazard rates are then preserved upon becoming aware by reverse Bayesianism. Since we focus on efficient mechanism design and conditional dominance, such properties are not needed in our present setting.

and information set. Property (ii) is just the standard requirement that strategies are measurable w.r.t. information (and awareness). At any two states and feasible history at those states, if her beliefs and awareness are the same in those states given the history, then her action must also be the same. For our purpose, it is sufficient to allow just for a notion of strategy that is akin to "stationary Markov" strategies in dynamic games as both the time period and the particular path will not matter for behavior.

For $i \in I$, we now denote by Σ_i the set of strategies of agent i in the game induced by the dynamic direct elaboration mechanism. We write $\sigma_{-i}(\omega, h_{-i}) := \prod_{j \neq i} \sigma_j(\omega, h_j)$ and $\Sigma_{-i} := \times_{j \neq i} \Sigma_j$.

For agent i, the *truth-telling and elaboration strategy* is defined by for $\omega \in \Omega$ and $h \in H_i(S_\omega)$, $\sigma_i^*(\omega, h) = \theta_i(\omega_{S_{\beta_i(\omega|S(h))}})$. When the state is $\omega \in \Omega$, agent i with the truth-telling and elaboration strategy initially reports her "true" payoff type $\theta_i(\omega_{S_{\beta_i(\omega)}})$ as perceive by her at her awareness level $S_{\beta_i(\omega)}$. At later information sets, $h_i \in H_i(S_\omega)$, her awareness level may be raised to $S_{\beta_i(\omega|S(h_i))} \succeq S_{\beta_i(\omega)}$. Consequently, she reports a more "elaborate" description of her true payoff type, $\theta_i(\omega_{S_{\beta_i(\omega|S(h_i))}})$. Although the agent herself at her type at ω can never anticipate during the game induced by the mechanism the emergency of particular information sets that she is initially unaware of, an information set h_i' may emerge with $S(h_i') \succ S_\omega$. In such a case, agent i's truth-telling and elaboration strategy requires $\sigma_i^*(\omega', h_i') = \theta_i(\omega'_{S_{\beta_i(\omega'|S(h_i'))}})$ such that $r_{S_\omega}^{S_{\omega'}}(\omega') = \omega$ and $(\rho_i)_{S_{\beta_i(\omega)}}^{S_{\beta_i(\omega'|h_i')}}(\theta_i(\omega'_{S_{\beta_i(\omega'|S(h_i'))}})) = \theta_i(\omega_{S_{\beta_i(\omega)}})$. That is, agents only report elaborations of their true type.

Given $\omega \in \Omega$, σ_i allows information set h_i if there exists a profile of i's opponents' strategies σ_{-i} such that $(\sigma_i(\omega), \sigma_{-i}(\omega))$ leads to h_i. Similarly, we say that given $\omega \in \Omega$, σ_{-i} allows information set h_i if there exists a strategy σ_i of agent i such that $(\sigma_i(\omega), \sigma_{-i}(\omega))$ leads to h_i. We denote by $\Sigma_i(\omega, h_i)$ the set of agent i's strategies that given ω allow information set h_i and by $\Sigma_{-i}(\omega, h_i)$ the set of profiles of strategies of agent i's opponents that allow information set h_i given ω.

We denote by $H_i(\omega, \sigma_i)$ the set of information sets of agent i that are allowed by strategy σ_i given ω.

Let $\tau_i : \Omega \times \Sigma_i \times \Sigma_{-i} \longrightarrow \Theta_i$ be the mapping that assigns to each $(\omega, \sigma_i, \sigma_{-i}) \in \Omega \times \Sigma_i \times \Sigma_{-i}$ the payoff type $\tau_i(\omega, \sigma_i, \sigma_{-i})$ of agent i that is reached with (σ_i, σ_{-i}) once the dynamic direct elaboration mechanism described above stops. (We noted already that it must stop since $\mathbf{S}$ is a finite lattice.)

We can have that $\tau_i(\omega, \sigma_i, \sigma_{-i})$ reaches a type that at ω agent i cannot anticipate yet. This is because agent i is unaware of that type but her awareness is raised of it at some point during the dynamic direct elaboration mechanism. For ω and appropriate h_i, the type anticipated by agent i from strategies (σ_i, σ_{-i}) is $\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})$. Similarly, $\tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})$ is the type profile of agent i's opponents that agent i anticipates from strategies (σ_i, σ_{-i}) at ω and h_i.

We say that the dynamic direct mechanism *truthfully implements the social choice* *function f in conditional dominant strategies* if for all $i \in I$, $\omega \in \Omega$, $h_i \in H_i(\omega, \sigma_i^*)$,

$$\begin{aligned} u_i(f(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega|S(h_i))}})) \geq \\ u_i(f(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega|S(h_i))}})) \end{aligned} \quad (4)$$

for all $\sigma_i \in \Sigma_i(\omega, h_i)$ and $\sigma_{-i} \in \Sigma_{-i}(\omega, h_i)$. Conditional dominance takes serious the idea that given a state, a history has been reached and that conditional on this history and state, the strategy very weakly dominates any other as far as the agent is able to anticipate at that state and history. Unconditional dominance is implied since it applies to all histories including the empty history. Conditional dominance is important because it allows for conditioning on the agent's awareness at various information sets.

A social choice function is *truthfully implemented under pooled awareness* if for any $\omega \in \Omega$, outcome $f((\theta_j(\omega_{\bigvee_{i \in I} S_{\beta_i(\omega)}}))_{j \in I})$ is implemented. Pooled awareness means that the awareness of all agents is joint together. Each agent's awareness at ω is given by $S_{\beta_i(\omega)}$. When awareness of all agents is pooled, then it is represented by the join, $\bigvee_{i \in I} S_{\beta_i(\omega)}$, which exists since $\mathbf{S}$ is a complete lattice. Pooled awareness does not mean full awareness as there might exist events of which no agent is aware of. Yet, pooled awareness is the highest awareness level one can hope for in such settings unless explicit investment in discovery is modeled as well.

If the mediator/mechanism designer has an awareness level higher than the meet of the lattice, then it would be straightforward to consider his awareness in the join above. In the dynamic elaboration mechanism, the designer would then report back to agents the pooled awareness level including his own awareness. However, as most of the literature on mechanism design, we do not model the mechanism designer as a player.

In the next sections, we consider special subclasses of dynamic direct elaboration mechanisms that are distinguished by their transfer functions.

7 Dynamic Elaboration Groves Mechanism

Consider for each agent $i \in I$ a transfer functions f_i defined as follows: Denote by $t_i^* \in \Theta_i$ and $t_{-i}^* \in \Theta_{-i}$ the final reports of agent i and i's opponents in the dynamic direct elaboration mechanism and by t_{-i}^1 the initial reports of opponents of agent i. The transfer function of agent i is given by

$$f_i(t_i^*, t_{-i}^*, t_{-i}^1) := \sum_{j \neq i} v_j(f_0(t_i^*, t_{-i}^*), t_j^*) + c_i(t_{-i}^1) \quad (5)$$

for some function c_i of *initial* reports of agent i's opponents. Mechanisms with transfer functions given by equation (5) can be viewed as a dynamic version of the Groves mechanism (Groves, 1973). That's why we call it the dynamic elaboration Groves mechanism.

Theorem 2 *If f is utilitarian ex post efficient, then f is truthfully implementable under pooled awareness in conditional dominant strategies in a dynamic elaboration Groves mechanism.*

Proof. Let f_0 be an utilitarian ex post efficient social choice function. Consider the dynamic direct elaboration mechanism implementing the outcome function f_0 and transfer functions $(f_i)_{i \in I}$ defined in equation (5).

We need to show inequality (4). Using equation (5), we can rewrite inequality (4) as for all $i \in I$, $\omega \in \Omega$, $h_i \in H_i(\omega, \sigma_i^*)$,

$$\begin{aligned} & v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega|S(h_i))}})) \\ & \quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i})), \tau_j(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i})) \\ & \quad \geq v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega|S(h_i))}})) \\ & \quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})), \tau_j(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})) \quad (6) \end{aligned}$$

for all $\sigma_i \in \Sigma_i(\omega, h_i)$ and $\sigma_{-i} \in \Sigma_{-i}(\omega, h_i)$. (Note that the $c_i(\cdot)$ term cancels out.)

Note that both sides of the inequality (6) represent utilitarian ex post welfare.

Observe that for all $i \in I$, $\omega \in \Omega$, $h_i \in H_i(\omega, \sigma_i^*)$, $\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}) = \theta_i(\omega_{S_{\beta_i(\omega|S(h_i))}})$ for all $\sigma_{-i} \in \Sigma_{-i}(\omega, h_i)$. This follows from the fact that σ_i^* is the truth-telling and elaboration strategy and that given ω and history h_i, agent i cannot anticipate to become aware of a particular space more expressive than $S_{\beta_i(\omega|S(h_i))}$. Observe also that for all $\omega \in \Omega$, $h_i \in H_i(\omega, \sigma_i^*)$, $\tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}) \in T_{-i}^{S_{\beta_i(\omega|S(h_i))}}$ for all $\sigma_{-i} \in \Sigma_{-i}(\omega, h_i)$. Again, this follows from σ_i^* being the truth-telling and elaboration strategy and that given ω and history h_i, agent i cannot anticipate that other agents being able to report types in a particular space more expressive than $S_{\beta_i(\omega|S(h_i))}$. Next, observe that for all $\omega \in \Omega$, $h_i \in H_i(\omega, \sigma_i^*)$, $j \neq i$, $\tau_j(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}) = (\rho_j)^{S_{\beta_i(\omega|S(h_i))}}_{S_{\beta_j(\omega_{S_{\beta_i(\omega|S(h_i))}}|S(h_i))}}(\tau_j(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}))$ for all $\sigma_{-i} \in \Sigma_{-i}(\omega, h_i)$. This follows from the fact that at stages $n > 1$ of the dynamic direct mechanism, agent j can only elaborate on her type reported in $n = 1$ but not report any other type inconsistent with her initially reported type. Thus, agent i's report affects the space in which agent j elaborates her previously reported type. Moreover, given ω and h_i, agent i cannot anticipate that agent j reports at a particular awareness level higher than $S_{\beta_i(\omega|S(h_i))}$. With these observations, inequality (6) follows now from f_0 being utilitarian ex post efficient (Remark 1).

Given that each agent reports her awareness truthfully in stage $n = 1$ and only elaborates her type in stage $n = 2$, the mechanism concludes after at most two stages. □

The reason for restricting the c_i functions to *initial* reports of agent i's opponents is that if they were to depend on later reports as well, then they would not cancel out in inequality (6). This is because agent i's report contains awareness that through the dynamic elaboration mechanism gets broadcasted to other agents who subsequently can revise their reports. When agent i deviates to another strategy (left-hand side of inequality (6)), then it could also influence the c_i term if this term can depend on later reports by agent i's opponents. This might destroy incentives to report truthfully.

The differences in transfers anticipated by agent i at ω and h_i given $\sigma_{-i} \in \Sigma_{-i}(\omega, h_i)$ from deviating from a truth-telling and elaboration strategy σ_i^* with a strategy $\sigma_i \in \Sigma_i(\omega, h_i)$ is

$$\sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i})), \tau_j(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i})) \\ - \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})), \tau_j(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i, \sigma_{-i})) \quad (7)$$

That is, the effect of deviation from the truth-telling and elaboration strategy by agent i on her transfers is exactly the externality she imposes on other agents with the deviation. Compared to the standard Groves mechanism, it is not just the externality on the outcome implemented but also the externality that agent i has other agents' ability to become aware of more events through the dynamic direct elaboration mechanism.

Latter feature that agent i affects other agents' valuations through the communication of awareness within the mechanism allows for interdependence of valuations. We know from Jehiel and Moldovanu (2001) that it is impossible to implement ex post efficient social choice functions in settings with interdependent valuations. Thus, it is surprising that we can implement efficiently. The reason for this possibility are two features of our unawareness setting: First, when agents misreport awareness, they can only lie towards less awareness, a feature that played already a role for the revelation principle under unawareness. Second, our definition of utilitarian ex post efficiency implies utilitarian ex post isotony in awareness, i.e., more awareness is better for society. Thus, in a Groves mechanism in which the agent's utility is essentially utilitarian social welfare, the agent has no incentives to misreport with less awareness as this can only affect other agents' valuations and thus utilitarian social welfare through less pooled awareness.

The proof of Theorem 2 reveals that for each agent, there is no reason not to reveal all her awareness with her first report and then only elaborate on her previously reported type in a second round. That is, implementation is possible within at most two steps.

Remark 2 *If f is utilitarian ex post efficient, then there is an outcome in conditional dominant strategies of the dynamic direct elaboration mechanism truthfully implementing f under pooled awareness in at most two stages.*

The reason for the qualification "at most two stages" is that in the special case of initial common awareness among all agents, the setting is similar to the standard setting of mechanism design and just one stage is sufficient. Of course, there are other conditional dominant strategies that could take more than two steps. There is no penalty for revealing later than earlier. Just outcomes matter. Instead, if utilities were discounted from stage to stage, then agents would have a strict incentive to pool awareness asap. However, since with non-trivial unawareness, it takes also at least two stages, some utilitarian welfare would inevitably be destroyed and efficient implementation under pooled awareness would be impossible.

Ideally, we would like our efficient mechanism to satisfy further properties such as participation constraints and budget balance. We say that a dynamic elaboration Groves mechanism with transfer functions $(f_i)_{i \in I}$ (defined in equation (5)) and an utilitarian ex post efficient outcome function f_0 satisfies *participation constraints* if for any agent $i \in I$, $\omega \in \Omega$, $h_i \in H_i(\omega, \sigma_i^*)$, and $h_j^1 \in H_j(\omega, \cdot)$ for $j \neq i$,

$$\begin{aligned} & v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}^*)), \theta_i(\omega_{S_{\beta_i(\omega|S(h_i))}})) \\ & \quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}^*)), \tau_j(\omega_{S_{\beta_i(\omega|S(h_i))}}, \sigma_i^*, \sigma_{-i}^*)) \quad (8) \\ & \quad + c_i\left(\left(\sigma_j^*(\omega_{S_{\beta_j(\omega)}}, h_j^1)\right)_{j \neq i}\right) \geq 0. \quad (9) \end{aligned}$$

This definition presumes an outside option that has a value of zero. Participation constraints could be satisfied for some choices for c_i and violated for others. The problem is that the choice of c_i is not allowed to depend on the ex post reports. Making the c_i term zero, will always satisfy participation constraints but makes the mechanism typically very costly to run.

Budget Balance

The dynamic elaboration Groves mechanism pays each agent the utilitarian ex post welfare under pooled awareness modulo a term that depends only on the other agents' initial reports. Depending on the size of this latter second term, running the dynamic elaboration Groves mechanism could require substantial subsidies by the mechanism designer. We say that a dynamic elaboration Groves mechanism with transfer functions $(f_i)_{i \in I}$ (defined in equation (5)) is ex post budget balanced if for all $S \in \mathcal{S}$ and $\boldsymbol{t} = (t_j)_{j \in I} \in \times_{j \in I} T_j^S$,

$$\sum_{i \in I} f_i \left(\boldsymbol{t}, \left(\left(\rho_{S_j}^S \right)_j (t_j) \right)_{j \in I \setminus \{i\}} \right) = 0 \quad (10)$$

for all $(S_j)_{j \in I}$ with $S_j \in \mathbf{S}(S)$ for all $j \in I$ and $\bigvee_{j \in I} S_j = S$. To understand the notation, recall that the transfer function f_i is a function of the final reports $\boldsymbol{t}$ and the initial opponents' reports t_{-i}^1. Since in the dynamic elaboration Groves mechanism, agents can only elaborate on prior reports, the initial reports must be projections of final reports. This explains f_i's second argument, $\left((\rho_{S_j}^S)_j(t_j)\right)_{j \in I \setminus \{i\}}$.

We are interested in conditions under which budget balance could be achieved. The following observation extends a result for standard Groves mechanisms due to Holmström (1977, Chapter 3.2.3) to our dynamic elaboration Groves mechanism under unawareness. The proof is an extension of the proof in Börgers (2015, Proposition 7.10).

Proposition 2 *Let f be a utilitarian ex post efficient social choice function. Then the dynamic elaboration Groves mechanism implements f with budget balance if and only if* *for every agent $i \in I$, there exists a function $g_i : \Theta_{-i} \longrightarrow \mathbb{R}$ such that for all $S \in \mathcal{S}$ and $\boldsymbol{t} = (t_j)_{j\in I} \in \times_{j\in I} T_j^S$*

$$\sum_{i\in I} v_i(f_0(\boldsymbol{t}), t_i) = -\sum_{i\in I} g_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right) \tag{11}$$

for all $(S_j)_{j\in I}$ with $S_j \in \mathbf{S}(S)$ for all $j \in I$ and $\bigvee_{j\in I} S_j = S$.

PROOF. "Only if": The proof is constructive. Using equations (5) and (10), the dynamic elaboration Groves mechanism is budget balanced if for all $S \in \mathcal{S}$ and $\boldsymbol{t} = (t_j)_{j\in I} \in \times_{j\in I} T_j^S$,

$$\sum_{i\in I}\left(\sum_{j\neq i} v_j(f_0(\boldsymbol{t}), t_j) + c_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right)\right) = 0 \tag{12}$$

for all $(S_j)_{j\in I}$ with $S_j \in \mathbf{S}(S)$ for all $j \in I$ and $\bigvee_{j\in I} S_j = S$. This is equivalent to

$$(|I|-1)\sum_{i\in I} v_i(f_0(\boldsymbol{t}), t_i) = \sum_{i\in I} c_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right)$$

$$\sum_{i\in I} v_i(f_0(\boldsymbol{t}), t_i) = \sum_{i\in I} \frac{c_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right)}{|I|-1}.$$

Set

$$g_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right) := -\frac{c_i\left(\left((\rho^S_{S_j})_j (t_j)\right)_{j\in I\setminus\{i\}}\right)}{|I|-1},$$

obtaining equation (11). This shows necessity.

"if": Assume equation (11) and define

$$c_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right) := -(|I|-1)g_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right).$$

Then

$$\sum_{i\in I}\left(\sum_{j\neq i} v_j(f_0(\boldsymbol{t}), t_j) - (|I|-1)g_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right)\right)$$

$$= (|I|-1)\sum_{i\in I} v_i(f_0(\boldsymbol{t}), t_i) - (|I|-1)g_i\left(\left(\left(\rho^S_{S_j}\right)_j (t_j)\right)_{j\in I\setminus\{i\}}\right) = 0.$$

This completes the proof of the proposition. □

Corollary 1 *If f is a utilitarian ex post efficient social choice function that is implemented in conditional dominant strategies in the dynamic elaboration Groves mechanism under pooled awareness with budget balance, then f is utilitarian ex post welfare constant in awareness. That is, for every $S, S' \in \boldsymbol{S}$ with $S' \preceq S$ and $\boldsymbol{t} = (t_i)_{i \in I} \in \times_{i \in I} T_i^S$,*

$$\sum_{i \in I} v_i(f_0(\boldsymbol{t}), t_i) = \sum_{i \in I} v_i(f_0\left(\rho_{S'}^S(\boldsymbol{t})\right), \left(\rho_{S'}^S\right)_i(t_i)).$$

This impossibility result does not say that there could not be cases in which budget balance is achieved. It just says that budget balance can not always be achieved if the social choice function is not utilitarian ex post welfare constant in awareness. Budget balance could be achieved for instance in settings in which the payoff type of an agent is known, he has the lowest awareness, and her valuation is constant in awareness.

8 Dynamic Elaboration Clarke Mechanism

In this section, we weaken budget balance to no deficit. That is, we allow the mechanism to run a surplus but not a deficit. To study when we can implement efficiently with no deficit, we consider a version of the Clarke or pivot mechanism (Clarke, 1971). First, for any agent $i \in I$, let f_0^{-i} be defined by for all $S \in \mathcal{S}$ and $t_{-i} \in \Theta_{-i}^S$,

$$\sum_{j \neq i} v_j(f_0^{-i}(t_{-i}), t_j) \geq \sum_{j \neq i} v_j(x_0, t_j) \tag{13}$$

for all $x_0 \in X_0^S$. That is, $f_0^{-i}(t_{-i})$ is an utilitarian ex post efficient outcome in X_0 when agent i does not exist and t_{-i} is the profile of payoff types of all agents but i.

Consider now a direct dynamic elaboration mechanism with transfer functions given by for any agent i, any initial reports $(t_i^1, t_{-i}^1) \in \Theta$, and any final reports $(t_i^*, t_{-i}^*) \in \Theta$,

$$f_i(t_i^*, t_{-i}^*, t_i^1, t_{-i}^1) \;:=\; \sum_{j \neq i} v_j(f_0(t_i^*, t_{-i}^*), t_j^*) + c_i(t_{-i}^*) + a_i(t_i^1, t_{-i}^1) \tag{14}$$

with

$$c_i(t_{-i}^*) \;:=\; -\sum_{j \neq i} v_j(f_0^{-i}(t_{-i}^*), t_j^*) \tag{15}$$

and

$$a_i(t_i^1, t_{-i}^1) \;:=\; \begin{cases} r_i(S([t_i^1])) & \text{if } S([t_i^1]) \succ S([t_j^1]), j \neq i \\ -\frac{1}{|I|-1} r_j(S([t_j^1])) & \text{if } S([t_j^1]) \succ S([t_k^1]), k \neq j \neq i \\ 0 \text{ otherwise} & \end{cases} \tag{16}$$

where $r_i(S)$ is defined inductively as follows:

$$r_i(\underline{S}) \;:=\; 0 \tag{17}$$

and for any $S \in \mathcal{S}$ with $S \succ \underline{S}$,

$$r_i(S) \ := \ \max_{S' \prec S, \tilde{t}'_{-i} \in T^{S'}_{-i}, \tilde{t}_{-i} \in T^S_{-i}} r_i(S') + \sum_{j \neq i} v_j(f_0^{-i}(\tilde{t}_{-i}), \tilde{t}_j) - \sum_{j \neq i} v_j(f_0^{-i}(\tilde{t}'_{-i}), \tilde{t}'_j). \quad (18)$$

To understand these transfers, first, note that $c_i(t^*_{-i})$ as defined in equation (15) just like in a standard Clarke mechanism. Together with the first term in equation (14), it makes the mechanism into a pivot mechanism. The issue of course is that in the dynamic direct elaboration mechanism, $c_i(t^*_{-i})$ depends indirectly also on the i's initial report t^1_i because player i can manipulate other's awareness. This might destroy agent i's incentive compatibility of reporting her full awareness in the first stage. This was also the motivation for letting c_i only depend on the other agents' *initial* reports only in the dynamic elaboration Groves mechanism of the prior section. For the dynamic elaboration Clarke mechanism, we take another route to restore incentive compatible reporting of initial awareness. We counter it with an additional adjustment term $a_i(t^1_i, t^1_{-i})$ that depends only on the agents' initial reports. In fact, inspection of equation (16) reveals that it only depends on the awareness levels of agents associated with their first reports. Roughly, a_i rewards agent i if initially she reports more awareness than other agent. It penalizes agent i if someone else initially reports more awareness. Otherwise, the adjustment term is zero.

We call a mechanism with transfer functions defined by equation (14) as dynamic elaboration Clarke mechanism.

Recall that we use the convention that payments are transfers *to* agents from the mechanism designer. We say that a dynamic elaboration Clarke mechanism runs *no deficit* if for all $S \in \mathcal{S}$ and $\boldsymbol{t} = (t_j)_{j \in I} \in \times_{j \in I} T^S_j$,

$$\sum_{i \in I} f_i \left(\boldsymbol{t}, \left(\left(\rho^S_{S_j} \right)_j (t_j) \right)_{j \in I} \right) \leq 0 \quad (19)$$

for all $(S_j)_{j \in I}$ with $S_j \in \mathbf{S}(S)$ for all $j \in I$ and $\bigvee_{j \in I} S_j = S$. To understand the notation, recall that the transfer function f_i is a function of the final reports $\boldsymbol{t}$ and the initial opponents' reports t^1_{-i}. Since in the dynamic elaboration mechanism, agents can only elaborate on prior reports, the initial reports must be projections of final reports. This explains f_i's second argument, $\left((\rho^S_{S_j})_j (t_j) \right)_{j \in I}$. No deficit weakens ex post budget balance by allowing for a surplus but not deficits.

To implement with no deficits, we need an assumption on the benefit function. This assumption is standard in the sense that an analogous assumption would be required for implementation with no deficit under full awareness.

Assumption 1 *For all $i \in I$, we require $v_i : \bigcup_{S \in \mathbf{S}} X^S_0 \times T^S_i \longrightarrow \mathbb{R}_+$ to be a nonnegative function.*

Our prior result on the dynamic elaboration Groves mechanism does not imply that also the dynamic elaboration Clarke mechanism can implement any utilitarian ex post efficient social choice function. This is because in our context, the dynamic elaboration Clarke mechanism is *not* a special case of the dynamic elaboration Groves mechanism. This is due to the fact that the dynamic elaboration Groves mechanism avoids the destruction of incentive compatible initial reporting of full awareness by letting the constant term in the transfers only depend on initial reports while the dynamic elaboration Clarke mechanism overwrites the destruction by adding an additional awareness dependent adjustment term.

Theorem 3 *Given Assumption 1, if f is a utilitarian ex post efficient social choice function, then f is truthfully implementable under pooled awareness in conditional dominant strategies in the dynamic elaboration Clarke mechanism that satisfies no deficit.*

Proof. Fix a state $\omega \in \Omega$. In a game with unawareness in extensive form, information set of players can be ordered by a precedence relation due to perfect recall (Heifetz, Meier, and Schipper, 2013b). Given ω, there is a unique first information set for each player. Any strategy profile must trivially reach the first information set of the player. Recall also that for any $i \in I$, $\beta_i(\omega) = \beta_i(\omega \mid S(h_i^1))$ for the first information set $h_i^1 \in H_i(\omega, \cdot)$ by definition.

We now consider two cases of possible deviations from the truth-telling strategy.

Case 1 (Deviations at the same awareness level): Suppose agent i deviates from the truth-telling strategy with strategy σ_i such at some $\omega \in \Omega$ at her first information set $h_i^1 \in H_i(\omega, \cdot)$, $\sigma_i\left(\omega_{S_{\beta_i(\omega)}}, h_i^1\right) = t_i \in T_i^{S_{\beta_i(\omega)}}$. That is, she reports a payoff type at the same awareness level as her type at ω.

We claim that she has no incentive to do so on matter what opponent's do and no matter which state occurs that she considers possible at this initial information set. That is, we claim

$$\begin{aligned}
& v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) + a_i\left(\sigma_i^*(\omega_{S_{\beta_i(\omega)}}, h_i^1), \left(\sigma_j\left(\omega'_{S_{\beta_j(\omega')}}, h_j^1\right)_{j \in I \setminus \{i\}}\right)\right) \geq \\
& v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})) + a_i\left(\sigma_i(\omega_{S_{\beta_i(\omega)}}, h_i^1), \left(\sigma_j\left(\omega'_{S_{\beta_j(\omega')}}, h_j^1\right)_{j \in I \setminus \{i\}}\right)\right) \qquad (20)
\end{aligned}$$

for each $(h_j^1)_{j \in I \setminus \{i\}} \in \times_{j \in I \setminus \{i\}} H_j(\omega', \cdot)$ and all $\omega' \in$ *supp* $\beta_i(\omega)$.

Since τ_{-i} depends on σ_i only through the awareness raised by agent i with σ_i, we have in this case that

$$c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) = c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})).$$

That is, these terms cancel out of inequality (20).

Next, observe that

$$a_i\left(\sigma_i^*(\omega_{S_{\beta_i(\omega)}}, h_i^1), \left(\sigma_j\left(\omega'_{S_{\beta_j(\omega')}}, h_j^1\right)_{j\in I\setminus\{i\}}\right)\right) = a_i\left(\sigma_i(\omega_{S_{\beta_i(\omega)}}, h_i^1), \left(\sigma_j\left(\omega'_{S_{\beta_j(\omega')}}, h_j^1\right)_{j\in I\setminus\{i\}}\right)\right)$$

for each $(h_j^1)_{j\in I\setminus\{i\}} \in \times_{j\in I\setminus\{i\}} H_j(\omega', \cdot)$ and all $\omega' \in$ *supp* $\beta_i(\omega)$. To understand these terms, note that agent i may be uncertain which first information set agent j reached since i is incompletely informed about the state ω. She considers all states in *supp* $\beta_i(\omega)$ possible. However, in the current case we are considering, her deviation strategy σ_i raises the same awareness as her truth-telling strategy at agent i's initial information set. Thus, by construction of the awareness adjustment function a_i, both terms must be equal no matter whether other agents raise more or less awareness. Hence, those terms cancel out in inequality (20). We are left with just the first two terms on each side. The claim now follows from the fact that f is utilitarian ex post efficient.

Case 2 (Deviations to lower awareness levels): Suppose agent i deviates from the truth-telling strategy with strategy σ_i such that at some $\omega \in \Omega$ at her first information set $h_i^1 \in H_i(\omega, \cdot)$, $\sigma_i\left(\omega_{S_{\beta_i(\omega)}}, h_i^1\right) = t_i \in T_i^S$ with $S \prec S_{\beta_i(\omega)}$. That is, she reports a payoff type at an awareness level lower than her type at ω. Two subcases can occur.

Subcase 2a (Her initially reported awareness level is higher than anyone's else): This is the case such that for all $j \neq i$ the information set $h_j^1 \in H_j(\omega, \cdot)$ is such that $S([\sigma_i(\omega, h_i^1]) \succ S([\sigma_j(\omega, h_j^1)])$. Thus, the awareness adjustment term in agent i's payoff is $r_i(S([\sigma_i(\omega, h_i^1]))$.

We claim that agent i has no incentive to not report her initial awareness level. That is,

$$\begin{aligned}
&v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + \sum_{j\neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) + r_i(S([\sigma_i^*(\omega, h_i^1])) \geq \\
&\quad v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + \sum_{j\neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})) + r_i(S([\sigma_i(\omega, h_i^1])) \qquad (21)
\end{aligned}$$

To see this, consider the right-hand side of inequality (21). Using the definition of c_i, we claim that we can bound the right-hand side from above by

$$\begin{aligned}
& v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
&\quad - \sum_{j \neq i} v_j(f_0^{-i}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) + r_i(S([\sigma_i^*(\omega, h_i^1]])) \geq \\
&\quad v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
&\quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
&\quad - \sum_{j \neq i} v_j(f_0^{-i}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) + r_i(S([\sigma_i(\omega, h_i^1]])) \qquad (22)
\end{aligned}$$

The first two terms on both sides cancel each other out. Note that σ_i^* is agent i's truth-telling strategy, $S([\sigma_i(\omega, h_i^1]]) = S_{\beta_i(\omega)}$. Thus, previous inequality reduces to

$$\begin{aligned}
& - \sum_{j \neq i} v_j(f_0^{-i}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) + r_i(S_{\beta_i(\omega)}) \geq \\
&\quad - \sum_{j \neq i} v_j(f_0^{-i}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) + r_i(S([\sigma_i(\omega, h_i^1]])) \qquad (23)
\end{aligned}$$

Finally, observe that by definition of r_i,

$$\begin{aligned}
& r_i(S_{\beta_i(\omega)}) = \\
& \qquad \max_{S' \prec S_{\beta_i(\omega)}, t'_{-i} \in T_{-i}^{S'}, t_{-i} \in T_{-i}^{S_{\beta_i(\omega)}}} \left(r_i(S') + \sum_{j \neq i} v_j(f_0^{-i}(t_{-i}), t_j) - \sum_{j \neq i} v_j(f_0^{-i}(t'_{-i}), t'_j) \right) \\
& \geq \; r_i(S([\sigma_i(\omega, h_i^1]])) + \\
& \qquad \max_{t''_{-i} \in T_{-i}^{S([\sigma_i(\omega, h_i^1)])}, t_{-i} \in T_{-i}^{S_{\beta_i(\omega)}}} \left(\sum_{j \neq i} v_j(f_0^{-i}(t_{-i}), t_{-i}) - \sum_{j \neq i} v_j(f_0^{-i}(t'_{-i}), t'_{-i}) \right) \\
& \geq \; r_i(S([\sigma_i(\omega, h_i^1]])) + \sum_{j \neq i} v_j(f_0^{-i}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
& \qquad - \sum_{j \neq i} v_j(f_0^{-i}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}}))
\end{aligned}$$

Therefore, inequality (23) and thus inequality (23) follow.

As a last step, note that the left-hand-side of inequality (21) is weakly larger than the left-hand-side of inequality (22). That is, we claim

$$
\begin{aligned}
& v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) + r_i(S([\sigma_i^*(\omega, h_i^1])) \geq \\
& \quad v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_j(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) + r_i(S([\sigma_i^*(\omega, h_i^1])) \qquad (24)
\end{aligned}
$$

Both the c_i and r_i terms cancel each other out and we are left with ex post utilitarian welfare expressions on both side. The inequality now simply follows from analogous arguments as in the proof of Theorem 2. This completes the proof of this subcase.

Subcase 2b (Her initially reported awareness level is lower than anyone's else): This is the case such that there is an agent $j \neq i$ with an information set $h_j^1 \in H_j(\omega, \cdot)$ such that $S([\sigma_j(\omega, h_j^1]) \succ S([\sigma_k(\omega, h_k^1)])$ for $k \neq j$, $h_k^1 \in H_k(\omega, \cdot)$. Thus, the awareness adjustment term in agent i's payoff is $-\frac{1}{|I|-1} r_j(S([\sigma_j(\omega, h_j^1]))$ for some j.

Assume that agent i is not aware of her unawareness and cannot anticipate that j reports an awareness level $S \not\preceq S_{\beta_i(\omega)}$. Recall that unawareness of an event implies that she is also unaware that others are aware of the event (see Heifetz, Meier, and Schipper, 2013a). Thus, in this case $S([\sigma_i^*(\omega, h_i^1]) \succeq S([\sigma_j(\omega, h_j^1)])$.

We claim that agent i has no incentive not to report her initial awareness level. That is,

$$
\begin{aligned}
& v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + \sum_{k \neq i} v_k(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \theta_k(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) + \mathbb{I}_{S([\sigma_i^*(\omega, h_i^1]) \succ S([\sigma_j(\omega, h_j^1])} r_i(S([\sigma_i^*(\omega, h_i^1)])) \geq \\
& \quad v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + \sum_{k \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \theta_k(\omega_{S_{\beta_i(\omega)}})) \\
& \quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})) - \frac{1}{|I|-1} r_j(S([\sigma_j(\omega, h_j^1)])) \qquad (25)
\end{aligned}
$$

To see this, note first that the first two terms on each side just represent ex post utilitarian welfare. Thus, the fact that first two terms of the left-hand side are weakly larger than the first two terms of the right-hand side follows from analogous arguments as in the proof of Theorem 2. We are left to show that

$$c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) + \mathbb{I}_{S([\sigma_i^*(\omega,h_i^1)]) \succ S([\sigma_j(\omega,h_j^1)])} r_i(S([\sigma_i^*(\omega, h_i^1)])) \geq \\ c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})) - \frac{1}{|I|-1} r_j(S([\sigma_j(\omega, h_j^1)])) \tag{26}$$

which is equivalent to

$$-\sum_{k \neq i} v_k(f_0^{-1}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) + \mathbb{I}_{S([\sigma_i^*(\omega,h_i^1)]) \succ S([\sigma_j(\omega,h_j^1)])} r_i(S([\sigma_i^*(\omega, h_i^1])) \geq \\ -\sum_{k \neq i} v_k(f_0^{-1}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})) - \frac{1}{|I|-1} r_j(S([\sigma_j(\omega, h_j^1)])) \tag{27}$$

Note that by the construction of r_i,

$$\begin{aligned} & r_i(S([\sigma_i^*(\omega, h_i^1)])) = \\ & \qquad \max_{S \prec S([\sigma_i^*(\omega,h_i^1)]), t_{-i} \in T_{-i}^S, t_{-i}^* \in T_{-i}^{S([\sigma_i^*(\omega,h_i^1)])}} \left(r_i(S) + \sum_{k \neq i} v_k(f_0^{-i}(t_{-i}^*), t_k^*) - \sum_{k \neq i} v_k(f_0^{-i}(t_{-i}), t_k) \right) \\ & \geq \max_{t_{-i} \in T_{-i}^{S([\sigma_j(\omega,h_j^1)])}, t_{-i}^* \in T_{-i}^{S([\sigma_i^*(\omega,h_i^1)])}} \left(\sum_{k \neq i} v_k(f_0^{-i}(t_{-i}^*), t_k^*) - \sum_{k \neq i} v_k(f_0^{-i}(t_{-i}), t_k) \right) \\ & \geq -\frac{1}{|I|-1} r_j(S([\sigma_j(\omega, h_j^1)])) \\ & \quad + \sum_{k \neq i} v_k(f_0^{-1}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) \\ & \quad - \sum_{k \neq i} v_k(f_0^{-1}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})) \end{aligned}$$

If $\mathbb{I}_{S([\sigma_i^*(\omega,h_i^1)]) \succ S([\sigma_j(\omega,h_j^1)])} = 1$, then inequality (27) follows from the above sequence of inequalities. Otherwise, $\mathbb{I}_{S([\sigma_i^*(\omega,h_i^1)]) \succ S([\sigma_j(\omega,h_j^1)])} = 0$, then inequality (27) follows from

$$\sum_{k \neq i} v_k(f_0^{-1}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})), \tau_k(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i})) = \\ \sum_{k \neq i} v_k(f_0^{-1}(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i})), \tau_k(\omega_{S_{\beta_i(\omega)}}, \sigma_i, \sigma_{-i}))$$

because in this case agent i can only anticipate $S([\sigma_i^*(\omega, h_i^1)]) = S([\sigma_j(\omega, h_j^1)])$.

Now assume that agent i is aware of her unawareness. She entertains the idea that another agent may report higher awareness than herself even though she cannot anticipate of what exactly the other agent would be aware that she is unaware (see Schipper, 2022). We claim that also in this case, the agent has no incentive not to initially report her awareness level. This is because her reported awareness level will not have any effect on first-stage transfers and on elaborations made later by herself or other agents. Thus, she might as well report her awareness level. This completes the proof of subcase 2b.

Note that deviations to higher or incomparable awareness levels are infeasible for the agent as she can only report of what she is aware of. Thus, the two main cases exhaust the set of all cases. Thus, we have shown that if f is a utilitarian ex post efficient social choice function, then f is truthfully implementable under pooled awareness in conditional dominant strategies in the dynamic elaboration Clarke mechanism.

What is left to show is that it satisfies no deficit. That is,

$$\sum_{i \in I} f_i(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \sigma_i^*(\omega, h_i^1), (\sigma_j^*(\omega, h_j^1))_{j \neq i}) \leq 0$$

which is equivalent to

$$\sum_{i \in I} \left(c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \theta_j(\omega_{S_{\beta_i(\omega)}})) \right) \leq - \sum_{i \in I} a_i(\sigma_i^*(\omega, h_i^1), (\sigma_j^*(\omega, h_j^1))_{j \neq i}) \quad (28)$$

Note that by construction,

$$\sum_{i \in I} a_i(\sigma_i^*(\omega, h_i^1), (\sigma_j^*(\omega, h_j^1))_{j \neq i}) = 0.$$

Moreover, for each $i \in I$,

$$-c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) \geq \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \theta_j(\omega_{S_{\beta_i(\omega)}})) \quad (29)$$

because $c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) = -\sum_{i \neq j} v_j(f_0^{-i}(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \theta_j(\omega_{S_{\beta_i(\omega)}}))$ and $f_0^{-i}(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*))$ is a utilitarian ex post efficient outcome among agents in $I \setminus \{i\}$. This completes the proof of the theorem. □

We say that a dynamic elaboration Clarke mechanism with transfer functions $(f_i)_{i \in I}$ (defined in equation (14)) and an utilitarian ex post efficient outcome function f_0 satisfies *ex post participation constraints at initial information sets* if for any agent $i \in I$, $\omega \in \Omega$, $h_i^1 \in H_i(\omega, \cdot)$, and $h_j^1 \in H_j(\omega, \cdot)$,

$$\begin{aligned} & v_i(f_0(\tau_i(\omega_{S_{\beta_i(\omega))}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \theta_i(\omega_{S_{\beta_i(\omega)}})) \\ & \quad + \sum_{j \neq i} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \tau_j(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) \\ & \quad + c_i(\tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) + a_i(\sigma_i^*(\omega, h_i^1), (\sigma_j^*(\omega_{S_{\beta_i(\omega)}}, h_j^1))_{j \neq i}) \geq 0. \end{aligned} \quad (30)$$

These are ex post participation constraints initially anticipated by agents.

We say agent i is *unaware of her unawareness* if for any $\omega \in \Omega$ and agent $j \neq i$ she considers only initial information sets $h_j^1 \in H_j(\omega_{S_{\beta_i(\omega)}}, \cdot)$ of agent j. That is, she does not anticipate that some agent could have higher awareness than her.[6]

Proposition 3 *If the agent is unaware of her unawareness, the dynamic elaboration Clarke mechanism satisfies ex post participation constraints at initial information sets.*

PROOF. Since agent i is unaware of her unawareness, $h_j^1 \in H_j(\omega_{S_{\beta_i(\omega)}}, \cdot)$. Thus, agent i does not anticipate a negative a_i term. Assume the worst case anticipated by i,

$$a_i(\sigma_i^*(\omega, h_i^1), (\sigma_j^*(\omega_{S_{\beta_i(\omega)}}, h_j^1))_{j \neq i}) = 0.$$

Then inequality (30) is equivalent to

$$\sum_{j \in I} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \tau_j(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) \geq$$
$$\sum_{j \neq i} v_j(f_0^{-i}(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \tau_j(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) \quad (31)$$

Suppose by contradiction that this inequality does not hold. Then

$$\sum_{j \in I} v_j(f_0^{-i}(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \tau_j(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) \geq$$
$$\sum_{j \neq i} v_j(f_0^{-i}(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \tau_j(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)) >$$
$$\sum_{j \in I} v_j(f_0(\tau_i(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*), \tau_{-i}(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*)), \tau_j(\omega_{S_{\beta_i(\omega)}}, \sigma_i^*, \sigma_{-i}^*))$$

The first inequality follows from v_i being nonnegative. Together with the last inequality, it implies a contradiction to f_0 being utilitarian ex post efficient. □

At the second period, an agent may realize that ex post participation constraints are violated, namely in the case when her opponents reported higher awareness in the first stage and her a_i term becomes sufficiently negative. Also, if agent i is aware of her unawareness, then she might anticipate a sufficiently negative a_i term in which case her ex post participation constraint might be violated already at the initial history.

...

[6] Although we do not model awareness of unawareness explicitly (since we focus here on unawareness of unawareness), the framework can be extended to explicitly model awareness of unawareness along the lines of Schipper (2022).

References

[1] Auster, S. (2013). Asymmetric awareness and moral hazard, Games and Economic Behavior 82, 503–521.

[2] Auster, S. and N. Pavoni (2021). Optimal delegation and information transmission under limited awareness, University of Bonn and Bocconi University.

[3] Bester, H. and R. Strausz (2001). Contracting with imperfect commitment and the revelation principle: The single agent case, Econometrica 69, 1077–1098.

[4] Börgers, T. (2015). An introduction to the theory of mechanism design, Oxford University Press.

[5] Brandenburger, A. and A. Friedenberg (2011). The relationship between rationality on the matrix and the tree, NYU Stern School of Business.

[6] Bull, J. and J. Watson (2007). Hard evidence and mechanism design, Games and Economic Behavior 58, 75–93.

[7] Clarke, E.H. (1971). Multipart pricing of public goods, Public Choice 8, 19–33.

[8] Deneckere, R. and S. Severinov (2008). Mechanism design with partial state verifiability, Games and Economic Behavior 64, 487–513.

[9] Filiz-Ozbay, E. (2012). Incorporating unawareness into contract theory, Games and Economic Behavior 76, 181–194.

[10] Francetich, A. and B.C. Schipper (2022). Rationalizable screening and disclosure under unawareness, University of Washington, Bothell, and University of California, Davis.

[11] Grant, S., Kline, J., and J. Quiggin (2012). Differential awareness, ambiguity, and incomplete contracts: A model of contractual disputes, Journal of Economic Behavior and Organization 82, 494–504.

[12] Green, J. and J.-J. Laffont (1986). Partially verifiable information and mechanism design, Review of Economic Studies 447–456.

[13] Green, J. and J.-J. Laffont (1979). Incentives in public decision-making, North-Holland.

[14] Green, J. and J.-J. Laffont (1977). Characterization of satisfactory mechanisms for the revelation of preferences for public goods, Econometrica 45, 427–438.

[15] Groves, T. (1973). Incentives in teams, Econometrica 41, 617–631.

[16] Heifetz, A., Meier, M., and B.C. Schipper (2013a). Unawareness, beliefs, and speculative trade, Games and Economic Behavior 77, 100–121.

[17] Heifetz, A., Meier, M., and B.C. Schipper (2013b). Dynamic unawareness and rationalizable behavior, Games and Economic Behavior 81, 50–68.

[18] Heifetz, A., Meier, M., and B.C. Schipper (2006). Interactive unawareness, Journal of Economic Theory 130, 78–94.

[19] Herweg, F. and K. Schmidt (2020). Procurement with unforeseen contingencies, Management Science 66, 2194–2212.

[20] Holmström, B. (1977). On incentives and control in organizations, Dissertation, MIT.

[21] Jehiel, P. and B. Moldovanu (2001). Efficient design with interdependent valuations, Econometrica 69, 1237–1259.

[22] Lee, J. (2008). Unforeseen contingencies and renegotiation with asymmetric information, Economic Journal 118, 678–694.

[23] Lei H. and X. Zhao (2021). Delegation and information disclosure with unforeseen contingencies, B.E. Journal of Theoretical Economics 21, 637–656.

[24] Maskin. E. and J. Tirole (1999). Unforeseen contingencies and incomplete contracts, Review of Economic Studies 66, 83–114.

[25] Meek Lange, M. (2011). Progress, Stanford Encyclopedia of Philosophy, Zalta, E.N. (ed.), `https://plato.stanford.edu/archives/sum2022/entries/progress/`.

[26] Meier, M. and B.C. Schipper (2022). Conditional dominance in games with unawareness, University of California, Davis.

[27] Meier, M. and B.C. Schipper (2014). Bayesian games with unawareness and unawareness perfection, Economic Theory 56, 219–249.

[28] Mill, J.S. (1868). A system of logic. Ratiocinative and inductive, Vol. II. 7th edition, London: Longmans et al. `https://www.gutenberg.org/ebooks/35421`.

[29] Schipper, B.C. (2022). Interactive awareness of unawareness, University of California, Davis.

[30] Schipper, B.C. (2021). Discovery and equilibrium in games with unawareness, Journal of Economic Theory 198, 105365.

[31] Shimoji, M. and J. Watson (1998). Conditional dominance, rationalizability, and game forms, Journal of Economic Theory 83, 161–195.

[32] Strausz, R. (2017). Mechanism design with partially verifiable information, Humboldt University zu Berlin.

[33] von Thadden, E.-L. and Zhao X.J. (2012). Incentives for unaware agents, Review of Economic Studies 79, 1151-1174.

[34] Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders, Journal of